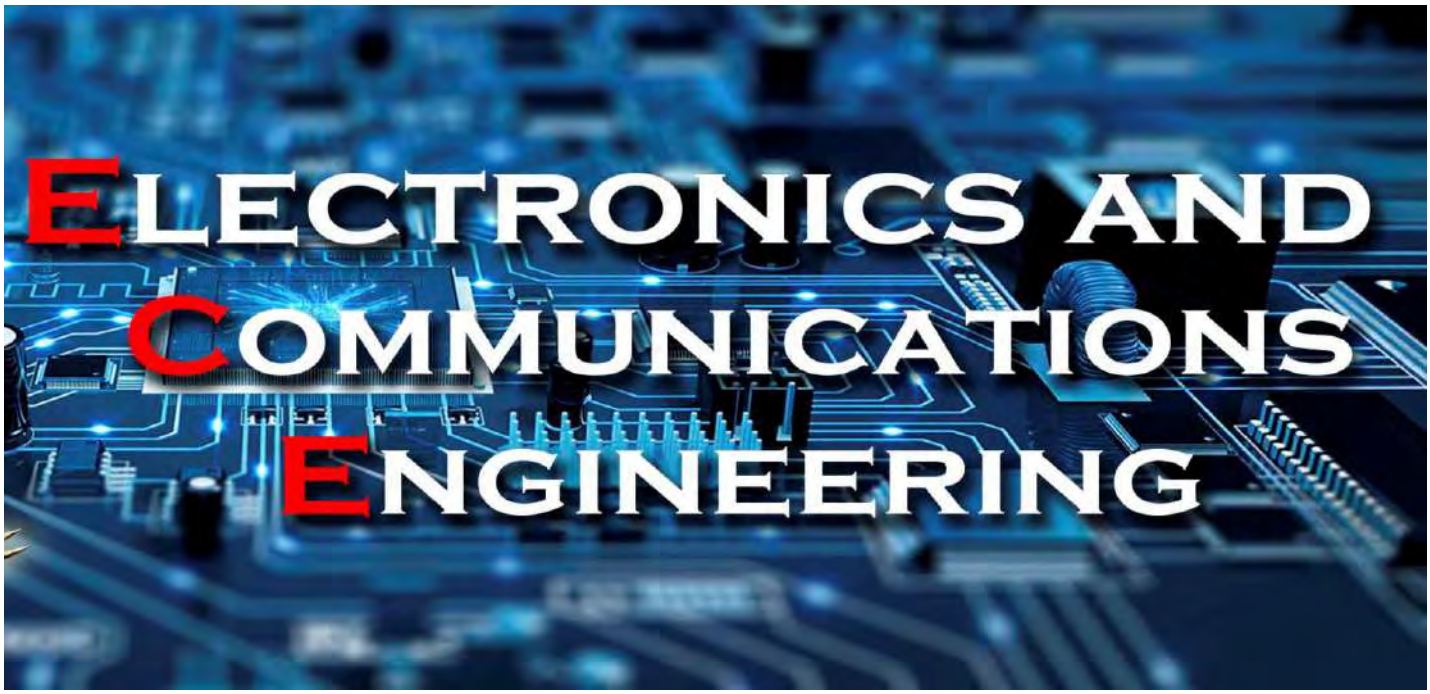
BMS
INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Avalahalli, Doddaballapur Main Road, Bengaluru – 560064

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

RESEARCH COMPENDIUM



ACY: 2022-23

RESEARCH

COMPENDIUM

Compiled by,

Dr. Dankan Gowda V

Dr. G. S Jayadeva
Prof. & HoD, Dept. of ECE

About Institution

In view of the growing demand for technical education and with the goal of establishing a premier technical education on par with international standards, a new technical institution by `name 'BMS Institute of Technology and Management' was established in 2002. Currently, BMSIT & M offers seven UG, three PG programs and Ph.D. /M.Sc. (Engg.) in seven disciplines. BMSIT & M considers research to be of equal importance as academics for the betterment of an institution. Research culture has been embraced well by the faculty members and research scholars at BMSIT and M. In this report, we present an overview of the research activities of Information Science and Engineering, BMSIT & M.

Vision

“To emerge as one of the nation’s finest technical institutions of higher learning to develop engineering professionals who are technically competent, ethical and environment friendly for betterment of the society.”

Mission

“Accomplish stimulating learning environment through high quality academic instruction, innovation and industry – institute interface.”

About Department

The Department of Electronics and Communication Engineering started in 2002-03 with the vision: “To Provide Quality Education in Electronics, Communication and Allied Engineering fields to serve as Valuable Resource for Industry and Society”. The department is a “Research Center” recognized by VTU, has well experienced and qualified faculty members who inspire the students to face the competitive world. 16 faculties are pursuing their Doctoral Degrees. Consistent performance by the students in VTU examinations is a reflection of the efforts by all faculty members. The department is equipped with latest equipment and laboratory amenities to meet the global standards.

Vision

“Provide Quality Education in Electronics, Communication and Allied Engineering fields to Serve as Valuable Resource for Industry and Society”.

Mission

1. Impart Sound Theoretical Concepts and Practical Skills.
2. Promote Inter-Disciplinary Research.
3. Inculcate Professional Ethics.

Programme Educational Objectives

Graduates of the programme will:

PEO1: Work as professionals in the area of Electronics and allied engineering fields

PEO2: Pursue higher studies and involve in the interdisciplinary research work

PEO3: Exhibit ethics, professional skills and leadership qualities in their profession.

Programme Outcomes

1. **Engineering knowledge:** Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization for the solution of complex engineering problems.

2. **Problem analysis:** Identify, formulate, research literature, and analyse complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.

3. **Design/development of solutions:** Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for public health and safety, and cultural, societal, and environmental considerations.

4. **Conduct investigations of complex problems:** The problems that cannot be solved by straightforward application of knowledge, theories and techniques applicable to the engineering discipline, that may not have a unique solution.

5. **Modern tool usage:** Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools, including prediction and modeling to complex engineering activities, with an understanding of the limitations.

6. **The engineer and society:** Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal, and cultural issues and the consequent responsibilities relevant to the professional engineering practice.

7. **Environment and sustainability:** Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.

8. **Ethics:** Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.

9. **Individual and team work:** Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.

10. **Communication:** Communicate effectively on complex engineering activities with the engineering community and with the society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.

11. **Project management and finance**: Demonstrate knowledge and understanding of t h e engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.

12. **Life-long learning:** Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

Programme Specific Outcomes

Graduates will be able to:

PSO1: Exhibit competency in embedded system domain.

PSO2: Exhibit competency in RF and signal processing domain.

International Journal Details

Sl. No.	Title of the Paper	Name of the Author/s	Name of the Journal	Details of publication with doi
1	A Super-resolution-based Approach for the Detection of Covid—19 Infection From Chest X-ray Images	Dr. M. C. Hanumantharaju and Seema S. Bhat	Resonance, 2023, 28(1), pp. 127–148	https://doi.org/10.1007/s12045-023-1530-7
2	FPGA Implementation of Breast Cancer Detection using SVM Linear Classifier	Dr. Lakshmi Sagar. H. S and Dr. M. C. Hanumantharaju	Multimedia Tools and Applications, 2023	https://doi.org/10.1007/s11042-023-15121-6
3	Twofold dynamic attention guided deep network and noise-aware mechanism for image denoising	Dr. Vijayalakshmi G V	Journal of King Saud University - Computer and Information Sciences, 2023, 35(3), pp. 87–102	https://doi.org/10.1016/j.jksuci.2023.02.003
4	X-band Miniaturized frequency reconfigurable travelling wave antenna	Dr. Anitha V.R	International Journal of Electronics Letters, 2023, 11(2), pp. 221–231	DOI: 10.1080/21681724.2022.2068194
5	Optimal design of improved H-infinity controller for MMC	Dr. Rashmi N	International Journal of Electronics, 2023	• https://doi.org/10.1080/00207217.2023.2173807
6	Computer vision based healthcare system for identification of	Dr. Dankan Gowda V	Measurement: Sensors 27	DOI:10.1016/j.measen.2023.100751

	diabetes & its types using AI			
7	A Novel Method of Enhancing Security Solutions and Energy Efficiency of IoT Protocols	Dr. Dankan Gowda V	International Journal on Recent and Innovation Trends in Computing and Communication 11 (4), pp. 325-335	DOI:10.17762/ijritcc.v11i4s.6575
8	Enhancing Microcomputer Edge Computing for Autonomous IoT Motion Control	Dr. Dankan Gowda V	International Journal on Recent and Innovation Trends in Computing and Communication 11 (3), pp. 58-67	DOI:10.17762/ijritcc.v11i3.6202
9	Multimodal Biometric Identification system using Random Selection of Biometrics	Dr. Dankan Gowda V	SSRG International Journal of Electrical and Electronics Engineering 10(1), pp. 63-73	10.14445/23488379/IJEEE-V10I1P106
10	Performance Analysis of Energy Efficiency and Security Solutions of Internet of Things Protocols	Dr. Dankan Gowda V	International Journal of Electrical and Electronics Research 11(2), 442-450	DOI:10.37391/ijeer.110226
11	A novel method of data compression using ROI for biomedical 2D images	Dr. Dankan Gowda V	Measurement: Sensors 24	https://doi.org/10.1016/j.measen.2022.100439
12	Human Emotion Recognition using Deep Learning with Special Emphasis on Infant's Face	Dr. Dankan Gowda V	International Journal of Electrical and Electronics Research 10(4), pp. 1176-1183	DOI:10.37391/ijeer.100466
13	Bilingual text detection from natural scene	Dr. Vijayalakshmi G V	Pattern Analysis and Applications,	https://doi.org/10.1007/s10044-022-01066-3

	images using faster R-CNN and extended histogram of oriented gradients		2022, 25(4), pp. 1001–1013	
14	Design and Implementation of a High-Speed D Flip Flop using CMOS Inverter Logic	Dr. Jayadeva G S	WSEAS Transactions on Electronics, 2022, 13, pp. 125–129	DOI: 10.37394/232017.2022.13.16
15	Design and FPGA Implementation of High Throughput and Low Latency Machine Learning based Approximate Multiplier for Image Processing Applications	Dr. Anil Kumar D	WSEAS Transactions on Systems and Control, 2022, 17, pp. 287–299, 33	DOI: 10.37394/23203.2022.17.33
16	Hybrid optimization model for design and optimization of microstrip patch antenna	Dr. Rashmi N	Transactions on Emerging Telecommunications Technologies, 2022, 33(12), e4640	https://doi.org/10.1002/ett.4640
17	The peak-to-average power ratio reduction using hybrid scheme with companding and discrete Hartley transform for orthogonal frequency division multiplexing system	Dr. Rashmi N	International Journal of Intelligent Enterprisethis link is disabled, 2022, 9(3), pp. 303–317	https://doi.org/10.1504/IJIE.2022.123754
18	Switching, optical, Raman, and Morphological investigations on Sn doped SiTe glasses and thin films	Dr. Jagannatha K B	Thin Solid Films, 2022, 759, 139472	https://doi.org/10.1016/j.tsf.2022.139472

19	Detection of Breast Cancer with Lightweight Deep Neural Networks for Histology Image Classification	Dr. Lakshmi Sagar.H.S and Dr. M. C. Hanumantharaju	Critical Reviews in Biomedical Engineering, 2022, 50(2), pp. 1–19	• DOI: 10.1615/CritRevBiomedEng.2022043417
20	Multi Objective Energy Based Improved Jellyfish Swarm Optimization for Effective Cluster Head Discovery in UWSN	Dr. Ambika R	International Journal of Intelligent Engineering and Systems, May 2022, pp. 509-518	DOI: 10.22266/ijies2023.0630.40
21	Wavelet energy-based adaptive retinex algorithm for low light mobile video enhancement	G. R. Vishalakshi, A. Shobharani & M. C. Hanumantharaju	The Imaging Science Journal	DOI: 10.1080/13682199.2023.2260663
22	Evaluating the Statistical Stability of POSIT Arithmetic and IEEE 754 Float to Accelerate Data for Detection of Breast Cancer	Dr. Lakshmi Sagar.H.S and Dr. M. C. Hanumantharaju	SSRG International Journal of Electrical and Electronics Engineering, 2022, 9(9), pp. 47–53	10.14445/23488379/IJEEE-V9I9P106
23	Probing the bandstructure dependent figures of merit in InAs/GaAs quantum dot photodetectors	Dr. Sabina Rahaman	Micro and Nanostructures, 2022, 168, 207328	https://doi.org/10.1016/j.micrna.2022.207328
24	Parametric Analysis of Compact UWB-MIMO Antenna with Improved	Dr.Anitha V R	Wireless Personal Communications, Vol No. 123, April 2022, pp.2209–2225	DOI:10.1007/s11277-021-09235-z

	Isolation Using Parasitic Reflectors and Protruded Ground Strips			
25	High-gain broadband ENZ metasurface-based RF lens with gradient refractive index for microwave and millimetre-wave applications	Dr.Anitha V R	International Journal of Communication System, Vol No. 35, Issue 17, August 2022,pp.1-13	DOI: 10.1002/dac.5325

International Conference

Sl. No.	Title of the Paper	Name of the Author/s	Name of the conference	ISBN/ISSN Number
1	Analysis of Deep Learning Architecture-Based Classifier for the Cervical Cancer Classification	R. Chandraprabha	Emerging Research in Computing, Information, Communication and Applications. Lecture Notes in Electrical Engineering	https://doi.org/10.1007/978-981-19-5482-5_23
2	Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection	Dr. Anitha V.R	Micro and Nanoelectronics Devices, Circuits and Systems. Lecture Notes in Electrical Engineering	DOI: 10.1007/978-981-19-2308-1_37
3	A Novel IoT Framework and Device Architecture for Efficient Smart city Implementation	Dr. Dankan Gowda V	7th International Conference on Trends in Electronics and Informatics, ICOEI 2023 - Proceedings	DOI: 10.1109/ICOEI56765.2023.10125677
4	IoT Wearable Breast Temperature Assessment System	Dr. Dankan Gowda V	Proceedings - 7th International Conference on Computing Methodologies and Communication, ICCMC 2023	DOI: 10.1109/ICCMC56507.2023.10083511
5	Implementation of a Machine Learning-based Model for Cardiovascular Disease Post	Dr. Dankan Gowda V	2023 International Conference for Advancement in Technology, ICONAT 2023	DOI: 10.1109/ICONAT57137.2023.10080833

	Exposure prophylaxis			
6	Technologies for Comprehensive Information Security in the IoT	Dr. Dankan Gowda V	2023 International Conference for Advancement in Technology, ICONAT 2023	DOI: 10.1109/ICONAT57137.2023.10080332
7	Symmetrized Feature Selection with Stacked Generalization based Machine Learning Algorithm for the Early Diagnosis of Chronic Diseases	Dr. Dankan Gowda V	Proceedings - 5th International Conference on Smart Systems and Inventive Technology, ICSSIT 2023	DOI: 10.1109/ICSSIT55814.2023.10061062
8	Customer Churn Analysis in Financial Domain using Deep Intelligence Network	Dr. Dankan Gowda V	IDCIoT 2023 - International Conference on Intelligent Data Communication Technologies and Internet of Things, Proceedings	DOI: 10.1109/IDCIoT56793.2023.10053473
9	Performance Analysis of 6x6 CNT Field Effect Transistor Array Model for Biomedical Applications	Dr.Anitha V R	International Conference On Composite Science and Technology (COMSAT 2022) during 22nd & 23rd August 2022, Johor, Malaysia	ISBN: 978-629-96647-0-3
10	Design and Development of Light Weight Fractal Tree	Dr.Anitha V R	International Conference On Composite Science and Technology (COMSAT 2022) during 22nd & 23rd August 2022, Johor, Malaysia	ISBN: 978-629-96647-0-3

	Antenna Structure			
11	Cloud Based Face and Speech Recognition for Access Control Applications	Dr.Anitha V R	AICTE-SPICES Sponcered International Conference on Energy Efficient Technologies for Sustainability on 27th & 28th April 2023 in Nagercoil, Kanyakumari.	
12	Performance Analysis of Parametric and Non-parametric Classifier Models for Predicting the Liver Disease	Dr. Vijayalakshmi G V	4th International Conference MISP 2022-Springer Nature Singapore	ISBN 978-981-99-0084-8

Book/Book Chapters

Sl. No.	Title of the Book /Book Chapters Published	Name of the Author/s	ISBN/ISSN Numbers	Publisher name
1	Book chapter: IoT Based Website for Identification of Acute Lymphoblastic Leukemia using DL Book: AI and IoT-based intelligent Health Care & Sanitation	Dr. Ambika R	ISBN: 978-981-5136-54-8 eISBN: 978-981-5136-53-1 (Online)	Bentham Science Publishers
2	An exhaustive analysis of energy harvesting absorbers and battery charging systems for the internet of things	Dr. Dankan Gowda V	ISBN13: 9781668449745 EISBN13: 9781668449769	IGI Global (contributed a chapter)
3	An enhanced method for running embedded applications in a power-efficient manner	Dr. Dankan Gowda V	ISBN13: 9781668449745 EISBN13: 9781668449769	IGI Global (contributed a chapter)
4	Wireless sensor and actuator networks-based reliable data acquisition mechanism	Dr. Dankan Gowda V	ISBN13: 9781668449745 EISBN13: 9781668449769	IGI Global (contributed a chapter)

5	Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future	Dr. Dankan Gowda V	Print ISBN 978-981-19-2467-5 Online ISBN 978-981-19-2468-2	Springer(contributed a chapter)
6	Implementation of a Deep Neural Network- Based Framework for Actigraphy Analysis and Prediction of Schizophrenia	Dr. Vijayalakshmi G V & Prof Chandra Prabha	9781003315452	CRC Press(contributed a chapter)
7	Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks	Dr. Vijayalakshmi G V	ISBN13: 9781668445587	IGI Global (Edited book)
8	Segmentation of optic disc from fundus image based on morphology and SVM classifier	Dr. Vijayalakshmi G V	ISBN13: 9781668475447	IGI Global (contributed a chapter)
9	Implementation of machine learning-aided speech analysis for speaker accent identification applied to audio forensics	Dr. Vijayalakshmi G V	ISBN13: 9781668445587	IGI Global (contributed a chapter)
10	Investigation of deep fake images using pre-trained CNN frameworks	Dr. Vijayalakshmi G V	ISBN13: 9781668445587	IGI Global (contributed a chapter)

11	Futuristic Trends in Network & Communication Technology	Dr. Dankan Gowda V	ISBN-10 : 8195935613	Iterative International Publishers (Edited book)
12	Chapter Title: Intelligent IoT Enabled Lighting System Book Title: Artificial Intelligence for Smart Cities and Villages:Advanced Technologies, Development, and Challenges	Dr. Dankan Gowda V	ISBN: 978-981-5049-26-8 (Print)	Bentham Science Publishers(contributed a chapter)
13	Chapter Title: IoT Enabled Smart Village for Sustainable Development Book Title: Artificial Intelligence for Smart Cities and Villages:Advanced Technologies, Development, and Challenges	Dr. Dankan Gowda V	ISBN: 978-981-5049-26-8 (Print)	Bentham Science Publishers(contributed a chapter)
14	IoT-Deep Learning-Based Detection of Cyber Security Threats	Dr. Dankan Gowda V	ISBN13: 9781668445587 EISBN13: 978166844560	IGI Global (contributed a chapter)
15	Deep learning-based automatic student authentication	Dr. Dankan Gowda V	ISBN13: 9781668445587 EISBN13: 978166844560	IGI Global (contributed a chapter)
16	Deep learning models for cyber security in IoT networks	Dr. Dankan Gowda V	ISBN13: 9781668445587	IGI Global (contributed a chapter)

			EISBN13: 978166844560	
17	Doppler Radar Technique for Geriatric Fall Detection	Dr. Rashmi N Mamatha K R	2524-7565	Algorithms for Intelligent Systems book series(AIS), Springer Nature (contributed a chapter)
18	SRR-loaded octagonal Sierpinski-based carpet-shaped antenna for multiband application	Anitha V R, K.Yogaprasad and Nanda Kumar M	DOI: 10.1201/9781003347057-6	Printed Antennas, Springer Nature (contributed a chapter)
19	Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection	M Lakshmi Prasanna and Anitha V R	DOI: 10.1007/978-981-19-2308-1_37	Micro and Nanoelectronics Devices, Circuits and Systems, Springer (contributed a chapter)

International Journal Details

International Conference

Book/Book Chapters

Dept. of ECE

Research Compendium

ACY:2022-23

A Super-resolution-based Approach for the Detection of Covid–19 Infection From Chest X-ray Images*

Seema S. Bhat and Hanumantharaju M. C

X-ray is the most accessible imaging modality for detecting Covid-19 infection. However, X-ray image resolution depends on the amount of radiation dose. The Lesser the dosage, the lower the resolution, the higher the noise and patient safety. Detecting Covid-19 infection would be more precise with high-resolution chest X-ray images. The current article explores an edge-preserving single-scale residual learning-based super-resolution method to enhance low-resolution chest X-ray images. We used unsharp masking to preserve small, medium, and high-scale details while super-resolving the given image. The method produces a clear view of the pulmonary opacities in chest X-ray images after super-resolution reconstruction. Statistical feature metrics of first and second-order showed superior quality reconstruction by the proposed method for the given Covid-19 chest X-ray images. Further, to measure the effectiveness of super-resolution, we used an Inception v3 based deep learning model to classify chest X-ray images of Covid-19, pneumonia, and normal class. The performance of the classification model with super-resolved chest X-ray images was tested against 400 images belonging to two different classes at a time. We obtained increased precision of 94% and 96% accuracy in detecting Covid-19 infection in chest X-ray images after super-resolution compared to 64% precision and 68% accuracy before super-resolution.

Seema S. Bhat is currently working as an Assistant Professor in Dept. of Information Scienece and Engineering in Dayananda Sagar College of Engineering, Bengaluru. Her research interests include magnetic resonance imaging and medical image processing.

Hanumantharaju M. C is currently working as a Professor in Dept. of Electronics and Communication in BMS Institute of Technology and Management, Bengaluru. His research interests include VLSI, algorithm design, reconfigurable IC design and image processing.

*Vol.28, No.1, DOI: https://doi.org/10.1007/s12045-023-1530-7

Introduction

Covid-19 is a disease caused by a new strain of severe acute respiratory syndrome coronavirus [1]. The use of reverse-transcriptase polymerase chain reaction (RT-PCR) and nucleic acid amplification tests (NAATs) along with the biomarkers are primarily used for Covid-19 detection and diagnosis [2]. However, the use of RT-PCR and biomarkers take longer and depend on the specimen taken, which could have faults. Imaging provides a faster solution, and the common imaging modalities used for Covid-19 diagnosis are chest computed tomography (CT), chest X-ray (CXR), and lung ultrasonography (LUS) [2]. We can also use magnetic resonance imaging (MRI) and positron-emitting tomography (PET) for pulmonary imaging [3]. Both cannot be used as first-line modalities due to longer scan times and higher costs. Also, PET scanning has a radiation burden. Likewise, LUS diagnosis accuracy depends on the operator's expertise and witnesses low specificity and sensitivity [4]. CT scans have been used frequently in Covid-19 screening, as they provide high anatomic resolution. But, CT scans propagate the menace of contamination and radiation exposure. However, chest X-ray prevents cross-infection among the staff and the patients. X-ray is cost-effective and accessible worldwide [2, 4].

CT scans have been used frequently in Covid-19 screening, as they provide high anatomic resolution. But, CT scans propagate the menace of contamination and radiation exposure.

Chest X-rays can be used from the onset of infection till it turns negative. The severity of lung abnormalities is found to peak at 5–10 days of symptom onset. The mean time from initial positive findings to negative chest X-ray is found to be 10 days(range 6–14 days). Radiologists visualize ground glass patterns, dense opacities obscuring vessels, and bronchial walls in the lungs in Covid-19 chest X-rays [5]. The opacities can damage both lungs at an early stage of Covid-19. Bilateral lower zone consolidation is also a symptom of Covid-19, usually appearing 10–12 days after the disease onset. One of the Covid-19 studies found these symptoms [6], and the study has claimed 69% sensitivity in finding Covid-19 symptoms with baseline chest X-ray images. In recent years, various machine learning models have been applied to solve different tasks, which otherwise was not possible.

In addition, multiple techniques have used deep learning for the enhancement, feature detection, segmentation, and detection of Covid-19 lung lesions [7, 8]. Artificial intelligence based Covid-19 infection detection offers an automated approach to disease diagnosis, thereby reducing reliance on the human eye. The features learned in deep architecture would detect invisible traces of infection, which could not be traced otherwise by the human eyes. These automated methods can be adapted in Covid-19 diagnosis centers as they are inexpensive, quicker, and simpler [9].

X-ray images have low resolution (LR) and possess noise due to low X-ray dosage. Low X-ray dosage is used to maintain patient safety [10]. Pulmonary diseases often have hazy structures. One can obtain more acceptable X-ray images with a higher radiation dose or an expensive imaging setup. Super-resolution is a class of techniques that enhance noisy LR images using a set of high-resolution (HR) images [10, 11]. Super resolution-based enhancement would be easier to get high-quality chest X-rays with minimal radiation dosage to the patients [10]. High-resolution X-ray images may increase the sensitivity and specificity of pathology detection and diagnosis.

X-ray images have low resolution (LR) and possess noise due to low X-ray dosage. Low X-ray dosage is used to maintain patient safety.

Most super-resolution methods are experimented with natural images and eventually extended their proficiencies in medical images. Textures in medical images are more complicated, affluent, and highly correlated [12, 13]. The medical image quality depends on imaging equipment, patient position, and the complicated and unpredictable shape of the scanned organs. Several methods have been proposed based on this aspect [14–16].

According to the literature, single-image super-resolution methods are categorized into three types. The first type is the interpolation method. High computational speed makes interpolation-based methods suitable for parallel computing [17, 18]. They do not produce valuable high-frequency details [18]. Therefore, it is hard to obtain sharp and high-resolution images with interpolation-based methods [18].

Super-resolution reconstruction (SRR) uses the HR reference im-

Super-resolution reconstruction (SRR) uses the high-resolution reference image to improve the low-resolution image.

age to improve the resolution of the LR image [19–23]. Reconstruction-based methods are typically dependent on well-designed priors. Most of them do not capture the complete characteristics of image priors, as they use shallow features of the image and manually chosen parameters [11, 14, 24]. Nevertheless, the hand-crafted tuning parameters make operations more complicated and prone to errors. However, without much information on image priors and repetitive patterns, these methods yield objectionable edges [11, 14].

In the third type, super-resolution images produce the corresponding HR version by learning the mapping function between LR and HR images [25–28]. Dictionary-based super-resolution methods consume more time to build dictionaries and reconstruct HR versions with limited real-time medical data. Also, optimizing the performance in every stage of learning and reconstructing is a tedious job.

Deep learning-based super-resolution methods extract image features implicitly to get high-resolution images.

Deep learning-based super-resolution methods extract image features implicitly to get HR images. Mapping is done automatically through convolution operations to enhance details. The first convolutional neural network (CNN) for super-resolution had three layers [29]. However, the deep learning literature has various deep network architectures for super-resolution [15, 29–34]. Deep architectures allow SR methods to expand their receptive field and capture more contextual details over large image regions. With the same weight parameters for all the convolutional layers, a deeply recursive SR technique in [30] did not change the receptive field size. Tai et al. stacked numerous residual recursive blocks to get a better model [32]. Kang et al. used ResNet architecture with wavelets to denoise and preserve texture in low-dose CT/ X-ray images [33]. Frequency localization properties of the wavelets preserve small-scale structures in the given low-dose X-ray images. Residual learning bypassed the low-frequency wavelet coefficients and, added with the denoised wavelet coefficients, made training more efficient. The larger receptive field enhanced the performance of the proposed method. However, the larger receptive field results in a deeper network.

Also, input images of lower thickness lost high-frequency details and produced blurred edges after reconstruction [33]. Xu et al. proposed GAN-based super-resolution for chest X-ray images with auxiliary label information for feature extraction and spectral normalization for good performance [15]. However, GAN-based SR methods are known to add noise while enhancing high-frequency details. Also, a few details were lost during the partial enlargement of regions of interest [15].

Taking more image contexts into account would miss high-frequency components in super-resolution [10, 11]. Expanding the receptive field deepens the architecture and enlarges the filter size [34]. This expansion also results in multiple parameters to tune while training and reduces the performance of the model [34]. Moreover, any medical imaging needs additional prior details compared to natural images. It is also complicated to enhance intricate structures and preserve sensitive details. Additionally, medical images have a relatively small dataset. It is not easy to collect clinical high- and low-resolution image pairs. The current work adapts the ResNet-based super-resolution network architecture proposed by Lim et al. [31]. In the second phase, we detect Covid-19 infection in the given chest X-ray image using an Inception v3-based classification model trained by super-resolved chest X-ray images. We use a residual learning-based convolution neural network (CNN) with unsharp masking to obtain SR Covid-19 chest X-ray images in this work. *Figure* 1 shows the single-scale residual network architecture used in the current work. All experiments were conducted with a GPU graphics card on Google Colaboratory with Windows 10 OS.

It is not easy to collect clinical high- and low-resolution image pairs.

The major highlights of this work are listed below:

1. We present a deep learning-based super-resolution method for Covid-19 chest X-ray images. High-resolution chest X-ray images would assist in the precise diagnosis of the disease.

2. We used a single-scale residual network model to reduce noise and amplify details in the low-resolution chest X-ray images. The residual network reduces noise and is followed by unsharp

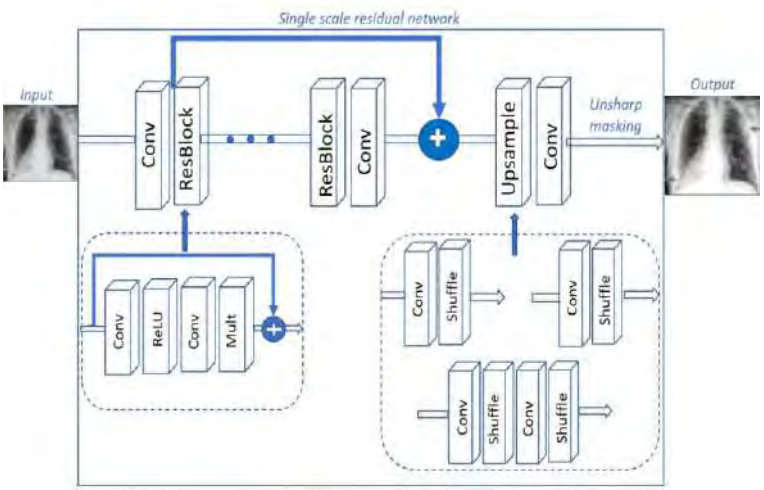


Figure 1. Workflow of the proposed method. We use edge-preserving single-scale residual network for Covid-19 chest X-ray super resolution.The network architecture is adpated from [31]. Unsharp masking is used to preserve details in the test samples.

masking, which is an edge-preserving high-pass filter that safeguards intricate details in chest X-ray images.

3. To improve the accuracy of Covid-19 chest X-ray image enhancement and quantitative assessment, we use the geometric self-ensemble strategy.

4. We could detect Covid-19 infection in a given chest X-ray image with more than 96% accuracy after super-resolution.

Materials and Methods

The proposed method works in two phases. In the first phase, we propose an edge-preserving super-resolution deep learning model to enhance low-resolution chest X-ray images. In the second phase, we detect Covid-19 infection in the given chest X-ray image using a classification model.

Phase 1: Edge-preserving Super-resolution Deep Learning Model

Dataset: The experiment uses the DIV2K dataset for training [35]. DIV2K has 800 training images, 100 validation images, and 100 test images of high quality with 2K resolution. The testing phase used selected samples from the COVID-19 chest X-

ray public dataset [36]. The dataset exclusively has 123 frontal view chest X-ray images. There are 244 total samples, including Covid-19 and other viral/ bacterial pneumonia in the posterior-anterior (PA) or vice-versa (AP) view of chest CT and X-ray images. It also has 38 samples of AP supine[1] view chest images. Though public Covid-19 chest X-ray images are available, they include non-Covid-19, pneumonia, and other chest X-ray images. There is no gold standard HR chest X-ray images specified in the test set. So, we reported and compared the performances on the DIV2K validation dataset.

[1] Supine anteroposterior view.

Network archirecture: *Figure* 1 shows the network architecture used to super-resolve the Covid-19 chest X-ray images. Each residual block in an edge-preserving single-scale residual network has two convolution layers. The first convolutional layer does feature extraction, and the last convolution layer does the nonlinear mapping. Eliminating batch normalization and final ReLU layers from the residual block in a single-scale residual network increases the super-resolution performance.

Each residual block in an edge-preserving single-scale residual network has two convolution layers.

Training and testing: The proposed super-resolution model uses 800 DIV2K training images. The model uses image patches of size 48×48 and normalized before training. Augmentation of the image patches includes 90° rotations and random horizontal flips. The training uses 30,000 steps with a constant learning rate of 1×10^{-4} and utilizes l_1 loss instead of l_2 for better convergence.

Testing was done on selected samples of matrix size—365×435—from the Covid-19 chest X-ray public dataset by Cohen [36]. Linear unsharp masking increases small-scale acutance to emphasize the texture and details of the given image [37]. Unsharp masking is an edge-preserving high-pass filter used in the testing phase of the proposed method to sharpen the edges of the reconstructed images [38]. The radius in unsharp masking decides the size of the edge to be enhanced. A larger radius affects small edges and vice versa. The magnitude and sharpness of the sharpened image depend on the amount[2] and threshold. A radius = 4, amount = 1 and threshold/scaling factor = 0.5 made reconstructed images sharper.

[2] The amount represents how intense the user wants the sharpening to be. A low amount setting results in rather mild sharpening lines, but a high amount setting produces extremely dark and extremely light lines.

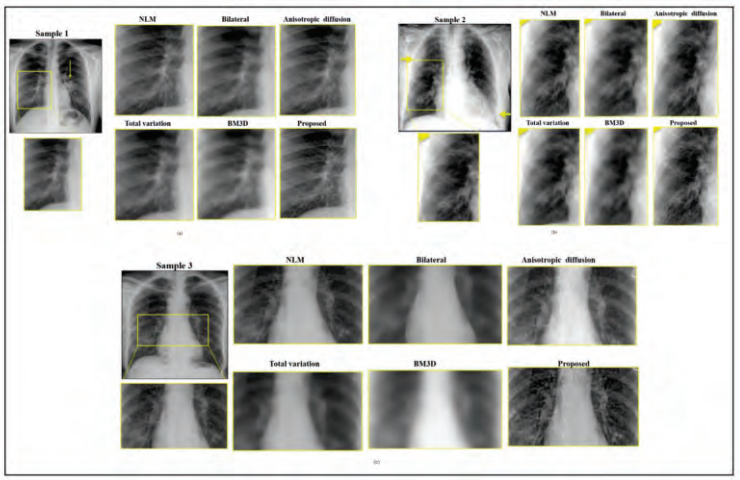


Figure 2. Visual represenation of the proposed super-resolution and edge preserving comparative methods of COVID-19 chest X-ray images.The figure shows visual quality of the edge-preserving single scale residual-based super resolution method and state of the art edge-preserving algorithms on chosen three samples.Loss of contour and other details are observed in the case of bilateral, total variation and BM3D reconstructed COVID-19 chest X-ray images.NLM and anisotropic diffusion witnessed loss of small and medium contrast details. The proposed method was able to preserve and enhance the details.

The final up-sampling layer includes a convolutional layer followed by a shuffle for $\times 2$, $\times 3$, and $\times 4$ upscaling factors. Each convolutional layer produces sixty-four feature maps. The depth of the network and the number of feature channels used are 32 and 256, respectively. The experiment used a scaling factor of 0.1.

The self-ensemble strategy is similar to [31] used in the testing phase. Input image I^{LR} to generate seven augmented inputs: $I^{LR}_{n,i} = T_i(I^{LR}_{n,i})$, where T_i represents transformation function, including identity by flipping and rotating. For every augmented image, the proposed model generated a corresponding super-resolved image $I^{SR}_{n,i}$ with $i = 1$ to 8. Later, the inverse transform is applied to get the original geometry. Final SR images are obtained by averaging the inverse transform $I^{SR}_{n,i} = \frac{1}{8}\sum_1^8 T_i^{-1} I^{SR}_{n,i}$.

The performance evaluation of the proposed method included a comparison with the state-of-the-art edge-preserving algorithms: BM3D, anisotropic diffusion total variation, bilateral filters, and nonlocal means.

Performance evaluation and comparative methods: The performance evaluation of the proposed method included a comparison with the state-of-the-art edge-preserving algorithms: BM3D, anisotropic diffusion total variation, bilateral filters, and nonlocal means [39–41]. These are edge-preserving filters, initially designed for realistic images, but extend their applications successfully to X-ray image enhancement and other medical imaging modalities [38–41]. The implementation of these algorithms

is officially available on the internet. All of these algorithms use the same set of normalized images from the test dataset for execution.

All the above-mentioned edge-preserving algorithms have at least one controlling parameter. We have used either maximum value or default values for the tuning parameters of these state-of-the-art edge-preserving algorithms. BM3D algorithms eliminate correlated stationary noise in the given image. It uses two tuning parameters for noise reduction: σ and hard threshold to map the pixel values. We have used the variance of the given image for parameter σ and default hard thresholding for image reconstruction.

Anisotropic diffusion algorithms require the number of iterations, conduction coefficient κ, stability parameter γ, and choosing an option between Perona–Malik diffusion equation 1 or 2. The current work uses iterations = 50, $\kappa = 50$, $\gamma = 0.2$, and equation 2 to reconstruct X-ray images. Total variation measures the similarity and dissimilarity of pixels to find the meaningful relationship between them. The smoothing parameters use a weight of 0.1 in the total variation experiment with Covid-19 chest X-ray images.

Anisotropic diffusion algorithms require the number of iterations, conduction coefficient κ, stability parameter γ, and choosing an option between Perona–Malik diffusion equation 1 or 2.

A bilateral filter needs the diameter of the pixel neighborhood and parameter σ for color and coordinate space. For larger values of color and coordinate σ values, greater will be the mixing of pixels in the given range. We have used diameter = 15, σ color, and coordinate space = 75 each in the present work.

NLM method uses luminance component regulator h, template window size. The weight of the template patch is computed in terms of pixels and search window size. The window size is used to compute the weighted average for a given pixel. The default values 3, 7, and 21 are used correspondingly for the NLM, as mentioned earlier.

In this work, *ImageJ*, a platform-independent image processing tool, was used for plotting edges in LR and HR chest X-ray images. We obtained sharper edges after super-resolution reconstruction. Finding edges in LR and SR X-ray images indicates

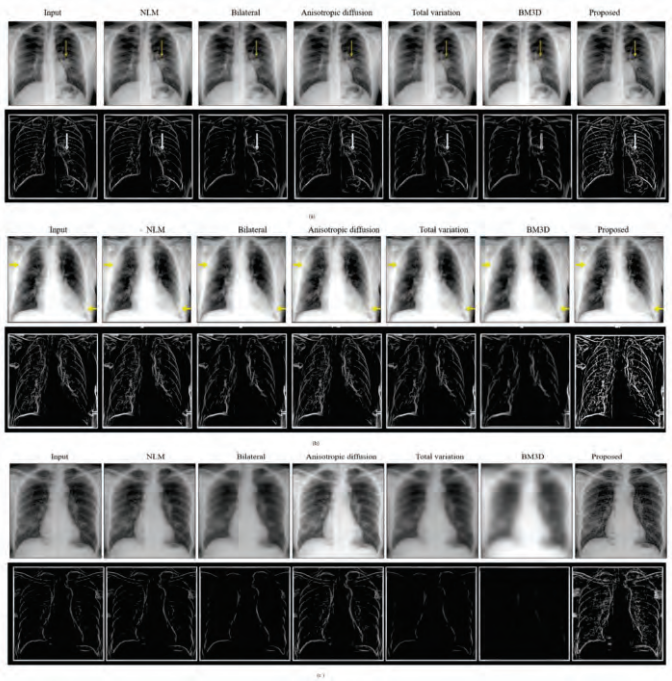


Figure 3. Visual representation of edge-preservation of the proposed and comparative methods. The comparison of small-, medium-, and high-scale edges of the proposed edge-preserving super-resolution method with the state-of-the-art algorithms for the chosen three samples is shown. Enhanced edges produced by the proposed super-resolution method can be observed for the chosen sample in the last column.

detail preservation by SR reconstruction. *Find Edges* application of *ImageJ* facilitated visualization of edge-preservation with comparative and proposed methods [42].

Evaluation metrics: We evaluated the proposed and comparative edge-preserving algorithms' performance with texture-related metrics. We do not use measures like PSNR, or SSIM, which are correlated with human perception, for two reasons. The first reason is that while performing this study, there were no gold standard reference Covid-19 chest X-ray images, and the second reason is that texture has more importance in medical diagnosis. A clear texture simplifies clinical diagnostics [15] [43–45]. We used selected blind reference metrics derived from first and second-order statistical feature measures. First-order statistical feature metrics are mean, standard deviation, and entropy. They are used to measure the contrast and brightness of reconstructed images. Entropy measures the content of the image; the higher the entropy richer the details in the given image.

Second-order statistical metrics—contrast, homogeneity, energy,

and correlation are obtained using the grey-level co-occurrence matrix (GCLM). They play a significant role in the measure of image sharpness [45]. The larger the changes in grey levels, the higher contrast and sharpness. GCLM energy defines the overall probability of unique gray-scale patterns in the image. To summarize, higher mean, standard deviation, entropy, contrast, and correlation indicate the reconstructed image's superior quality, whereas energy and homogeneity decrease with image quality.

Second-order statistical metrics—homogeneity, energy, and correlation are obtained using the grey-level co-occurrence matrix. They play a significant role in the measure of image sharpness.

Phase 2: Classification Model for Covid-19 Infection Detection

We evaluated the effectiveness of the super-resolution method for the given chest X-ray images before and after super-resolution using a classification model. We utilized VGG16 and Inception v3 models as base models. We used normal, Covid-19, and pneumonia chest X-ray images from the COVID-19 Radiography Database by Kaggle. It had 3616 Covid-19 positive cases along with 10,192 normal and 1345 pneumonia chest X-ray images. The classification model labelled the given chest X-ray image under covid and non-covid (normal and pneumonia) class.

Training and testing: In the first phase, the training and testing experiments included Covid-19 and normal chest X-ray image pairs from the Kaggle database. The model was trained on 200 samples of each Covid-19 and normal images before and after super-resolution. All the images used for training were augmented, normalized, and resized to 224×224 dimension. The batch size (BS) was set to 16, and a learning rate of 1e-4 was used throughout the training. The training labeled each chest X-ray image to covid or non-covid class based on classification probability. We plotted a receiver operating characteristic (ROC) curve to visualize the performance of the classification. Also, a confusion matrix was created to compare the actual target chest X-ray images with the images predicted by the classification model. We computed the precession, F1-score, recall, and accuracy of the classification model for both covid and non-covid classes. We measured the testing loss and testing accuracy of the classification model. In the second phase, training and testing were repeated for

Covid-19 and pneumonia chest X-ray image pairs before and after super-resolution. In total, we tested 800 chest X-ray images belonging to Covid-19 and normal and Covid-19 and pneumonia classes. Test loss and test accuracy were computed before and after super-resolution.

Performance evaluation: To evaluate the performance of the classification model, we used four common measures. They are—precision, recall, accuracy, and F1 score—as described as below [46]:

$$Precision = \frac{TP}{(TP + FP)}. \tag{1}$$

$$Recall = \frac{TP}{(TP + FN)}. \tag{2}$$

$$F1\ score = 2 * \frac{(Precision * Recall)}{(Precision + Recall)}. \tag{3}$$

$$Accuracy = \frac{(TN + TP)}{(TN + TP + FN + FP)}. \tag{4}$$

Here, TP is the count of true positive samples, TN is the count of true negative samples, FP is the count of false positive samples, and FN is the count of false negative samples from a confusion matrix.

Results and Discussion

The experiment involved three randomly chosen samples from the test set. The first column in *Figure* 2 shows the test images chosen with a posterior-anterior (PA) view of the patient. The rest of the columns in *Figure* 2 show reconstructed images of chosen edge-preserving filters and the proposed method. We enlarged randomly chosen regions to view the lung bronchioles, alveolus, and ground-glass opacities in the reconstructed images.

Edge-preserving algorithms are supposed to retain details while removing the noise from the given image.

Edge-preserving algorithms are supposed to retain details while removing the noise from the given image. For the given sample, there wasn't much difference in the details observed in the NLM and anisotropic diffusion (AD) filter in the magnified regions. Noise present in small-scale features of X-ray images foils accurate averaging of coefficients and generates artifacts in NLM filters. In the anisotropic diffusion method, the regularization parameter's optimal value depends on the prior knowledge of the noise. AD smoothens flat regions, and the tuning regularization parameter is complex. However, noise reduction resulted in blurred edges in the bilateral, total variation, and BM3D filters. It is hard to observe the contour and other details in their reconstructed counterpart. The reason is, in high gradient regions, bilateral filters tend to produce cartoonish texture. At the same time, the performance of BM3D worsens with higher levels of noise in the images. The total variation filters tend to lose contrast during image restoration; hence blurring small- and medium-scale edges are observed.

The reconstructed results from the proposed method, edges, and other details were well preserved, as shown in *Figure* 2(a–c). The larger receptive fields can capture small- and medium-scale edges in X-ray images, and residual learning can propagate these features to the deepest layers. Also, the proposed method learns features implicitly; it does not require prior knowledge of the noise distribution. Single-scale residual blocks can effectively eliminate noise. Unsharp masking is a postprocessing technique that preserves high-frequency details in the reconstructed image.

We generated the edge maps to visualize the amount of edge preservation of each technique mentioned earlier. *Figure* 3(a–c) shows edge maps of the proposed and comparative edge-preserving filters. NLM and anisotropic filters were able to retain most of the edges like in the given three samples. However, bilateral, total variation, and BM3D reconstructed images suffered from loss of low-frequency edges. Unsharp masking is a powerful way of enhancing edges in the given images. In the proposed method, unsharp masking helped to retain low-scale and high-scale details

(see last column in *Figure* 3 (a–c)).

We evaluated the proposed and comparative methods using statistical feature measures of first and second orders because the standard reference image was unavailable. The first-order metrics include individual pixels to evaluate the performance. The mean and standard deviation of the reconstructed images show pixel intensities in the given image. The proposed method obtained a higher mean and standard deviation than its LR counterpart and some comparative filters. Entropy indicates information present in the given image. The proposed method's higher entropy values for all three samples indicate high-intensity levels and high contrast in reconstructed images. The plots in *Figure* 4 summarizes first-order statistical measures for the given samples and their reconstructed images. The proposed method exhibits the best entropy and standard deviation values compared to other edge-preserving algorithms. Total variation and BM3D obtained pretty good values. However, the mean, standard deviation, and entropy values obtained using NLM, bilateral, and anisotropic diffusion algorithms were similar to those of the input samples.

The second-order statistical measures incorporate the neighborhood of pixels to evaluate the performance of the image processing techniques and impact the sharpness of the given image. We obtained the second-order statistics by constructing a GLCM matrix for each of the above-mentioned edge-preserving methods. GCLM details the positions of the pixels with similar gray levels in the image. It is obvious to notice higher contrast and correlation values for the proposed method in *Figure* 5. Also, the proposed method obtained lower energy and homogeneity values. Lower energy and homogeneity values indicate a lower localized change in the image. We obtained higher contrast and correlation values for the proposed method, indicating superior sharpness and brightness.

We evaluated the performance of the proposed super-resolution using a classification model. *Figure* 4 shows the sample images belonging to Covid or non-Covid classes. *Figure* 5 shows selected

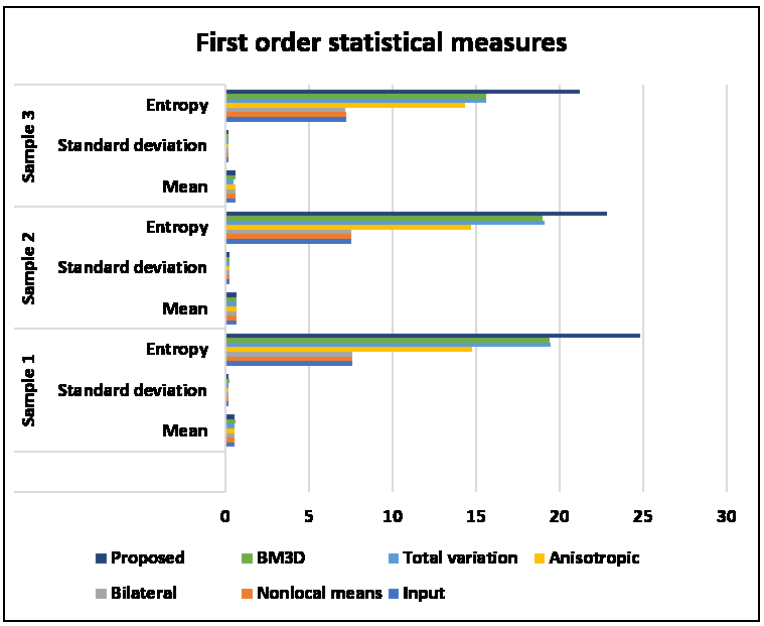


Figure 4. Image reconstruction results using first-order statistical metrics. We computed the mean, standard deviation, and entropy of the proposed edge-preserving super-resolution method with state-of-the-art edge-preserving algorithms. The proposed method obtained a higher mean and entropy than the other methods.

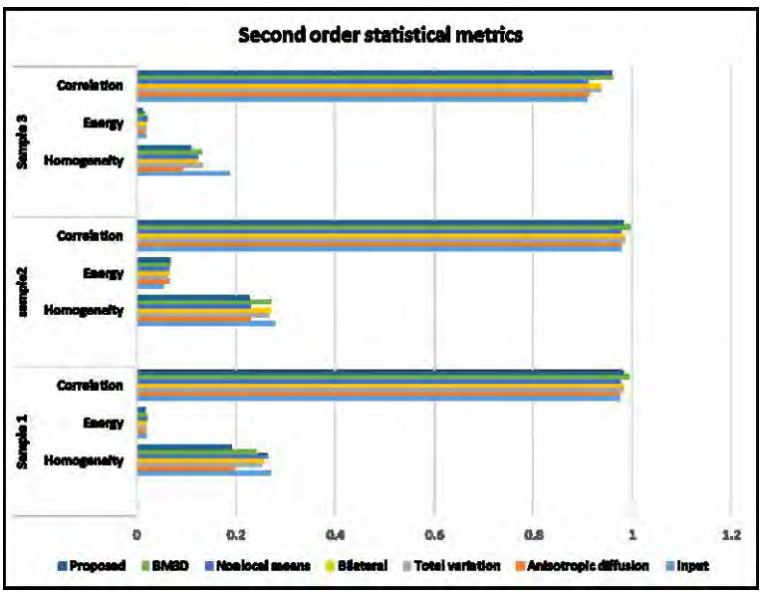


Figure 5. Performance evaluation using second-order statistical metrics. We computed homogeneity, energy, and correlation of the proposed edge-preserving super-resolution method with state-of-the-art edge-preserving algorithms. The proposed method's good quantitative values ensured enhanced contrast and brightness after super-resolution.

images obtained after model prediction whether the test image is Covid or non-Covid with its probability. The model classifies the

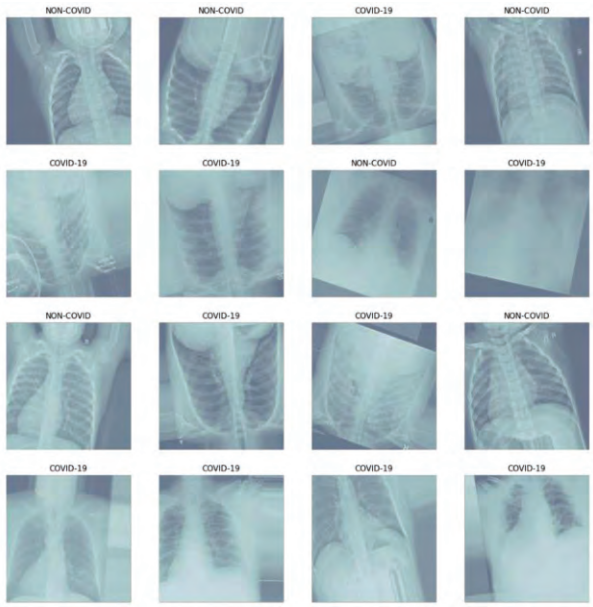


Figure 6. Sample images from training and testing with their class labels. The labels are assigned by the classification model by extracting features from each image during the training phase.

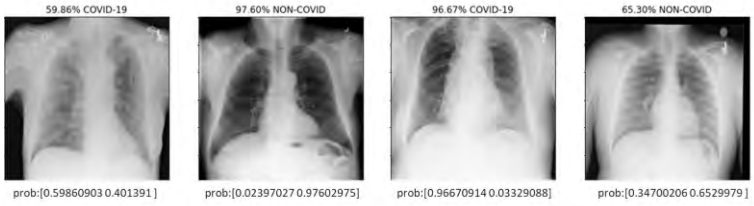


Figure 7. Test image prediction samples with their probability value. A probability higher than 0.5 is considered to be the Covid class. The probabilities are calculated during the training of the classification network.

test images as covid if the probability is more than 0.5. Otherwise, the model classifies the test image as non-Covid.

We computed the test loss and test accuracy before and after super-resolution for 800 chest X-ray images belonging to Covid-19 and normal and Covid-19 and pneumonia classes. We obtained a test loss of 0.7160 and a test accuracy of 0.6399 before super-resolution. Higher test accuracy of 0.7599 and lower test loss of 0.3759 were attained after super-resolution.

Figure 6 shows ROC curves obtained for classifying Covid-19 and normal chest X-ray images. The curve is closer to the upper

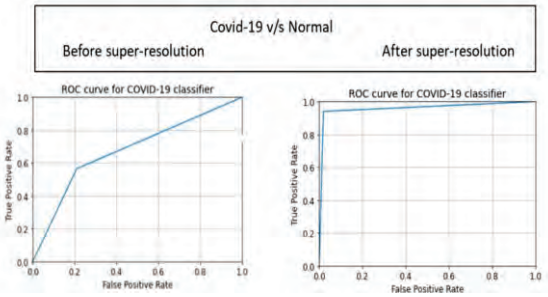


Figure 8. ROC curve for Covid-19 versus normal chest X-ray images. The curve closer to the upper left corner indicates higher test accuracy. The curve closer to the corner after super-resolution indicates higher classification accuracy.

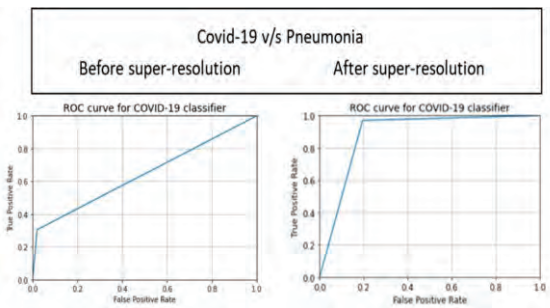


Figure 9. ROC curve for Covid-19 versus pneumonia chest X-ray images. The curve closer to the corner after super-resolution shows higher classification accuracy in predicting Covid and non-Covid images.

left corner, which indicates higher test accuracy. It is evident that ROC curve obtained after super-resolution has increased sensitivity in detecting Covid-19. *Figure* 7 shows ROC curves plotted for Covid-19 and pneumonia image classification. Obviously, a better ROC curve is noticed after super-resolution. *Figure* 8 and 9 show confusion matrix Covid and non-Covid class images. The proposed method labeled 196 Covid-19 images out of 200 images after super-resolution compared to before super-resolution values 127 and 158.

Table 1 and 2 shows precision, recall, and F1 score values before and after super-resolution. We obtained higher precision and recall values after super-resolution. The classification accuracy increased from 0.68 to 0.96 after super-resolution in Covid-19 and normal class. At the same time, we obtained an increase in accuracy for Covid-19 and pneumonia cases.

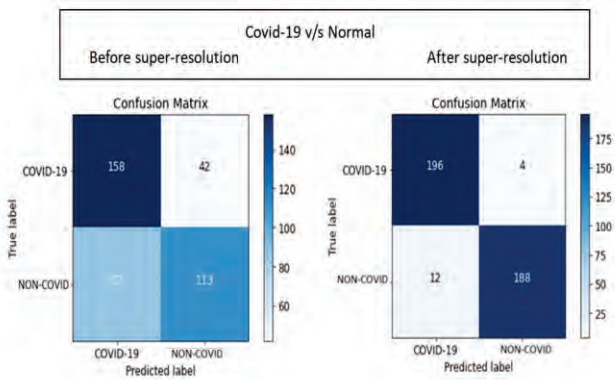


Figure 10. Confusion matrix Covid-19 versus normal chest X-ray images. An increase in classification accuracy after super-resolution can be observed.

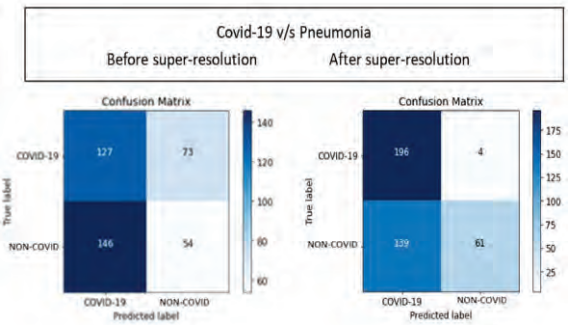


Figure 11. Confusion matrix Covid-19 versus pneumonia chest X-ray images. An increase in classification accuracy after super-resolution can be observed.

Conclusion

The proposed deep learning-based edge-preserving super-resolution method is used to obtain high-resolution Covid-19 chest X-ray images to detect Covid-19 infection. Unsharp masking during the testing phase preserves details, and a single-scale residual network reduces inherent noise in Covid-19 chest X-ray images. The

Table 1. Precision, recall, F1 score, and accuracy of classification model before and after super-resolution for Covid-19 versus normal class. Improved values in all the classification metrics can be observed after super-resolution.

		2 way: Covid v/s Normal			
		Precision	Recall	F1 score	Accuracy
Before super-resolution	Covid-19	0.64	0.79	0.71	0.68
	Normal	0.73	0.56	0.64	
After super-resolution	Covid-19	0.94	0.98	0.96	0.96
	Normal	0.98	0.94	0.96	

		2 way: Covid v/s Pneumonia			
		Precision	Recall	F1 score	Accuracy
Before super-resolution	Covid-19	0.47	0.64	0.54	0.45
	Pneumonia	0.43	0.27	0.33	
After super-resolution	Covid-19	0.59	0.98	0.73	0.64
	Pneumonia	0.94	0.30	0.46	

Table 2. Precision, recall, F1 score, and accuracy of classification model before and after super-resolution for Covid-19 versus pneumonia class. We obtained higher classification metrics as expected after super-resolution.

experimental results show that the proposed method can effectively reconstruct Covid-19 features compared to other state-of-the-art techniques. Higher entropy, contrast values, and lower homogeneity energy values in performance evaluation ensured superior quality reconstruction of the proposed method. The classification model trained using super-resolved chest X-ray images obtained higher accuracy of 96%. However, analysis has been done on limited data, and only preliminary results have been obtained. The enhancement schemes are purely exploratory and do not include clinical trials/medical validations of the approach. Also, there is a potential risk of generating features not part of the target images. Network optimization and effective mapping between low- and high-resolution feature space can be potential future work.

Suggested Reading

[1] **Na Zhu et al., A novel coronavirus from patients with pneumonia in China, 2019. *New England Journal of Medicine*, 20;382(8), pp.727–733, 2020.**

[2] **Tinotenda A Harahwa et al., The optimal diagnostic methods for Covid-19. *Diagnosis*, Vol.7, No.4, pp.349–356, 2020.**

[3] **Brandon K K Fields, Natalie L Demirjian, Habibollah Dadgar and Ali Gholamrezanezhad, Imaging of Covid-19: CT, MRI, and PET, In *Seminars in Nuclear Medicine*, Vol.51, pp.312–320. Elsevier, 2021.**

[4] **Liqa A Rousan, Eyhab Elobeid, Musaab Karrar and Yousef Khader, Chest X-ray findings and temporal lung changes in patients with Covid-19 pneumonia, *BMC Pulmonary Medicine*, Vol.20, No.1, pp.1–9, 2020.**

[5] **V N Manjunath Aradhya, Mufti Mahmud, D S Guru, Basant Agarwal, and M Shamim Kaiser, One shot cluster based approach for the detection of Covid-19 from chest X-ray images, *Cogn Comput.*, 13(4), pp.873–881, 2021.**

[6] **Ho Yuen Frank Wong et al., Frequency and distribution of chest radiographic findings in patients positive for Covid-19, *Radiology*, Vol.296, No.2, pp.E72–E78, 2020.**

[7] Seema S Bhat, M C Hanumantharaju, and V N Manjunath Aradhya, Super resolution techniques for Covid-19 chest X-ray images, In *Computer Communication, Networking and IoT*, pp.549–558. Springer, 2021.

[8] Sudhen B Desai, Anuj Pareek, and Matthew P Lungren, Deep learning and its role in Covid-19 medical imaging, *Intelligence-based Medicine*, Vol.3, 100013, 2020.

[9] Shashank Vaid, Reza Kalantar and Mohit Bhandari, Deep learning Covid-19 detection bias: Accuracy through artificial intelligence, *International Orthopaedics*, Vol.44, pp.1539–1542, 2020.

[10] Masato Shimizu, Hidetoshi Kariya, Tomio Goto, Satoshi Hirano and Masaru Sakurai, Super-resolution for X-ray images, In *2015 IEEE 4th Global Conference on Consumer Electronics (GCCE)*, pp.246–247, IEEE, 2015.

[11] Yoshiki Sano, Takuma Mori, Tomio Goto, Satoshi Hirano and Koji Funahashi, Super-resolution method and its application to medical image processing, In *2017 IEEE 6th Global Conference on Consumer Electronics (GCCE)*, pp.1–2, IEEE, 2017.

[12] Yang Li, Wei Liang, Haibo An, Yinlong Zhang and Jindong Tan, C-arm based image-guided percutaneous puncture of minimally invasive spine surgery, In *2017–10th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)*, pp.1–5. IEEE, 2017.

[13] Concetto Spampinato, Simone Palazzo, Daniela Giordano, Marco Aldinucci and Rosalia Leonardi, Deep learning for automated skeletal bone age assessment in X-ray images, *Medical Image Analysis*, Vol.36, pp.41–51, 2017.

[14] Kensuke Umehara, Junko Ota, Naoki Ishimaru, Shunsuke Ohno, Kentaro Okamoto, Takanori Suzuki, and Takayuki Ishida, Performance evaluation of super-resolution methods using deep-learning and sparse-coding for improving the image quality of magnified images in chest radiographs, *Open Journal of Medical Imaging*, Vol.7, No.3, pp.100–111, 2017.

[15] Liming Xu, Xianhua Zeng, Zhiwei Huang, Weisheng Li and He Zhang, Low-dose chest X-ray image super-resolution using generative adversarial nets with spectral normalization, *Biomedical Signal Processing and Control*, Vol.55, p.101600, 2020.

[16] Chenyu You et al., CT super-resolution gan constrained by the identical, residual, and cycle learning ensemble (GAN-circle), *IEEE Transactions On Medical Imaging*, Vol.39, No.1, pp.188–203, 2019.

[17] Michel Bätz, Andrea Eichenseer, Jürgen Seiler, Markus Jonscher and Andre Kaup, Hybrid super-resolution combining example-based single-image and interpolation-based multi-image reconstruction approaches, In *2015 IEEE International Conference On Image Processing (ICIP)*, pp.58–62. IEEE, 2015.

[18] Robert Keys, Cubic convolution interpolation for digital image processing, *IEEE Transactions On Acoustics, Speech, and Signal Processing*, Vol.29, No.6, pp.1153–1160, 1981.

[19] Michal Irani and Shmuel Peleg, Super resolution from image sequences, In *Proceedings 10th International Conference on Pattern Recognition*, Vol.2, pp.115–120. IEEE, 1990.

[20] Zhouchen Lin and Heung-Yeung Shum, Fundamental limits of reconstruction-

based super-resolution algorithms under local translation, *IEEE Transactions On Pattern Analysis and Machine Intelligence*, Vol.26, No.1, pp.83–97, 2004.

[21] Henry Stark and Peyma Oskoui, High-resolution image recovery from image-plane arrays, using convex projections, *JOSA A*, Vol.6, No.11, pp.1715–1726, 1989.

[22] Richard R Schultz and Robert L Stevenson, Extraction of high-resolution frames from video sequences, *IEEE Transactions On Image Processing*, Vol.5, No.6, pp.996–1011, 1996.

[23] Krishna Ratakonda and Narendra Ahuja, POCS based adaptive image magnification, In *Proceedings 1998 International Conference on Image Processing*, ICIP98 (Cat. No. 98CB36269), pp.203–207, IEEE, 1998.

[24] Mufti Mahmud et al., Deep learning in mining biological data, *Cogn. Comput.*, Vol.13, No.1, pp.1–33, January 2021.

[25] Sam T Roweis and Lawrence K Saul, Nonlinear dimensionality reduction by locally linear embedding, *Science*,Vol.290, No.5500, pp.2323–2326, 2000.

[26] Radu Timofte, Vincent De Smet and Luc Van Gool, Anchored neighborhood regression for fast example-based super-resolution, In *Proceedings of the IEEE International Conference on Computer Vision*, pp.1920–1927, 2013.

[27] Junjun Jiang, Xiang Ma, Chen Chen, Tao Lu, Zhongyuan Wang and Jiayi Ma, Single image super-resolution via locally regularized anchored neighborhood regression and nonlocal means, *IEEE Transactions on Multimedia*, Vol.19, No.1, pp.15–26, 2016.

[28] Jianchao Yang, John Wright, Thomas S Huang, and Yi Ma, Image super-resolution via sparse representation, *IEEE Transactions On Image Processing*, Vol.19, No.11, pp.2861–2873, 2010.

[29] Chao Dong, Chen Change Loy, Kaiming He and Xiaoou Tang, Image super-resolution using deep convolutional networks, *IEEE Transactions on Pattern Analysis and Machine Intelligence*, Vol.38, No.2, pp.295–307, 2015.

[30] Jiwon Kim, Jung Kwon Lee, and Kyoung Mu Lee, Deeply-recursive convolutional network for image super-resolution, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.1637–1645, 2016.

[31] Bee Lim, Sanghyun Son, Heewon Kim, Seungjun Nah and Kyoung Mu Lee, Enhanced deep residual networks for single image super-resolution, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition Workshops*, pp.136–144, 2017.

[32] Ying Tai, Jian Yang and Xiaoming Liu, Image super-resolution via deep recursive residual network, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.3147–3155, 2017.

[33] Eunhee Kang, Won Chang, Jaejun Yoo and Jong Chul Ye, Deep convolutional framelet denosing for low-dose CT via wavelet residual network, *IEEE Transactions on Medical Imaging*, Vol.37, No.6, pp.1358–1369, 2018.

[34] Yulun Zhang, Yapeng Tian, Yu Kong, Bineng Zhong, and Yun Fu, Residual dense network for image super-resolution, In *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp.2472–2481, 2018.

[35] Radu Timofte, Shuhang Gu, Jiqing Wu, and Luc Van Gool, NTIRE 2018 challenge on single image super-resolution: Methods and results, In *Proceedings of*

the IEEE Conference on Computer Vision and Pattern Recognition Workshops, pp.852—863, 2018.

[36] J P Cohen, P Morrison, L Dao, K Roth, TQ Duong and M Ghassemi, Covid-19 image data collection: Prospective predictions are the future, *The Journal of Machine Learning for Biomedical Imaging*, 2, pp.1–38, Dec 2020.

[37] Radu Ciprian Bilcu and Markku Vehvilainen, Constrained unsharp masking for image enhancement, In: Elmoataz, A., Lezoray, O., Nouboud, F., Mammass, D. (eds) *Image and Signal Processing*, ICISP 2008. *Lecture Notes in Computer Science*, Vol.5099. Springer, Berlin, Heidelberg. https://doi.org/10.1007/978-3-540-69905-7_2.

[38] Tommaso Cerciello, Paolo Bifulco, Mario Cesarelli and Antonio Fratini, A comparison of denoising methods for x-ray fluoroscopic images, *Biomedical Signal Processing and Control*, Vol.7, No.6, pp.550–559, 2012.

[39] Eric Michel-González, Min Hyoung Cho and Soo Yeol Lee, Geometric nonlinear diffusion filter and its application to X-ray imaging, *Biomedical Engineering Online*, Vol.10, No.1, pp.1–16, 2011.

[40] Dang Thanh et al., A review on CT and X-ray images denoising methods, *Informatica*, Vol.43, No.2, 2019.

[41] Mugahed A Al-Antari, Mohammed A Al-Masni, Mo-hamed K Metwally, Dildar Hussain, Se-Je Park, Jeong-Sik Shin, Seung-Moo Han, and Tae-Seong Kim, Denoising images of dual energy X-ray absorptiometry using non-local means filters, *Journal of X-ray Science and Technology*, Vol.26, No.3, pp.395–412, 2018.

[42] Wayne S Rasband et al., *Imagej*, 1997.

[43] Norman Burningham, Zygmunt Pizlo and Jan P Allebach, Image quality metrics, *Encyclopedia of Imaging Science and Technology*, 2002.

[44] M C Hanumantharaju, V N Manjunath Aradhya and G Hemantha Kumar, Understanding and analysis of enhanced Covid-19 chest X-ray images, In *Intelligent Systems and Methods to Combat Covid-19*, pp.47–55. Springer, 2020

[45] V L Jaya and R Gopikakumari, IEM: A new image enhancement metric for contrast and sharpness measurements, *International Journal of Computer Applications*, Vol.79, No.9, 2013.

[46] Cyril Goutte and Eric Gaussier, A probabilistic interpretation of precision, recall and f-score, with implication for evaluation, In *European Conference on Information Retrieval*, pp.345–359. Springer, 2005.

Address for Correspondence
Seema S. Bhat
#77, Ankura
1st main, 2nd stage
Arakere Micolayout
Bengaluru 560 076, India.
Email:
seemasbhat@dayanandasagar.edu
Hanumantharaju M. C
Email:
mchanumantharaju@bmsit.in

Multimedia Tools and Applications
https://doi.org/10.1007/s11042-023-15121-6

FPGA implementation of breast cancer detection using SVM linear classifier

H. S. Laxmisagar[1] · M. C. Hanumantharaju

Received: 31 August 2021 / Revised: 17 November 2022 / Accepted: 13 March 2023

Abstract
The Support Vector Machine (SVM) can be used to perform linear and nonlinear operations to solve regression and classification problems. The SVM algorithm is straightforward, generating a line or a hyperplane that can be used for separating different classes of data. However, due to its high computational complexity, SVM is a time-consuming algorithm when modeled solely with software. Various researchers attempted to implement SVM in hardware particularly on field-programmable gate array (FPGA) platforms in order to achieve high performance at lower cost and power consumption. As a result, the algorithm is unsuitable for embedded real-time applications. Therefore, SVM linear classifier is implemented on hardware which decreases the latency and executes the task in real time. In this paper, an SVM linear classifier with pipeline architecture is proposed for fast processing in Verilog HDL using a single-precision IEEE standard 754 number format. In order to perform a study related to hardware resource utilization and timing for the WBCD breast cancer datasets. The various performance metrics such as resource utilization, on-chip power consumption, and static timing analysis with constraints are evaluated. The accuracy rate is computed both using software and hardware for performance evaluation. The pipelined SVM architecture is designed using Verilog HDL, and then it is synthesized using the Vivado simulation tool. The design is configured to the Xilinx KC705 Kintex-7 evaluation board for implementation. This paper mainly focuses on the design of an SVM linear classifier with pipelined architecture for FPGA implementation. The FPGA-based two-class SVM classifier can perform fast data classification due to the advanced parallel calculation feature provided by FPGA. The classification system operates in a linear fashion. The simulation and synthesis results show that the SVM linear classification system can be able to classify data effectively.

Keywords Breast Cancer · SVM · IEEE 754 format · Vivado tool · Verilog HDL · FPGA

✉ H. S. Laxmisagar
sagar8.hs@bmsit.in

[1] Department of Electronics and Communication Engineering, BMS Institute of Technology Management, Bengaluru, India

1 Introduction

Breast cancer is seen as a serious threat to women's lives and health. It is one of the most common types of cancer in women around the world [29]. Breast cancer is the most predominant cancer among Egyptian women and is found to be 29% of National Cancer Institute cases. The lump is limited to the Breast in the early stages, but it can move to lymph hubs [24] in the armpit and other parts of the body such as the bones, liver, lungs, and cerebrum in the later stages. The illness is treatable in the early stages, but it is not recognized until later stages, which is the primary reason for the death of so many women around the world. Breast cancer detection at an early stage is critical for successful treatment and to reduce the mortality rate. Breast cancer affects the majority of women worldwide today compared to males as per World Health Organization. Breast cancer is the leading cause of death in women among all cancers. In the United States, 12.15% of women [27] will develop this cancer during their lifetime. For the last two decades, machine-learning researchers have been working on improving classifier effectiveness. At the same time, machine-learning research has resulted in a new generation of cutting-edge supervised machine-learning [14] classification algorithms, such as support vector machines (SVMs).

SVMs are a powerful supervised learning tool that offers cutting-edge accuracy at the expense of high computational complexity [26]. The linear SVM classification system suffers from the dimensionality of the problem and linear dependencies on the number of Support Vectors. SVMs have demonstrated [4] high accuracy and have been widely used in recent years. The standard versions of the SVM algorithm, on the other hand, are very time-consuming and computationally intensive, presents a challenge to investigate other hardware architectures than CPU. SVMs are a popular supervised learning method that provides cutting-edge accuracy in a variety of classification tasks. SVM classification system is a robust [25] supervised machine learning algorithm that achieves good accuracy rate with several classification problems and is extensively used in embedded applications [2]. However, due to the inherent complexity of the computations required, embedding SVM classifiers is difficult. This motivates the SVM to be implemented on hardware platforms in order to achieve high-performance computing at a low cost and power consumption.

FPGAs have confirmed substantial performance and acceleration with competent hardware execution results for a variety of applications, overtaking other comparable platforms such as general-purpose processors and graphics processing units (GPUs) [3]. As a result, FPGAs have become increasingly popular for realizing and accelerating SVM on hardware for online and embedded classification. These FPGA performance characteristics make them ideal for implementing SVM linear classifier algorithms, and they are thus responsible for overcoming some of the critical limitations and challenges of other research methods.

The major contribution to the research work are as follows:

- Development of an SVM Linear Classifier system with pipelined architecture for fast data classification.
- The main components of the proposed architecture include a computing modules such as a 32-bit floating point adder and multiplier, BRAM IP configured for storing and processing data, Adder tree, and decision device.
- The Verilog hardware description language (HDL) is used for register transfer level (RTL) modeling style which is adopted in this work conforms to industry standards.

- Performance enhancement techniques such as pipelining and parallel processing are incorporated into each of the modules in order to increase the throughput of the system.
- Entire design is implemented for Xilinx FPGA device, KC705 Kintex-7 evaluation board.
- The proposed framework is capable of achieving a maximum delay of 1.391 ns and a minimum delay of 0.577 ns after applying a timing constraint of about 2.00 ns by creating "timing.xdc" file.
- The accuracy of prediction about 91% in linear one has been achieved with maximum clock frequency of 100 MHz.

The rate of accuracy can be greatly increased by incorporating exact dot product accumulator compared to other implementation schemes. In Section 2 describes the background concept related to Breast cancer, Support vector machine and FPGAs, which are necessary for understanding the proposed system for the detection of Breast cancer. The algorithm implementation with mathematical modeling and flowchart is discussed in Section 3. With reference to Section 3 architecture implementation with floating point multipliers and adder tree is discussed in Section 4. In Section 5, the proposed design simulation, synthesis and implementation are presented in more details. Finally, a brief conclusion and future implementation is presented in Section 6.

2 Existing work

Selvathi et al. [30] proposed integrating a computerized diagnostic system for breast cancer detection in FPGA using Artificial Neural Network (ANN). The features of the Wisconsin Breast Cancer Database are used to train and test the Multilayer Perceptron Neural Network (MLPNN). The Verilog coding is done to get the netlist of ANN and synthesized by FPGA Virtex 5 board. Finally, the netlist is mapped to the FPGA, and the hardware functionality is tested. The proposed system has a correct classification rate of 90.83%. To appreciate melanoma detection on a chip, Lopes et al. [19] proposed a hardware/software co-design to execute the SVM classification system onto FPGA. The SVM implementation on a hybrid FPGA (Zynq) platform using the recent Ultra-Fast High-Level Synthesis design methodology to achieve good accuracy. The hardware implementation results show a 97.9% classification accuracy and a significant hardware acceleration rate of 21 with resource utilization of 3% and power consumption of about 1.69 W.

Baez et al. [6] proposed the multiclass SVM classifier algorithm, which is designed using C programming and implemented on both ZC7020 (ZedBoard) and ZC7045 (ZC706) devices. The classifier was tested on hyperspectral images from a brain cancer database. This system not only decreases the resource utilization by using less than 20% of the FPGA and required power consumption by 23% when compared to the full implementation.

Mohammadi et al. [23] proposed a hardware implementation of a fully parallel SVM using Stochastic Gradient Descent for FPGA implementation. The FPGA implementation of an SVM classifier with SGD attains speedups of about 10,000 when compared to a software approach running on a quad-core processor and up to 319 related to state-of-the-art FPGA implementations while using less hardware resources utilization. The implementation results demonstrate that the suggested method is a feasible solution for very challenging difficulties like big data analysis.

Fiolhais et al. [16] proposed a fused dot product processor with exact operand arithmetic that can output an incomplete result per cycle. To achieve full accuracy, the processor employs a long interval accumulator with full fixed-point precision. The proposed processor was built in a Zynq 7020–1 FPGA, in which a single-precision core can perform with an output of one partial result per 90 MHz clock cycle and occupies approximately 5 K LUTs. Because of the complexity of their algorithms, floating point operations are difficult to implement on FPGAs. Many scientific problems, on the other hand, necessitate the use of floating point arithmetic with high levels of accuracy in their calculations. Bassoli et al. [8] investigated FPGA implementations of IEEE single-precision floating-point addition and multiplication. Customizations were made in order to save the chip area. The binary parallel adder and multiplier, which is a digit-serial multiplier, was proposed by Louca et al. [20]. Prototypes were built on Altera FLEX8000s, with peak rates of 7 MFlops for 32-bit addition and 2.3 MFlops for 32-bit multiplication.

One of the most significant limitations of current embedded GPUs, as addressed by the majority of researchers [10, 28, 32, 34], is the limited existing resources like limited memory, cache, registers, and cores). As a result, GPUs are challenging to deploy in embedded environments, motivating a shift toward FPGA implementations. One of the most significant limitations of existing embedded FPGAs [4] is their major hardware development effort and time-to-market. The comparison of related results about developments of the existing algorithm in different areas is tabulated in Table 1.

Numerous researchers [1, 5, 22] have contributed their work to the implementation of the SVM machine learning algorithm, which is becoming increasingly popular in research due to the promising opportunities it provides in this field. This method is developed and used in a variety of statistical and intelligent fields, including regression, time series analysis, pattern recognition systems, and so on. It has been widely used in a variety of fields, including bioinformatics, text and document classification, pattern recognition, and image recognition. Kim et al. [18] proposed a single-precision floating-point (SPFP) multiplication system to reduce the number of logic blocks. The suggested work estimates only one of the operands and iteratively compensates the calculation error to achieve acceptable error ranges in applications. Bassoli et al. [7] proposed a new floating-point accumulation circuit based on a model. The model was implemented in Simulink, and results of simulation showed that it performed better in terms of speed and occupied area than other models.

Most of the author proposed CAD system development for the detection of Breast cancer and very few papers report FPGA implementation for detection of various types of cancer. Lopes et al. [19] implemented a linear SVM classifier onto an FPGA to recognize detection of melanoma on a chip. Very few researchers have been reported power consumption results, making it difficult to achieve a low-power embedded system. As per our knowledge no paper is found related to implementation of efficient SVM classification system for the detection of Breast cancer onto FPGA and most existing implementations other than Breast cancers are not realized on modern FPGAs board. As it is evident from the literature survey of state-of-the-art methods. There is a trade-off between classification performance, speed and hardware resource usage, which necessitates more optimum solutions. There is a primary trade-off between preserving increase in classification accuracy and meeting embedded system constraints. Due to the lack of flexibility during off line training using MATLAB or some other software methods, the challenging task of implementing the linear support vector machine (SVM) classifier with pipelined floating point adder and multiplier for fast processing for FPGA implementation is proposed. In this article, linear SVM classification system is designed for diagnosing breast cancer problem and is implemented for Xilinx KC705 Kintex-7 FPGA using

Table 1 Comparison of related results about developments of the existing algorithm in different areas

Author	Technique	Application	FPGA Device	Accuracy in %	Limitations
Zachariah Carmichael et al., 2019 [12]	DNN Using the Posit Number System	Data classification	Xilinx Virtex-7 FPGA device	Not Computed	DNN accelerators support low-precision numerical formats. 8-bit posits with 8-bit IEEE 754 format
Selvathi, D et al. 2018 [31]	ANN implementation using FPGA	Breast Cancer Classification	XILINX Virtex 5 FPGA device	90.83%	The area, power and timing optimization is not done
Batista et al. 2020 [9]	Pseudo-logarithmic number representation	Data classification	XILINX Virtex 5 FPGA device	–	The area, power and timing optimization is not done Accuracy rate not computed Expensive multipliers with simpler adders are used in the design
Hamed F et al. 2019 [13]	Deep Neural Networks	Data classification	Xilinx Virtex-7 FPGA device	Not Computed	Accuracy is not measured Area, power and timing optimization is not done 16-bit posits with 16-bit IEEE 754 format

Verilog HDL. It is suggested for future implementation that instead of performing rounding as number of multiplier which results in lot of inaccuracy just perform one time rounding after all multiplied data is loaded into an exact dot product accumulator of size 32-bit x number of floating point multipliers.

3 SVM classifier algorithm

3.1 Mathematical modeling of SVM algorithm

The main goal of the linear SVM Classifier is to develop a more efficient computational method of separating hyperplanes between several classes of data through learning during training phase [33]. As shown in Fig. 1, SVM can be used to identify hyperplanes between different classes of data and looks like a linear function in a high dimensional feature space. The hyperplanes are separated such that the distance between the classes should be wide to increase accuracy rate.

The classification function [21] needs to be implemented which is very much essential during data training of the SVM algorithm. The SVM algorithm for implementation can be expressed with several equations but to implement a more efficient model particular equations need to be presented.

$$f(x) = W^T x + b \tag{1}$$

where,

'W' denotes a weight vector and normal to the line.
'b' is the bias value.

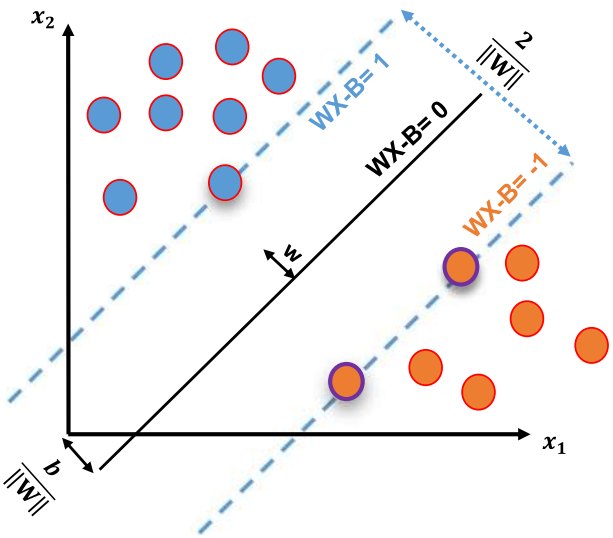


Fig. 1 Support vector machine separating Hyperplane [35]

In case of thirty features in 3D it looks like a plane called hyperplane. For the linearly separable data $\boldsymbol{X_i}$ labelled into two categories y_i = {−1,1} as per WBCD standard dataset, weight vector 'w' can be calculated such that the discriminant function separates the categories for i = 1, .., N.

The support vector machine (SVM) prediction rule is given by,

$$Y = sign\left(W^T x + b\right) \tag{2}$$

The SVM learning happens through input training data (x_i, y_i) = {(x_1, y_1),(x_2, y_2),…..,(x_n, y_n)} which consist of samples taken from two classes of data labeled as 'M' and 'B'. The distance between two lines can be computed by considering the following two equations which helps in optimization.

The two hyperplanes 'l1' and 'l2' can be expressed as:

$$l1 : W^T x + b = 1 \tag{3}$$

$$l2 : W^T x + b = -1 \tag{4}$$

If 'b' value becomes equal to zero, then line passes through origin.

The maximum distance between two hyperplanes can be calculated using Eqs. (3) and (4) as follows,

$$Max\left(\frac{2}{\|W\|}\right) \tag{5}$$

Then the problem of Maximizing the margin is same as minimizing ||w||.

Optimization formulation:

In order to perform optimization with N linear inequalities constraints substitute ||w|| with $\frac{1}{2}$ ||w|| ^{2}, then the constrained optimization function can be expressed as:

$$\underset{w,b}{\text{Minimize}} \quad \frac{\|W\|^2}{2} \tag{6}$$

The above equation is called quadratic objective function subject to:

$$y_n\left(W^T x_n + \mathrm{b}\right) \geq 1, \mathrm{n} = 1, 2, 3, \ldots\ldots\ldots, \mathrm{N} \tag{7}$$

It is possible to solve constrained optimization problem using Lagrange multipliers [17, 36] to get the classification function:

$$\mathrm{L(W)} = \frac{\|W\|^2}{2} - \sum_{n=1}^{N} \alpha_n g_n(w, b) \qquad \alpha_1, \alpha_2 \ldots\ldots\ldots, \alpha_n \geq 0 \tag{8}$$

To solve the SVM classification problem with PFGA, it is required to create a compatible algorithm expression rather than using the equations and functions.

3.2 FPGA implementation of linear SVM classifier

For FPGA implementation, the UCI Machine Learning Repository for breast cancer dataset is used. The physician Dr. William H. Wolberg at the University of Wisconsin Hospital in Madison, USA, created the WBCD dataset. The data need to be analysed to select training data set of about 80% and test data of about 20%. It is required to perform conversion from real number into IEEE 754 single precision floating point format.

The floating-point unit is designed in IEEE754 single precision floating-point format and the hardware design is not dependent on dataset. However, the weight value is computed for each feature using python script executed in online Jupyter notebook. In Vivado simulation tool IP need to be configured for loading patient data and weight value in block RAM for further processing. The IP need to be configured for differential clock in order to perform write/ read operation in block RAM w.r.t positive/ negative edge of clock input by decoding address lines addr1 and addr2. The ROM controller is designed to control read / write operation into BRAM. The proposed FPGA design flow implements the most efficient adder tree. The design flow of the SVM linear characteristic is presented in Fig. 2. It is required to implement floating point adder and multiplier because hardware does not understand input real numbers. The pipelined four stage multiplier and six stage adder is to be designed for faster processing. The adder tree is designed in order to compute the inner product and helps in classifying two classes of data. The design of adder tree is not dependent on dataset characteristics. The output of the adder tree is added with bias value 'b' results in making decision by decision device comparing with threshold value '0'. The decision device output may be positive, negative or zero. If the SVM output is positive or equal to '0' then decision is taken w.r.t 'Malignant'. Similarly, if the SVM output is negative or less than '0' then decision is

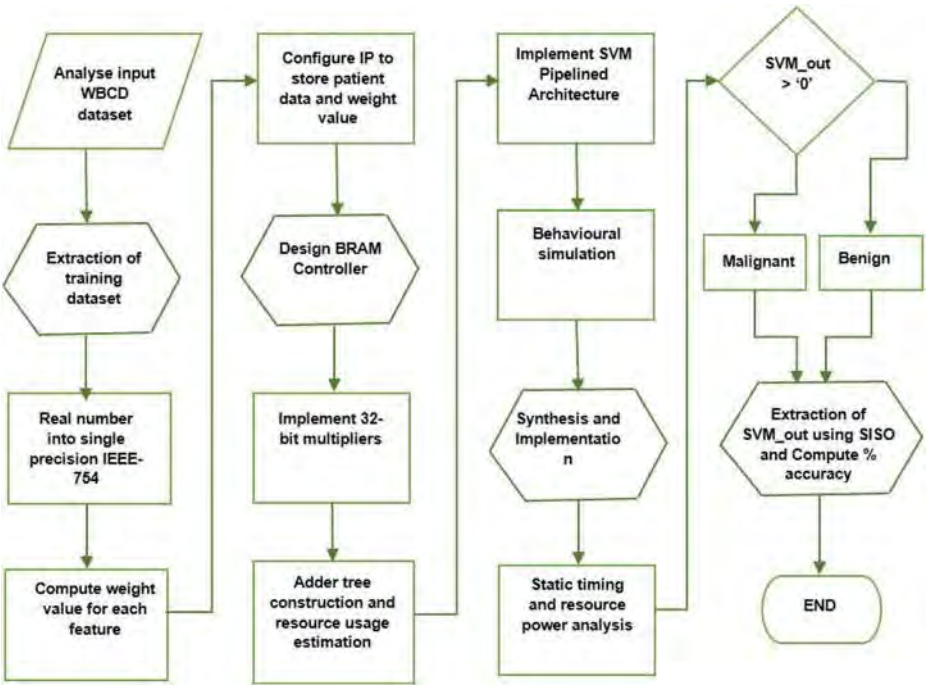


Fig. 2 The Proposed flowchart for FPGA implementation of linear SVM classifier

taken w.r.t 'Benign'. This results in more inaccuracy due to comparing with only positive, negative and '0' for classification. In order to overcome from this problem confusion matrix, need to be incorporated in order to increase the percentage accuracy by considering true positive or negative and false positive or negative. The SVM output data can be extracted from Vivado simulation tool by designing SISO architecture in Verilog and then the obtained data is in row vector may be mapped into column vector for classification. The comparison is done by writing program in MATLAB and to compute percentage accuracy.

3.3 SVM linear classifier algorithm

SVMs are supervised learning models [15] that analyses data for classification and regression analysis in machine learning. The SVM algorithm was developed by Vladimir Vapnik at AT&T Bell Laboratories. The pseudocode of SVM linear Classifier algorithm is illustrated in the following section.

The SVM Linear Classifier Algorithm

Input ports declaration:	clkp, clkn, reset
Output ports declaration:	svm_out
Sub module instantiation-1:	clk_wiz_0 for differential clock
Sub module instantiation-2:	rom_cn1
Sub module instantiation-3:	patient_data_rom
Sub module instantiation-4:	svm_weights_rom
Floating point multiplier:	Parallel processing of thirty multipliers and multiplication happens through four stages. **Stage-1:** Exponent calculation **Stage-2:** Multiplication of Significands **Stage-3:** Normalize the product **Stage-4:** Round off
Floating point adder:	Pipelined operation is implemented for fast processing and adder tree operation happens through six stages. **Stage-1:** unpacking operation **Stage-2:** Floating point align and swapping operation **Stage-3:** Floating point invert operation **Stage-4:** Normalize the output from stage-3 **Stage-5:** Round off the output from stage-4 **Stage-6:** Packing operation
SVM output: Last stage output of adder tree is compared with decision value to get SVM output	

4 Proposed architecture for SVM linear classifier algorithm

The proposed algorithm for classification of WBCD dataset to identify Benign or Malignant.

(1). The first step is to analyze WBCD dataset.
(2). The second step is to alter real number into IEEE 754 single precision format for floating point representation.
(3). In the third step the patient data and weight value is loaded into block RAM by loading two coefficient files such as 'patient_data.coe' and 'weigts.coe'.

(4). In fourth step the design of floating point adder, multiplier and ROM controller.
(5). In the fifth step linear SVM architecture need to be implemented with adder tree.
(6). Final step is to add bias value to the output of the adder tree and then perform classification into benign and malignant.

Further, it is investigated that in order to reduce the error due to rounding and to increase percentage accuracy with additional hardware cost, it is required to design an exact dot product accumulator in between multiplier and adder tree to store the output of each multiplier without rounding. It is suggested for future implementation that instead of performing rounding as the number of multiplier which results in a lot of inaccuracy just perform one time rounding after all multiplied data is loaded into an accumulator of size 32-bit x number of floating point multipliers.

4.1 System architecture

The proposed FPGA architecture for the SVM linear classifier for the detection of Breast cancer is depicted in Fig. 3. The training dataset in SVM consisting of two input vectors such as patient data vector range from Data_f [959:928] to Data_f [31:0] and weight value corresponding to each feature range from Coeff_w [959:928] to Coeff_w [31:0] is loaded into internal FPGA memories. There are total of 30 different features are present in WBCD dataset, which is very much essential for detection of Breast cancer. There are total of fifteen 32-bit multipliers are required to compute multiplication between two 32-bit binary numbers. All the multiplier units work in parallel. The multiplication happens through four stages and after each multiplication rounding has to be performed for 32-bit precision. The accuracy reduces slightly as the number of rounding increases. The output of the multiplier is given as input to the adder tree which performs addition of intermediate results. The addition happens through six stages and the output of the adder tree is

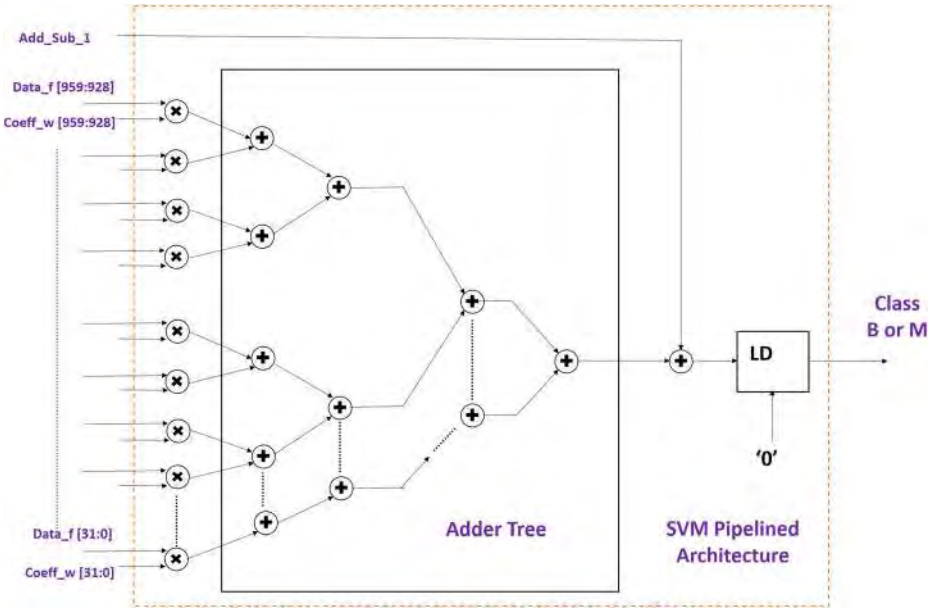


Fig. 3 SVM linear classifier architecture

added with bias value equal to '1'. The SVM model identifies the data points Support Vectors (SVs) that can be used to create a separation model for identification of different classes. The main goal of the linear SVM Classifier is to develop a more efficient computational method of separating hyperplanes between several classes of data through learning during training phase.

The architecture's adder tree constructs the classification function f(X) with the bias 'b' and polynomials such as patient data and weight value. Once testing data is passed into the FPGA board, the trained model uses it to perform a classifying procedure. The results are compared using decision function to obtain either benign or malignant. To complete the calculation in order to obtain SVM output, this design employs 1 sign bit, 8 bits exponent, and 23-bits mantissa.

The WBCD dataset is used for accelerating data using FPGA. A total of 80% of the data is used for training and the remaining 20% of the data is used for testing. The 20% of data for testing is separately loaded into BRAM with patient coefficient data and corresponding weight value for each feature as shown in Fig. 4. Later once the model is trained to do the correct classification, it is tested with a test value.

4.2 Design of BRAM controller

The design of block RAM controller is very much essential to load the both Breast cancer patient data and as well as computed weight value using python script corresponding to each feature in Wisconsin Breast Cancer Dataset. The loaded data can be used as input further by designing linear SVM classifier for creating training model to classify two classes of data. The single ended clock in XILINX kC705 can be generated by configuring clk_wizard IP in Vivado simulation tool. In Fig. 5, it is observed that single ended clock 'clk' can be generated using two differential clock inputs 'clkp' and 'clkn', which can be used to perform write and read operation into Block RAM. It is required total 960 bits of data storage elements to store each patient data. Because each feature value can be represented with 32 bits of data.so there are total 30 features are present in WBCD standard dataset. The total number of bits required for each patient is 32-bits multiplied with FN (Total number of feature).

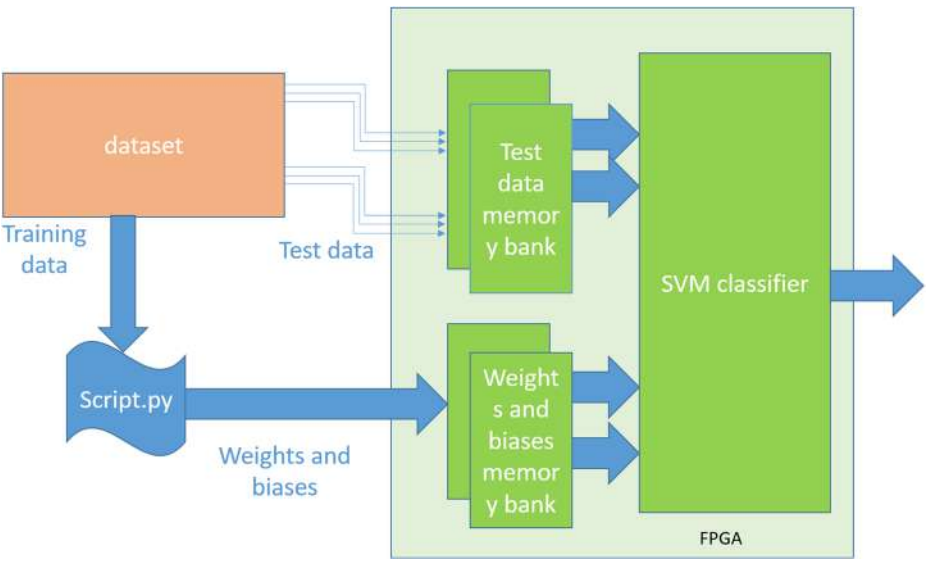


Fig. 4 SVM classifier system with BRAM for training and testing

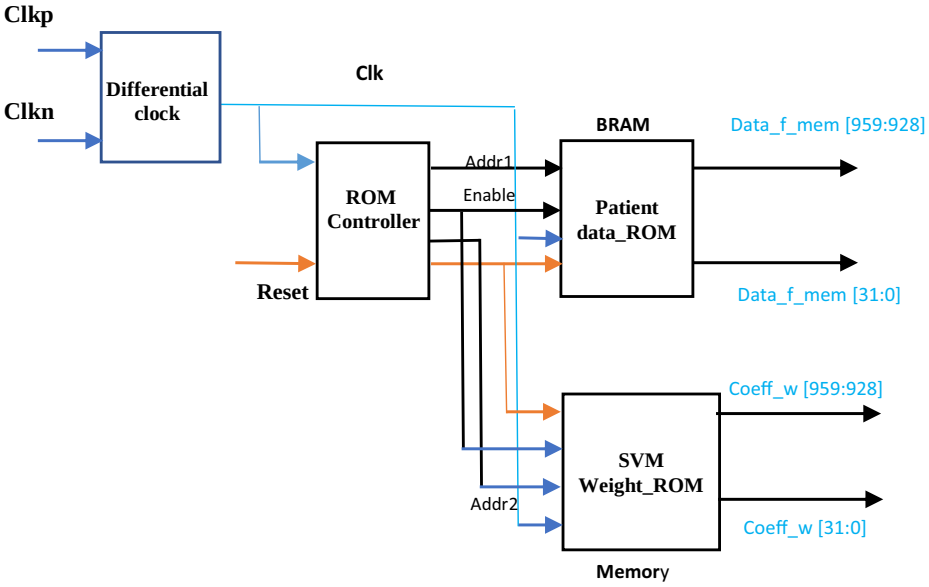


Fig. 5 Design of BRAM controller

The total memory requirement to store training data:

Memory requirement in terms of bits = 32-bits * FN * Total number of patients.

Two address lines are used such as addr1 and addr2 which is 10 bits wide to access BRAM for performing read or write operation for further processing to classify between two classes of data.

4.3 Computing modules for linear SVM classifier

The computing modules designed for the system are 32-bit signed floating-point adders and multipliers with 1 sign bit, 8-bit exponent and 23-bit mantissa. A floating point package is used to build the computing units as shown in Fig. 6. The floating-point math packages are based on the VHDL 1164 numeric_std package and use the package's signed and unsigned arithmetic. Because the numeric_std package is well supported by simulation and synthesis tools, they are extremely efficient.

By applying floating-point package to designed computing units, the computational accuracy is slightly dropped because of the restriction of the fixed 32-bits. This error is due to rounding after each multiplication and further it is investigated that by using dot product accumulator after multiplication rounding error can be reduced. Instead of using rounding after each multiplication store the multiplied output in the accumulator and at the end perform rounding which results in reducing the error and improvement in accuracy. Overall, the total computational accuracy is satisfied while the computation resource consumption is limited within an acceptable range.

5 Experimental results and discussions

For FPGA implementation, the UCI Machine Learning Repository for breast cancer dataset is used. The physician Dr. William H. Wolberg at the University of Wisconsin Hospital in

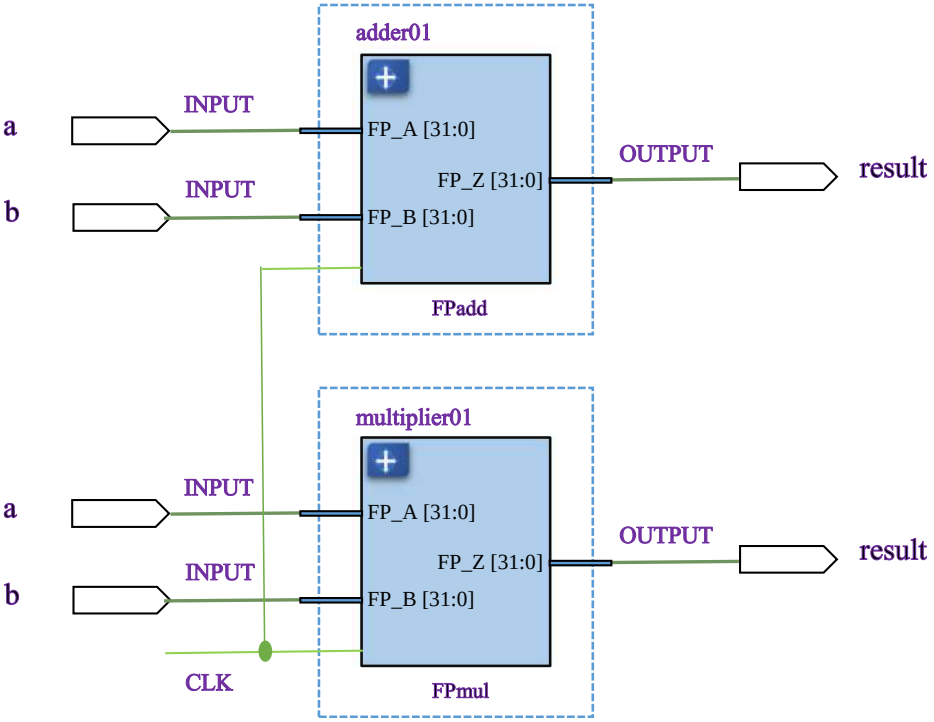


Fig. 6 Design of floating point multiplier and adder

Madison, USA, created the WBCD dataset [11]. For each image, the standard error, mean and "worst" or largest values of these features were computed, yielding 30 features. For example, third field represents Mean Radius, thirteenth field represents Radius SE, and twenty third field represents Worst Radius.

This study seeks to determine which features are most useful in identifying malignant or benign cancer. The objective is to determine whether the breast cancer is benign or malignant. Here, the standard WBCD dataset is analyzed for 569 patients with 30 best features to categorize between two classes of data for the detection of Breast Cancer. The Summary of the Testing and Training data is as shown in Table 2.

5.1 Simulation output waveform analysis of top module of SVM linear classifier

The pipelined SVM architecture is designed using Verilog HDL, and then it is synthesized by Vivado (ver. 2017.4) simulation tool. The design is implemented through translating, mapping, placing and routing, and generating bit stream. At the end, it is configured to the Xilinx

Table 2 Summary of the testing and training data

Training in %	Benign	Malignant	Testing in %	Benign	Malignant
80	268	186	20	89	25
70	169	148	30	133	39
60	133	139	40	174	55

KC705 Kintex-7 evaluation board. Given below are the simulation results for the single precision IEEE standard 754 based SVM linear classifier algorithm.

The Simulation result of SVM linear classifier depicts data loaded into BRAM for training with respect to positive edge of clock as shown in Fig. 7. The initial delay for clock from 0 ns to 1000 ns is called as clock sink delay, it is due to PLL. The width of addr1 and addr2 is required 10 bit to accommodate 569 patient's data stored in BRAM to perform read / write operation. There are 30 features present in WBCD dataset and each feature is represented with 32-bit floating point number. So total bandwidth requirement to accommodate each patient data is 960 bits wide. There are two coefficient files with file name extension '.coe' is created to load the data into BRAM after configuration.

The SVM output obtained in simulation window as shown in Fig. 8 represented with hexadecimal number representation indicated by placing a marker. The SVM output in hexadecimal representation is '3f3994f2'. It has to be converted to equivalent decimal value using IEEE-754 Floating Point Converter in order to perform comparison manually for understanding purpose. The obtained decimal value after conversion is '0.724928975105' refers to positive value. So, the SVM classified output considered as 'Malignant' corresponding to the address which is decoded is '21' and the corresponding patient ID as per standard WBCD is 842,302 which belongs to the class Malignant.

The floating point 32-bit number multiplication between patient data and weight value is executed in parallel, which produces output at 4100 ns as shown in Fig. 9. The floating point 32-bit number multiplication happens through four stages and addition in an adder tree happens through six stages in pipelined. The floating point adder and multiplier takes three clock cycles for execution.

The simulation window shown in Figs. 10 and 11 depicts pipelined adder tree operation with different stage of operation. The adder tree intermediate output at different stages is indicated with multiple marker line in the simulation waveform.

The adder tree operation happens through six stages as depicted in the following table.

The adder tree operation happens through six stages as shown in Table 3 and takes three clock cycle delay, which is about 300 ns between each stage to obtain the adder output.

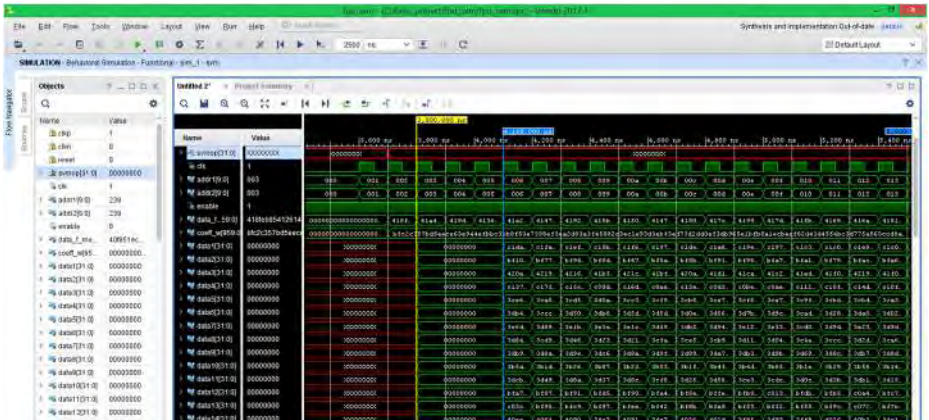


Fig. 7 The output waveform for top module of SVM classifier: depicts data loaded into BRAM at 3700 ns

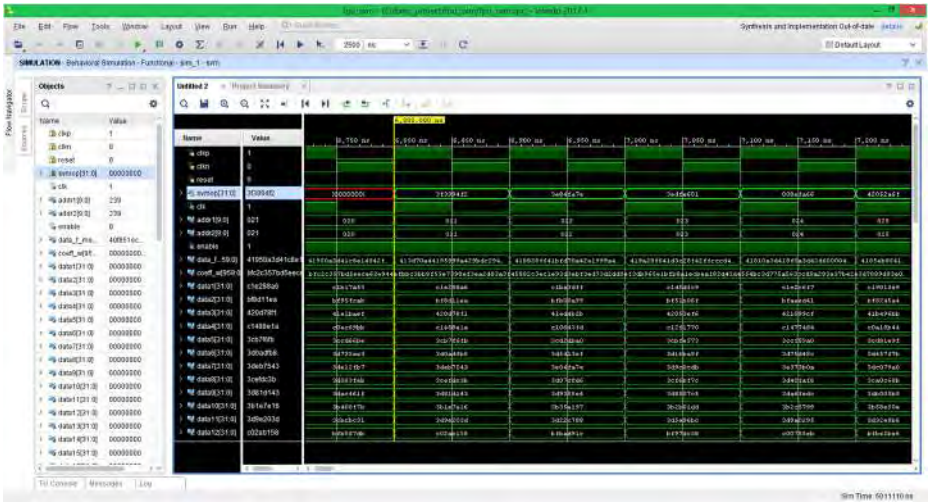


Fig. 8 Simulation result of SVM output set as top level module obtaining at 6800 ns

5.2 Logic synthesis of top module SVM classifier system

It becomes very easy to check and analyse the different metrics such as area, power and timing after synthesis and implementation. The implementation result is more accurate to evaluate performance analysis of the system. The Fig. 12 illustrates the project summary starting from synthesis to final FPGA implementation.

In Vivado simulation RTL analysis is done for the proposed system by exploring schematic view by elaborating design after behavioral simulation is depicted as shown in Fig. 13. The design before optimization is represented in terms of generic symbols such as multipliers, adders, counters, OR gates, and AND gates that are independent of the FPGA device.

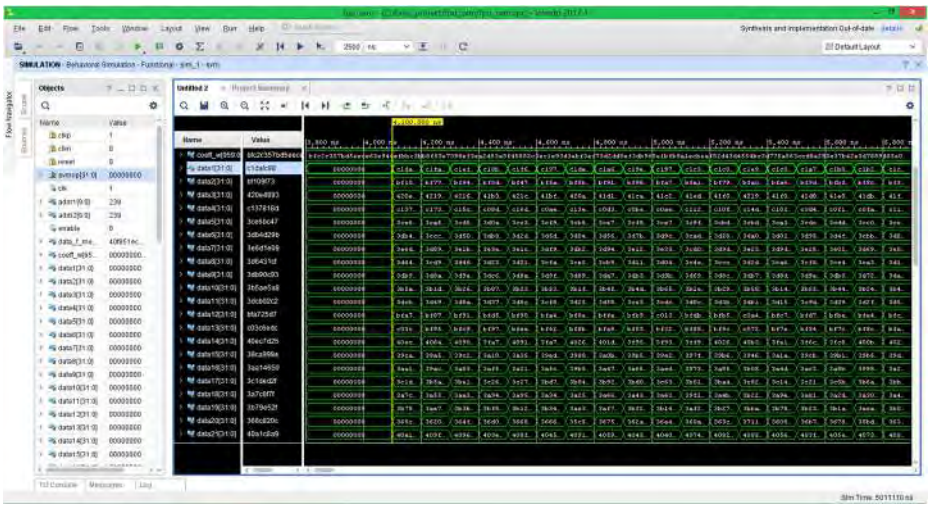


Fig. 9 The Output waveform for top module of SVM classifier: floating point multiplier output at 4100 ns

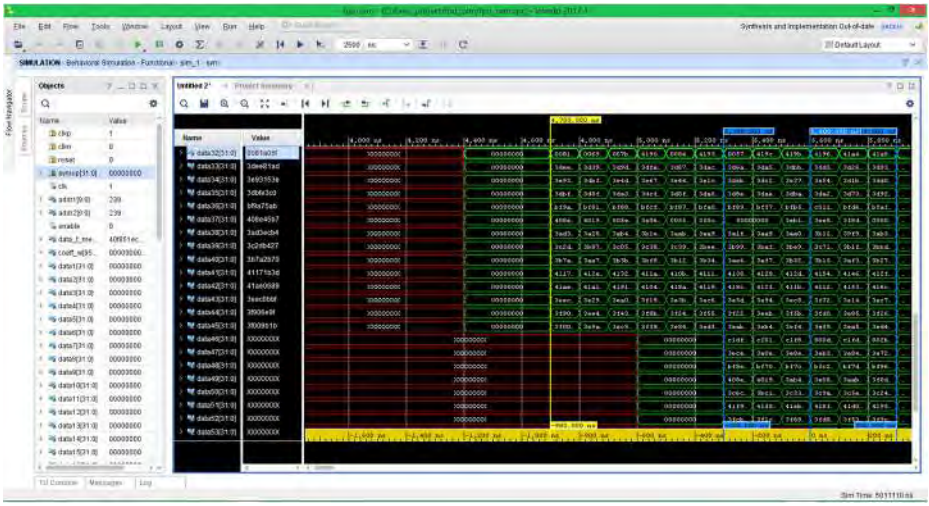


Fig. 10 Output waveform for adder tree pipelined operation producing Stage-1 output at 4700 ns

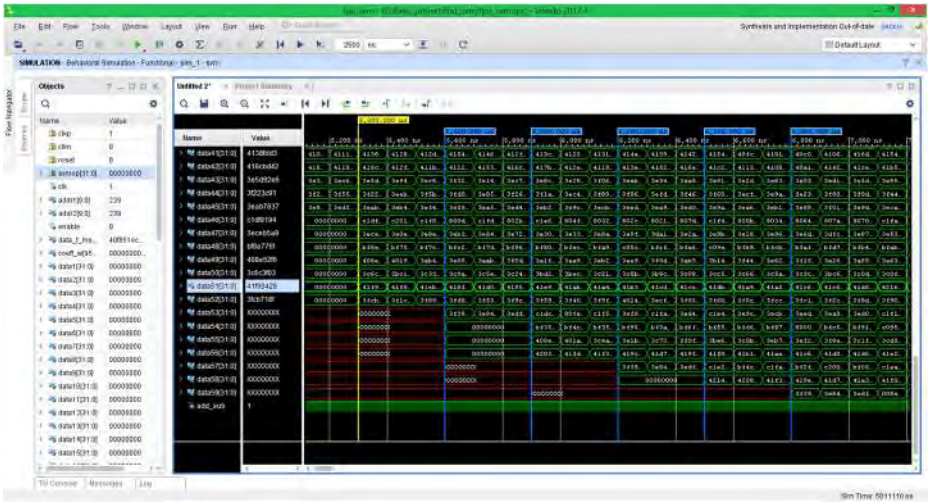


Fig. 11 Simulation waveform of adder tree pipelined operation in SVM linear classifier at different stages

Table 3 Operation of adder tree through six stages

Total number of Stages	Intermediate output	Simulation output
Stage 1	Data 46 to data 52	5300 ns
Stage 2	Data 53	5600 ns
Stage 3	Data 54 to 56	5900 ns
Stage 4	Data 57	6200 ns
Stage 5	Data 58	6500 ns
Stage 6	Data 59 (SVM output)	6800 ns

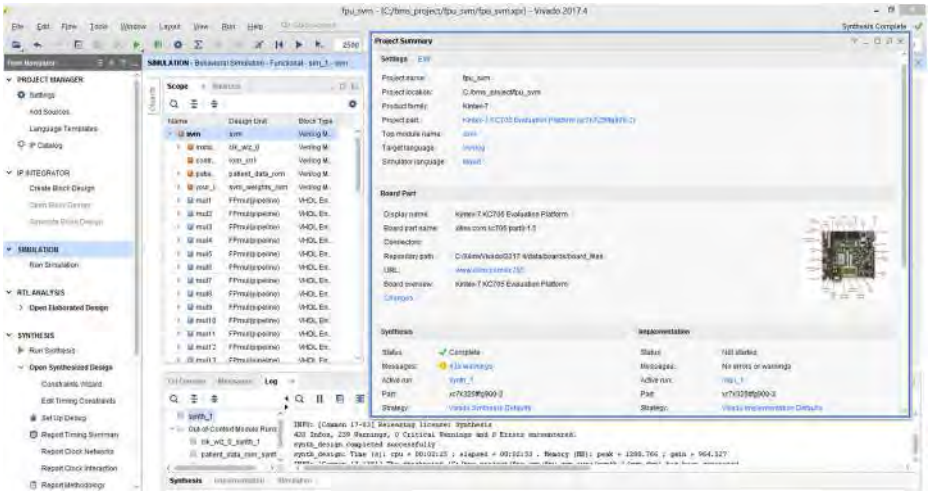


Fig. 12 Project summary of synthesis and implementation

The schematic is a graphical representation of the netlist. The total number of cell are present after synthesis is 64. The 'clkp' and 'clkn' are differential clock inputs and the global reset signal such as 'reset' is used to reset the system. The designed architecture in Section 4 matches with the captured schematic diagram and the signals presented in the simulation window. In RTL schematic view each small modules interconnected to top module in terms of blocks.

In Vivado simulation schematic analysis is done for the proposed system for to view graphical representation of the netlist and better understanding of the design by exploring

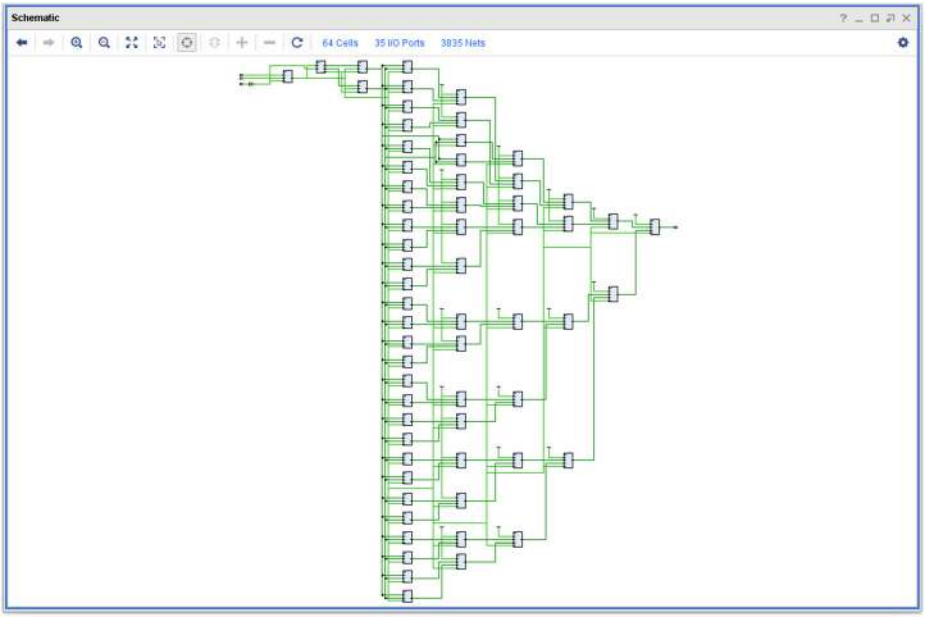


Fig. 13 RTL schematic view of the SVM linear classifier

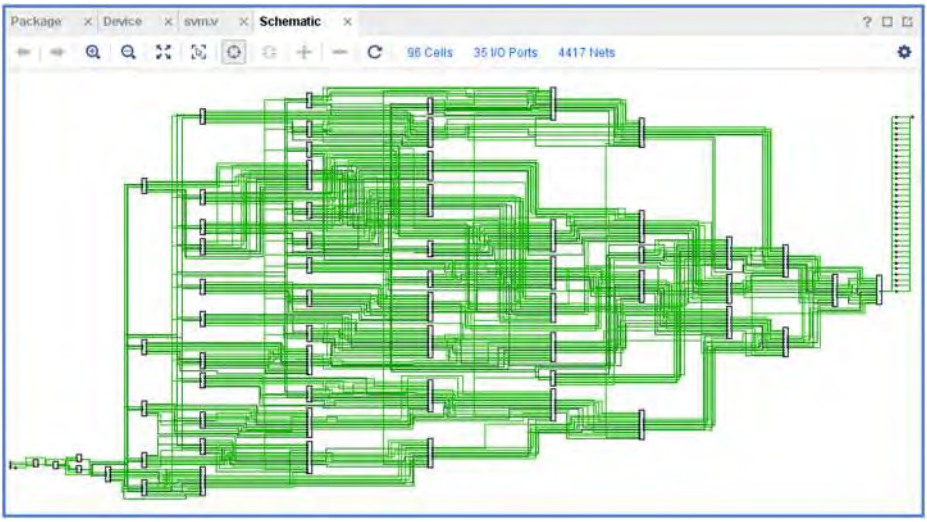


Fig. 14 Schematic view of top module SVM classifier after synthesis

schematic view after performing synthesis as shown in Fig. 14. The gates generated by the synthesis tool are visible in the Synthesized design. The schematic is the simplest way to visualise the gates in a timing path after implementation. The Schematic window displays the selected logic cells or nets. The total number of cell are present after synthesis is 96.

The netlist which is generated after performing synthesis by configured to the Xilinx KC705 Kintex-7 evaluation board is depicted as shown in Fig. 15. The various net and system components connections which is made virtually for selected device is very complex for analysis. The details of each net and component can be visualized by highlighting with chosen color for understanding purpose. The selected net which can be seen at RHS i.e. A_EXP [29] and A_EXP [23] can be visualized with white routing marking. The implemented design details of device on Xilinx FPGA XC7K325tffg900–2 is shown in Fig. 16.

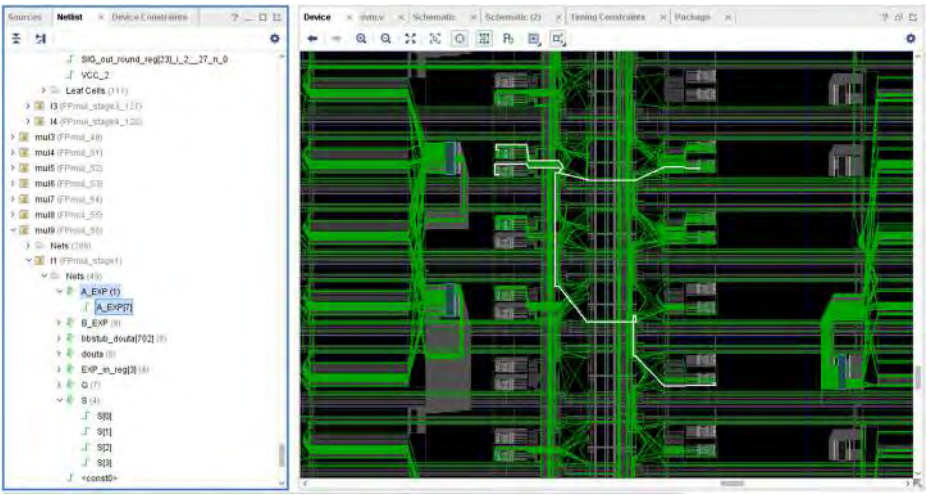


Fig. 15 The Netlist is generated after synthesis

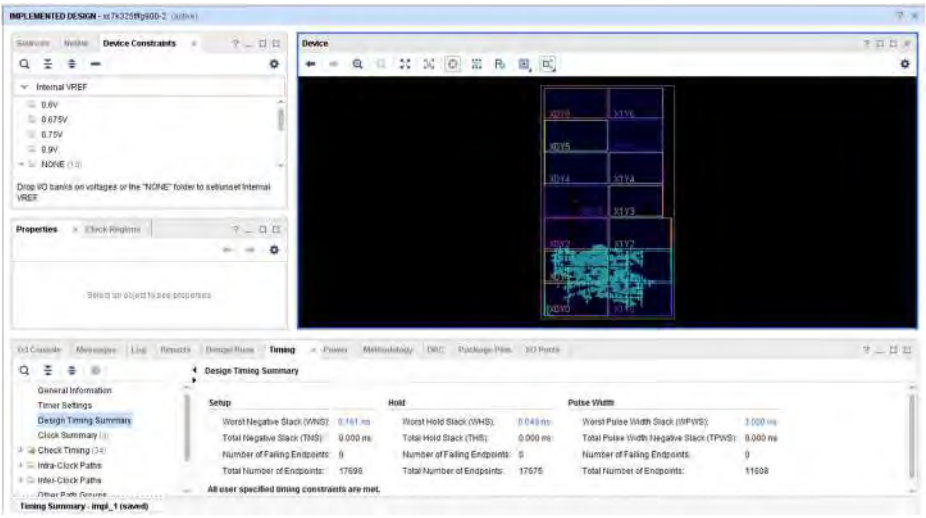


Fig. 16 Implemented design details of device on xilinx FPGA XC7K325tffg900–2

The total power estimation from synthesized netlist depicts total on-chip power is about 1.174 W and the dynamic power is about 1.007 W is as shown in Fig. 17.

The BRAM configured to store coefficient files related to patient data vector and weight vector for further processing, it is utilizing 12% of resources compared to other resources. The total device utilization in percentage is about 49% within a band of 25% resource utilization after implementation as shown in Fig. 18. The summary of device utilization after post-synthesis and implementation is depicted in Fig. 19. The graph is plotted corresponding to each device utilization. The LUT utilizing resources is about 9% within a band of 25% of total device utilization after post-synthesis.

The power estimation result after post-synthesis and implementation is shown in Fig. 18. The "On-Chip Power" report is comprised of dynamic power and static power. The Static power for proposed design after post implementation is 0.124 W, which is 14% of total on-chip power due to power consumption in the steady state resulting from intrinsic transistor

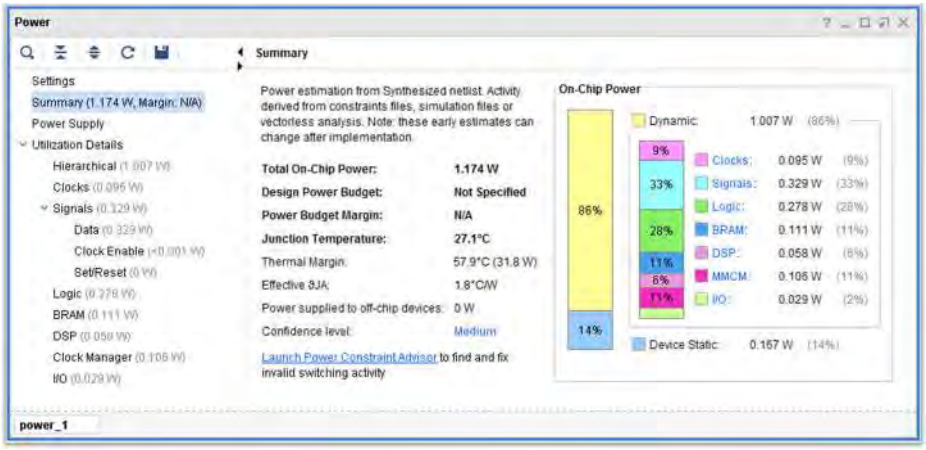


Fig. 17 Snapshot for power estimation from synthesized netlist

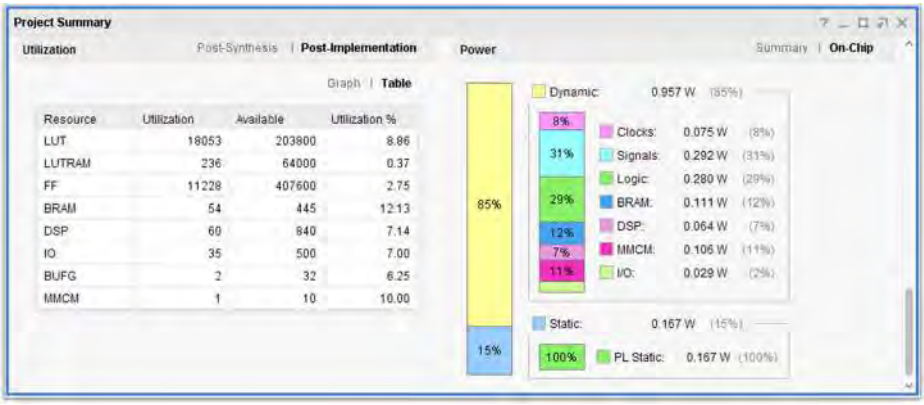


Fig. 18 Resource utilization and power estimation from post-implementation

leakage. It may depend on various parameters like voltage, device process and temperature. The static power is a constant value under the same condition which is independent of logic design.

The dynamic power for the proposed design is 0.957 W, which is 85% of the total on-chip power due to the custom design's internal switching activity, which varies with each clock cycle. When the voltage is constant, dynamic power is proportional to the amount of active resource and the clock frequency. The power consumed due to clock propagation signals is 0.075 W. The power consumed by the combinational and sequential logic circuits is 0.280 W. The signal power is nothing but the power consumed by the wiring made inside is 0.292 W. The total on-chip power consumed for the proposed system after post implementation is 1.124 W.

The "Slack" simply means the difference between the expected arrival of a signal and the actual arrival of a signal. In Vivado simulation tool there are three sets of checks can be performed which are Worst negative slack (WNS), Worst pulse width slack (WPWS) and Worst hold slack (WHS). In all the three different checks negative means fails to attain timing constraints and positive means pass to attain timing constraints. The proposed design meets the timing constraints as total Hold Slack (THS), total negative Slack (TNS) and total Pulse Width Negative Slack (TPWS) is equal to '0' ns as shown in Fig. 20. The total critical path delay

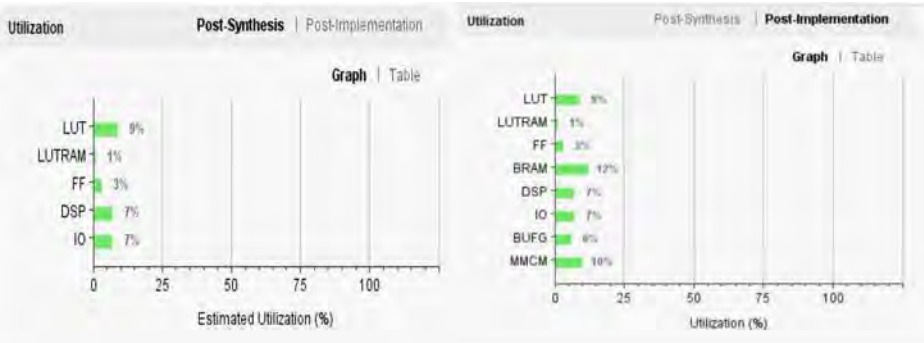


Fig. 19 Device utilization graph related to post-synthesis and implementation

Fig. 20 Design timing summary

from clock input port to SVM output is calculated with maximum delay 1.391 ns and minimum delay 0.577 ns after applying timing constraint to 2.00 ns by creating "timing.xdc" file.

5.3 Analysis of SVM output

The output of the SVM linear classifier is obtained with radix hexadecimal number which is tabulated as shown in Table 4. Here only first 20 samples have considered for experimental analysis. The obtained experimental result is compared with WBCD labeled dataset.

The obtained hexadecimal number need to be converted into real number for comparison to find out the classified output whether it belongs to 'Benign' or 'Malignant'. If the SVM output is positive or equal to '0' then decision is taken w.r.t 'Malignant'. Similarly, if the SVM output is negative or less than '0' then decision is taken w.r.t 'Benign'. This results in more inaccuracy due to comparing with only positive, negative and '0' for classification. In order to overcome from this problem confusion matrix, need to be incorporated in order to increase the percentage accuracy by considering true positive or negative and false positive or negative.

Table 4 Output of SVM linear classifier algorithm

SVM output in HEX	SVM Classified output	WBDS Label data	Classification accuracy	Simulation time in ns	SVM output in real number
3f3994f2	M	M	Yes	6800	0.724929
3e84fa7e	M	M	Yes	6900	0.2597236
3edfe601	M	M	Yes	7000	0.4373017
008efa66	M	M	Yes	7100	1.31E-38
42052a6f	M	M	Yes	7200	33.29144
801b7778	B	M	No	7300	−2.52E−39
4198b2ac	M	M	Yes	7400	19.08724
80255a78	B	M	No	7500	-3.43E-39
419fb3d8	M	M	Yes	7600	19.96281
001576d8	M	M	Yes	7700	1.97E-39
41a70a21	M	M	Yes	7800	20.87995
41 ac4967	M	M	Yes	7900	21.53584
41cdba85	M	M	Yes	8000	25.71607
41cf95ae	M	M	Yes	8100	25.94809
41d0c3e1	M	M	Yes	8200	26.09564
805047c0	B	M	No	8300	−7.37E-39
41acdbd6	M	M	Yes	8400	21.60734
41ac693a	M	M	Yes	8500	21.55138
41d00846	M	M	Yes	8600	26.00404
41c3135e	M	B	No	8700	24.38446

The SVM output data can be extracted from Vivado simulation tool by designing SISO architecture in Verilog and then the obtained data is in row vector may be mapped into column vector for classification. The comparison is done by writing program in MATLAB and to compute percentage accuracy. The correct classification rate of proposed system is 91.08%. The classification accuracy further can be increased by using exact dot product accumulator with additional hardware cost.

This paper describes a linear SVM classifier system based on FPGA for acceleration of data. The simulation and synthesis results shows that the SVM linear classification system designed for FPGA implementation works faster for binary classification system with high accuracy. The SVM linear classifier system designed for data acceleration meets the proposed requirement.

6 Conclusion

The proposed parallel implementation of the SVM algorithm for detection of breast cancer using Vivado simulation tool as a training method is presented. This implementation's main goal is to attain a high rate of data processing in order to meet the requirement of computationally intensive applications. As a result, all possible computations were parallelized. Finally, based on our interpretations using the synthesis results, we can conclude that implementing this technique in hardware would result in significant performance enhancements over the current state of the art. Based on the simulation and synthesis results presented, we successfully implemented the SVM Classifier for IEEE 754 number representation. The simulation and synthesis results exhibits that the SVM linear classification system is more effective in fast data classification. According to our experimental results, maximum frequency of 100 MHz in linear classification, and classification accuracy of 91.08% in linear one has been achieved. The total on-chip power such as static and dynamic power is 1.124 W. The total critical path delay from clock input port to SVM output is calculated with maximum delay 1.391 ns and minimum delay 0.577 ns after applying timing constraint to 2.00 ns. The LUT utilizing resources is about 49% within a band of 25% of total device utilization after post-implementation. The accuracy further can be increased by designing a dot product accumulator to store data after multiplication and then performing rounding at the end instead of performing rounding after each multiplication results in low accuracy. As per our knowledge this algorithm provides high accuracy rate compared to other classification algorithms which is implemented for breast cancer detection using FPGA. Furthermore, this study can be extended to various other algorithms. Along with implementation on FPGA to further improve the performance by fine-tuning the SVM parameters, area placement, latency and power consumption.

The FPGA implementation of the SVM linear classifier with exact dot product accumulator provides a way for improving the percentage of accuracy compared to state of the art and provides more flexibility to alter the structure in the future. Because of essential characteristics of the FPGA, the FPGA is able to reconfigure the hardware in a faster way. The exact dot product accumulator can be designed to reduce rounding error in order to improve the accuracy after performing multiplication between two floating point numbers such as patient data or support vectors and weight vector.

Abbreviations *SISO,* Serial in Serial Out.; *SVM,* Support Vector Machine; *THS,* Total Hold Slack.; *RAM,* Random Access Memory; *CNN,* Convolutional Neural Network; *WNS,* Worst negative slack.; *FPU,* Float Point Unit; *TPWS,* Total Pulse Width Negative Slack.; *LUT,* Look Up Table; *WPWS,* Worst pulse width slack.; *DSP,* Digital Signal Processing; *CPU,* Central Processing Unit; *SDSoC,* Software-Defined System-On-Chip.; *WHS,* Worst hold slack.; *FIFO,* First in First Out; *SoC,* System on Chip; *FF,* Flip Flop; *TNS,* Total negative Slack.; *BRAM,* Block RAM; *FPGA,* Field Programmable Gate Array; *DMA,* Direct Memory Access

Data availability Data sharing not applicable to this article as no datasets were generated or analyzed during the current study.

Declarations

Conflict of interest All authors certify that they have no affiliations with or involvement in any organization or entity with any financial interest or non-financial interest in the subject matter or materials discussed in this manuscript. The authors have no financial or proprietary interests in any material discussed in this article.

References

1. Adamowicz E, Walczyk K, Bernat T, Danenas P, Garsva G, Stępień P (2002) Support vector machines and their application in credit risk evaluation process
2. Afifi S, GholamHosseini H, Sinha R (2019) A system on chip for melanoma detection using FPGA-based SVM classifier. Microprocess Microsyst 65:57–68. https://doi.org/10.1016/j.micpro.2018.12.005
3. Afifi S, GholamHosseini H, Sinha R (2020) Fpga implementations of svm classifiers: a review. SN Comput Sci 1(3):1–17
4. Amezzane I, Fakhri Y, El Aroussi M, Bakhouya M (2020) Hardware acceleration of svm training for real-time embedded systems: Overview. Recent Adv Math Technol pp 131–139, Springer
5. Baccarini LMR, Silva VVRE, de Menezes BR, Caminhas WM (2011) SVM practical industrial application for mechanical faults diagnostic. Exp Syst Appl 38(6):6980–6984
6. Baez A, Fabelo H, Ortega S, Florimbi G, Torti E, Hernandez A, Leporati F, Danese G, Callico GM, Sarmiento R (2019) High-level synthesis of multiclass SVM using code refactoring to classify brain cancer from hyperspectral images. Electron 8(12):1494
7. Bassoli M, Bianchi V, De Munari I (2019) A Simulink Model-based Design of a Floating-point Pipelined Accumulator with HDL Coder Compatibility for FPGA Implementation. Appl. Electron. Pervading Ind. Environ. Soc. ApplePies 2019. Lect. Notes Electr. Eng. 1–9, in press
8. Bassoli M, Bianchi V, De Munari I (2020) A model-based design floating-point accumulator. case of study: Fpga implementation of a support vector machine kernel function. Sensors 20(5):1362. https://doi.org/10.3390/s20051362
9. Batista GC, Oliveira DL, Saotome O, Silva WLS (n.d.) A low-power asynchronous hardware implementation of a novel SVM classifier, with an application in a speech recognition system. Microelectron J
10. Blaiech, Ahmed Ghazi, Khaled Ben Khalifa, Carlos Valderrama, Marcelo AC Fernandes, Mohamed Hedi Bedoui (2019) A survey and taxonomy of FPGA-based deep learning accelerators. J Syst Archit 98: 331–345
11. Breast Cancer Wisconsin (Diagnostic) Data Set (n.d.) https://www.kaggle.com/uciml/breast-cancer-wisconsin-data. Accessed 5 Jan 2020
12. Carmichael Z, Langroudi HF, Khazanov C, Lillie J, Gustafson JL, Kudithipudi D (2019) Deep positron: a deep neural network using the posit number system. In: 2019 Design, Automation & Test in Europe Conference & Exhibition (DATE), pp 1421–146. https://doi.org/10.23919/DATE.2019.8715262
13. Carmichael Z, Langroudi HF, Khazanov C, Lillie J, Gustafson JL, Kudithipudi D (2019) Performance-Efficiency Trade-off of Low-Precision Numerical Formats in Deep Neural Networks. In: Proceedings of the Conference for Next Generation Arithmetic 2019 [Internet]. New York, NY, USA: Association for Computing Machinery; (CoNGA'19)
14. Chen W, Pourghasemi HR, Kornejady A, Zhang N (2017) Landslide spatial modeling: introducing new ensembles of ANN, MaxEnt, and SVM machine learning techniques. Geoderma 305:314–327. https://doi.org/10.1016/j.geoderma.2017.06.020
15. Cortes C, Vapnik V (1995) Support-vector networks. Mach Learn 20(3):273–297

16. Fiolhais L, Neto H (2018) An efficient exact fused dot product processor in FPGA. In: 2018 28th International Conference on Field Programmable Logic and Applications (FPL), pp 327–3273. https://doi.org/10.1109/FPL.2018.00062
17. Khan, F, Khan MA, Abbas S, Athar A, Siddiqui SY, Khan AH, Saeed MA, Hussain M (2020) Cloud-based breast cancer prediction empowered with soft computing approaches. J Healthcare Eng 2020. https://doi.org/10.1155/2020/8017496
18. Kim, S, Rutenbar RA (2019) An area-efficient iterative single-precision floating-point multiplier architecture for FPGA. In Proceedings of the 2019 on Great Lakes Symposium on VLSI, pp 87–92
19. Lopes FF, Ferreira JC, Fernandes MAC (2019) Parallel implementation on FPGA of support vector machines using stochastic gradient descent. Electronics 8(6):631. https://doi.org/10.3390/electronics8060631
20. Louca, L, Cook TA, Johnson WH (1996) Implementation of IEEE single precision floating point addition and multiplication on FPGAs. In FCCM, pp 107–116
21. Mavroforakis ME, Theodoridis S (2006) A geometric approach to support vector machine (SVM) classification. IEEE Trans Neural Netw 17(3):671–682. https://doi.org/10.1109/TNN.2006.873281
22. Meng W, Dou Y (2021) An optimization algorithm for computer-aided diagnosis of breast Cancer based on support vector machine. Front Bioeng Biotechnol 9:581
23. Mohammadi M, Krishna A, Nalesh S, Nandy SK (2018) A hardware architecture for radial basis function neural network classifier. IEEE Trans Parallel Distrib Syst 29(3):481–495
24. Nadikattu RR (2020) A comparative study between simulation of machine learning and extreme learning techniques on breast Cancer diagnosis. Available at SSRN 3615092
25. Papadonikolakis M, Bouganis C-S (2008) A scalable FPGA architecture for non-linear SVM training. In: 2008 International Conference on Field-Programmable Technology, pp 337–340. https://doi.org/10.1109/FPT.2008.4762412
26. Papadonikolakis M, Bouganis C-S (2010) A novel FPGA-based SVM classifier. In 2010 International Conference on Field-Programmable Technology, pp 283–286. IEEE
27. Patil PP, Kotrappa S (2020) A novel approach to detect microcalcification for accurate detection for diagnosis of breast Cancer. in internet of things, smart computing and technology: a roadmap ahead, pp 81–94. Springer, Cham
28. Pietron M, Wielgosz M, Zurek D, Jamro E, Wiatr K (2013) Comparison of GPU and FPGA implementation of SVM algorithm for fast image segmentation. In: Architecture of Computing Systems–ARCS 2013: 26th International Conference, Prague, Czech Republic, February 19-22, 2013. Proceedings 26, 292–302
29. Ragab DA, Attallah O, Sharkas M, Ren J, Marshall S (2021) A framework for breast cancer classification using multi-DCNNs. Comput Biol Med 131:104245 Elsevier
30. Selvathi D, Nayagam D (2016) FPGA implementation of on-chip ANN for breast cancer diagnosis. Intell Decis Technol 10:1–12. https://doi.org/10.3233/IDT-160261
31. Selvathi D, Nayagam RD (2016) FPGA implementation of on-chip ANN for breast cancer diagnosis. Intell Decis Technol 10(4):341–352
32. Siddiqui F, Amiri S, Minhas UI, Deng T, Woods R, Rafferty K, Crookes D (2019) Fpga-based processor acceleration for image processing applications. J Imaging 5(1):16
33. Subasi A (2020) Practical machine learning for data analysis using Python. Academic Press
34. Wang C, Gong L, Yu Q, Li X, Xie Y, Zhou X (2016) DLAU: a scalable deep learning accelerator unit on FPGA. IEEE Trans Comput Aided Des Integr Circ Syst 36(3):513–517
35. Yao Y, Yang L, Yu Y, Xu H, Lv W, Li Z, Chen X (2013) K-SVM: an effective SVM algorithm based on K-means clustering. J Comput 8(10):2632–2639
36. Yue S, Li P, Hao P (2003) SVM classification: its contents and challenges. Appl Math A J Chin Univ 18: 332–342

Journal of King Saud University – Computer and Information Sciences 35 (2023) 87–102

HOSTED BY

Contents lists available at ScienceDirect

Journal of King Saud University – Computer and Information Sciences

journal homepage: www.sciencedirect.com

Twofold dynamic attention guided deep network and noise-aware mechanism for image denoising

Zihao Chen [a], Alex Noel Joseph Raj [a], Vijayarajan Rajangam [b], Wei Li [a], Vijayalakshmi G.V. Mahesh [c], Zhemin Zhuang [a,*]

[a] Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Department of Electronic Engineering, College of Engineering, Shantou University, China
[b] Centre for Healthcare Advancement, Innovation and Research, Vellore Institute of Technology Chennai, India
[c] BMS Institute of Technology and Management, India

ARTICLE INFO

Article history:
Received 27 July 2022
Revised 6 January 2023
Accepted 4 February 2023
Available online 11 February 2023

Keywords:
Twofold CNN
Image denoising
Attention mechanism
Feature enhancement

ABSTRACT

Convolutional neural networks are given extensive attention towards noise removal due to their good performance over traditional denoising algorithms. With shallow conventional neural networks, the feature extraction ability is not profound. While employing deeper networks, network performance improves with the cost of additional computational requirements. In this paper, we propose an attention-guided twofold denoising network to remove the noise present in the image. The proposed network incorporates dilation convolution to enlarge the receptive fields and improves the feature extraction ability. Also, the presence of attention mechanism strengthens the extracted features and restores the image details during the noise removal. To demonstrate the superiority of the twofold structure, the proposed network is compared with the state-of-the-art denoising models. The experimental results prove that the proposed deep network achieves good peak signal to noise ratio and structural similarity index for different noise levels.

1. Introduction

Image denoising is a low-level pre-processing task in the field of computer vision which enables recovery of high quality images from the noisy source images (He et al., 2018). Different kinds of noise, such as photonic noise, shot noise, current noise and thermal noise, are introduced during image acquisition, storage and transmission. The noises are introduced due to the influence of imaging sensor material properties, illumination environment, electronic components in imaging, storage and communication systems. Typically, the noise degraded image y is modeled as $y = x + v$, where x is the noise free image and v is additive white Gaussian noise (AWGN). Image denoising can help in improving the accuracy of high-level computer vision tasks like image classification, segmentation (Arbelaez et al., 2010), and recognition (Romano et al., 2017) and is widely required in various fields, such as medical image analysis (Jifara et al., 2019) and target detection (Redmon et al., 2016). The prime objective of image denoising is to solve the problem of image quality degradation due to noise interference and to effectively improve the image quality and signal-to-noise ratio, so as to present better information available in the original image. Therefore, as an important pre-processing tool, image denoising algorithms have been extensively studied.

Traditional image denoising algorithms, such as BM3D (Dabov et al., 2007) and WNNM (Gu et al., 2014), have achieved good results in image denoising tasks, but the optimization complexity of the methods makes the denoising process time consuming. In BM3D, the algorithm uses block matching to aggregate the similar structures in the image, and later employs orthogonal transformation followed by filtering to perform 3D inverse transformation to get the denoised image. With the development of deep learning architectures and the widespread use of convolutional neural networks (CNNs) in recent years, image denoising techniques have also evolved rapidly. CNNs are often trained on large datasets with the noisy images as the inputs and deliver a corresponding denoised image as the desired output. During training, the loss function adjusts the network weights in order to reduce the difference between the outputs and the target values. Deep CNNs have

* Corresponding author.
E-mail address: zmzhuang@stu.edu.cn (Z. Zhuang).

Peer review under responsibility of King Saud University.

Production and hosting by Elsevier

https://doi.org/10.1016/j.jksuci.2023.02.003
1319-1578/

been widely used in image denoising and super-resolution due to its superior performance (Yamanaka et al., 2017). Zhang et al. proposed a deep CNN for image denoising (DnCNN) using batch normalization (BN) (Ioffe and Szegedy, 2015b) and residual learning (RL) (Wang et al., 2015). The experiments conducted on DnCNN demonstrate that the models with residual learning and normalization have faster and more consistent convergence with good denoising performance than the traditional models. Zhang et al. proposed a fast and flexible denoising network (FFDNet) (Zhang et al., 2018) by utilizing a single denoising network to handle multiple noise levels (Tai et al., 2017). Most deep learning denoising networks need to train multiple models for different noise levels, while FFDNet uses images with different noise levels to train the model, enabling a single model to remove multiple noise levels. However, FFDNet requires various noise levels as input data during processing and therefore ineffective for denoising images with unknown noise levels. Furthermore, the development of generative adversarial networks (GAN), has accelerated the image denoising research. For example, Denoise GAN proposed by Alsaiari et al. (2019) used a trained GAN based on a residual network to generate clean images through the generator. Here, the discriminator improves the quality of the image produced by the generator by comparing it with the original image. Deep learning denoising algorithms are based on learning methods, providing better performance due to end-to-end training (Zhang et al., 2021) with optimized time consumption. Whereas, traditional rule based algorithms require prior knowledge of the degradation, and thus offer less generalization with more time-consumption. Therefore, deep learning-based denoising methods have been a preferred choice by many researchers.

In this paper, we introduce an image denoising as an end-to-end learning task, where noisy images are presented as the inputs to the network and noise-free images are used as the learning targets. The incumbent noise is separated from the images using a attention-guided twofold convolutional neural network (ATDNet), where the traditional convolutinal layers are replaced by residual networks. The proposed network comprises of a sparse CNN structure and non-sparse CNN structure. The structure enables the network to extract additional noise information (more contextual information), mainly through the sparse network that incorporates dilation convolution layers that accurately predict the noise distribution which further allows the denoised image to retain more of its original details. As different noise features are combined, the feature information extracted from the network provides both the distribution and conceptual information of noise providing denoised images that are more closer to the original image. In addition, an attention mechanism is also introduced to the sparse layer that controls the weights that reduce the density of noise features and enhances the overall robustness of the model. Eventually, the denoised image closely represents the original image devoid of noise content.

The main contributions are summarized as follows:

- We propose a novel network for image denoising comprising of top and bottom networks arranged in parallel. The model includes dilated convolution and attention mechanisms to enable stronger noise extraction ability than conventional networks such as DnCNN (Zhang et al., 2018), ECNDNet (Tian et al., 2019) and SANet (Zhang et al., 2019).
- The proposed twofold network fuses extracted noise features with the features of the input image. The attention mechanism adjusts weights from the features extracted by the top network to enhance the noise extraction ability of the overall network. The ablation experiments shown in Section 4.4 illustrate that the proposed technique substantially improves the performance of image denoising algorithm.
- In addition to peak signal to noise ratio (PSNR) and structural similarity index (SSIM), as denoising performance metrics, we have also used a metric referred as Intra-SSIM that is more in line with human visual characteristics. The Intra-SSIM demonstrates the recovery details through contour reconstruction and is helpful in understanding the image quality when the PSNR and SSIM values between the original and denoised images are small.

This paper is organized as follows. Section 2 reviews the deep learning networks for denoising which is followed by residual learning, dilated convolution, and attention mechanism. The proposed architecture and its modules are presented in Section 3 followed by the loss function and evaluation standard. Experimental configuration and results, with the details of dataset and implementation, are presented in Section 4. This section also compares the performance of the proposed network with the state of the art denoising methods. Conclusion is given in Section 5.

2. Related work

2.1. Deep learning for image denoising

Deep learning is widely used in image denoising due to its superior performance. Jain and Seung (2008) pioneered the use of CNN to denoise the noisy image, but the denoising performance was inferior when compared to traditional methods due to its simple structure consisting of five-layers of neural network. Xie et al. (2012) proposed auto-encoders for image denoising that combines sparse coding and deep network training. The model can handle complex noise tasks as superimposed text type, however, the denoising effect is not significant compared with traditional algorithms such as K-singular value decomposition (K-SVD) (Elad and Aharon, 2006). By linking convolution and deconvolution, (Mao et al., 2016) created a very deep residual encoder-decoder network to eliminate noise in the image. This model uses convolutions to extract features from the input image which is followed by a deconvolution to obtain a denoised image. Further, the use of skip connections solves the gradient disappearance problem. With the idea of increasing the number of CNN layers to enhance the expression ability and to predict the noise, DNCNN (Zhang et al., 2017) is designed with a 17-layer deep network consisting of convolution (Conv), batch normalization (Ioffe and Szegedy, 2015b) and Rectified Linear Unit (ReLU) followed by global residual learning technique to present a denoised image. In order to balance the computational cost and accuracy, Zhang et al. proposed an image restoration CNN (IRCNN) (Kai et al., 2017) method and successfully predicted the denoised image by combining the optimization and discriminant learning methods. The denoising performance of the model depends on the model structure and the chosen optimization technique. When the primary structure successfully extracts the noise, the model's performance is good and the recovered image is comparable to the ground truth. However, if the denoising task becomes more complex such as removal of blind and real noise, the model will fail to recover the noise free image. Subsequently, the problems associated with (Kai et al., 2017), were solved by Pan et al. (2018) who proposed a dual convolutional neural network to extract features that complement each other and thus enhance the recovery ability of the model. In addition, Rawat et al. (2021) constructed a complex-valued convolutional neural network using complex-valued convolutional filters, wherein the complex-valued operations replaced the real-valued operations. The architecture was subjected to the task of medical image denoising, which exhibited an improvement of the PSNR to 37.201 dB compared to Dncnn's 35.099 db at a noise level of

15 db. Fan et al. fused Swin Transformer (Liu et al., 2021a) with UNet to construct SUNet (Fan et al., 2022) for image denoising. The Swin Transformer module extracted global information from the image and provided a PSNR improvement of 5% when compared to the conventional UNet architecture.

The feature extraction ability of the network can be improved by increasing the network's receptive field. However, increasing the depth alone to acquire a larger receptive field, will lead to a loss of contextual information and lower the networks denoising performance (He et al., 2016). Since, the denoising ability of the network is to accurately learn the noise distribution within the image. This analogy seeks to design a twofold convolutional network that incorporates dilated convolution and attention mechanisms to enhance the receptive field and improve the denoising ability of the model.

2.2. Residual learning and dilated convolution

Deeper neural network will lead to accuracy saturation and training difficulties. Therefore, deep residual learning for image recognition (ResNet) (He et al., 2016) was introduced. Residual learning models include an identity mapping layer to solve the gradient disappearance or explosion issues by connecting a shallow network to the deep network which allows fused features to serve as the inputs to the next layer. Due to better performance, the residual learning methods were widely used in image super-resolution and image denoising tasks(Geng et al., 2021). Zhang et al. proposed a residual dense network for image restoration (RDN) (Zhang et al., 2018) that uses short-connection to connect each layer of the network to the previous layer. It effectively solves the problem of gradient disappearance and realizes feature reuse.

Generally, more convolution layers in the model will improve the size of the receptive field at the cost of more parameters that eventually decreases the network's performance (Lu et al., 2018). Therefore, pooling operations were often used to reduce the size of the feature layers and thus enhancing the performance of the model. However, pooling layers cause loss of spatial data requiring interpolation procedures for feature reconstruction. To compensate, the dilated convolution technique was proposed by Yu and Koltun (2015) that allows the convolution functions to obtain information from larger receptive field with a lower computational parameters. Dilated convolution has been successfully used in problems such as image segmentation (Chen et al., 2016), speech synthesis (Oord et al., 2016), and machine translation (Kalchbrenner et al., 2016). Based on the fact that residual learning and extended convolution can improve the denoising performance of CNN, a twofold CNN for image denoising is proposed in this paper.

2.3. Attention mechanism

Visual interpretation can empower the user to understand the decision-making ability of a deep convolutional neural network, but it is inadequate to improve the performance of CNN (Fukui et al., 2019). Hence, an end-to-end attention mechanism was proposed by Zhu et al. (2018). In attention-oriented mechanism, the historical feature map is fused with the current feature map to enhance the ability of feature extraction and performance of the model. SISR-CA-OA (Chen et al., 2019) used a channel attention mechanism to fuse the features extracted from different directions. In terms of accuracy of the recovered noise free image and computational efficiency, the framework outperforms the single image super-resolution (SISR) algorithm, which is void of any attention mechanism. Inspired by the above, we also utilize the attention mechanism to enhance feature extraction process in the proposed network.

3. Proposed network

In this section, we introduce the architecture of ATDNet, the structure and parameter settings of each block and its role within the network. The proposed model consists of a top and a bottom network allowing features to be extracted in parallel (Zheng et al., 2018). The extracted features are combined to enhance the denoising ability of the model. The denoising procedure of ATDNet is illustrated in Fig. 1 as follows. (1) The noisy image Y is fed as the input to both the top and bottom network of ATDNet. The networks deliver two noisy features, NF_{top} and NF_{bottom}, respectively; (2) The feature NF_{top} is combined with Y to obtain $\overline{NF}_{top}$ and feature NF_{bottom} is subtracted from Y to obtain $\overline{NF}_{bottom}$. The need for addition and subtraction of the noise features with the input is presented in Section 3.1 (3) $\overline{NF}_{top}$ and $\overline{NF}_{bottom}$ are fused (NF_{fused}); then passed through an attention-guided mechanism and later multiplied with the NF_{top}, to render the final noise feature NF_{final} (4) The noise NF_{final} is removed from the input image Y by residual learning to obtain a clean image X.

3.1. Network architecture

The ATDNet consists of five main blocks: a Sparse Block (SB), an Enhancement Block (EB), a Global Feature Block (GFB), a Fusion Block (FB) and an Attention-guided Block (AGB). The overall block diagram is shown in Fig. 2 where the top and bottom networks consists of 16 and 17 layers respectively.

Sparse Block: The layers 1–12 of the top network are structured as SB and illustrated in Fig. 3. In layers 2, 5, 9, and 12, Dilated convolutions (Dila_Conv) are employed instead of normal convolutions and each of these layer includes Dila_Conv+

BN + ReLU operations. The other layers follow Conv + BN + ReLU, where Conv refers to conventional convolution operations. The convolution kernels are of size 3×3, the depth of the convolution layers is 64 and the dilation factor of Dila_Conv is 2. This configuration provides a receptive field with a size equivalent to twice the size of receptive field of traditional convolution kernels. We set the padding of the Dila_Conv layer to 2, and the other convolution layers as 1 to carefully extract the boundaries. The use of dilated convolutions within the SB abridges the need for the deep network, since information acquired from the receptive field of Dila_Conv is equivalent to the a deep networks. Also, the use of Dila_Conv reduces the complexity and the overall computational cost of the model. Furthermore, the introduction of BN within Conv and ReLU layers improves the training speed and the learning rate of the model (Ioffe and Szegedy, 2015a).

If the input gray image is of spatial resolution 256×256, then the overall computing process of top network is expressed as:

$$NF_{top}^{[l]} = \psi^{[l]}\left[\gamma^{[l]}\left(NF_{top}^{[l-1]} * K_{(fXf,c,D_k)}^{[l]}\right)\right] \tag{1}$$

where $NF_{top}^{[l-1]}$ represents the output of layer $l-1$ of the top network. $\psi^{[l]}$ is activation function of l-th layer, $\gamma^{[l]}$ represents batch normalization process and $K_{(fXf,c,D_k)}^{[l]}$ is the convolution kernel of l-th layer. where f and c are the size and depth of the convolution kernel respectively, D_k represents the dilation factor of the dilated convolution. Specially, $NF_{top}^{[0]}$ being the image in the input and $D_k = 1$ means use traditional convolution. Since the output of SB,

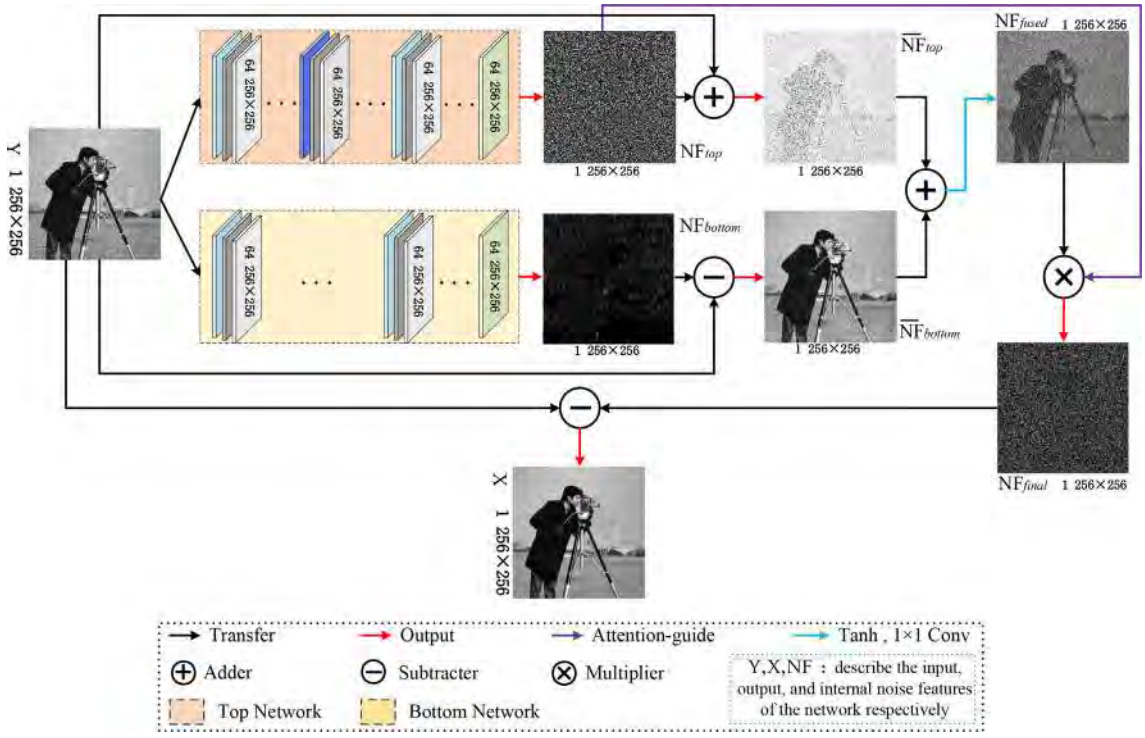


Fig. 1. The feature extraction steps of ATDNet.

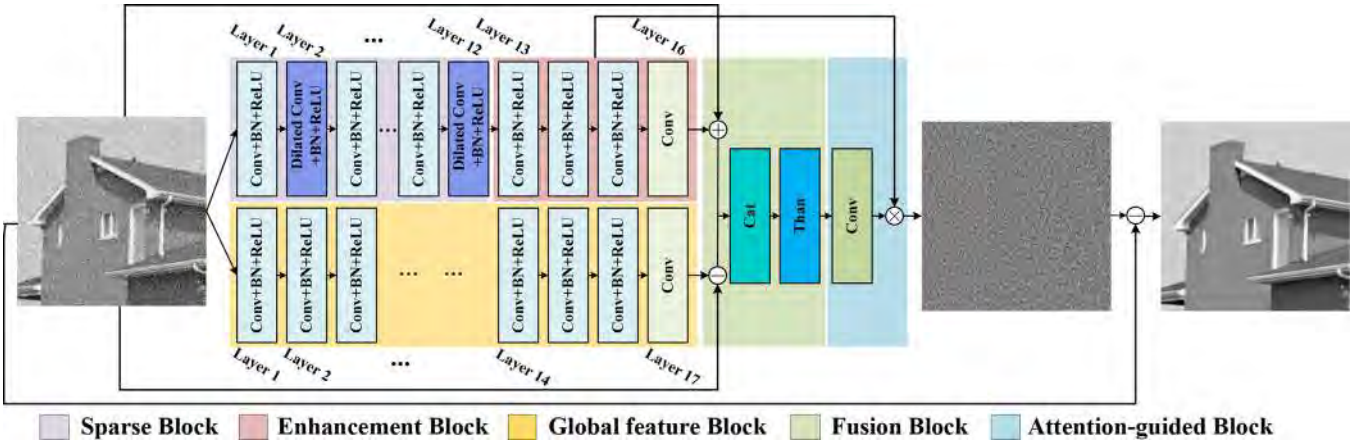


Fig. 2. The architecture of the proposed ATDNet. The top network consists of SB and EB, and the bottom network consists of GFB layers.

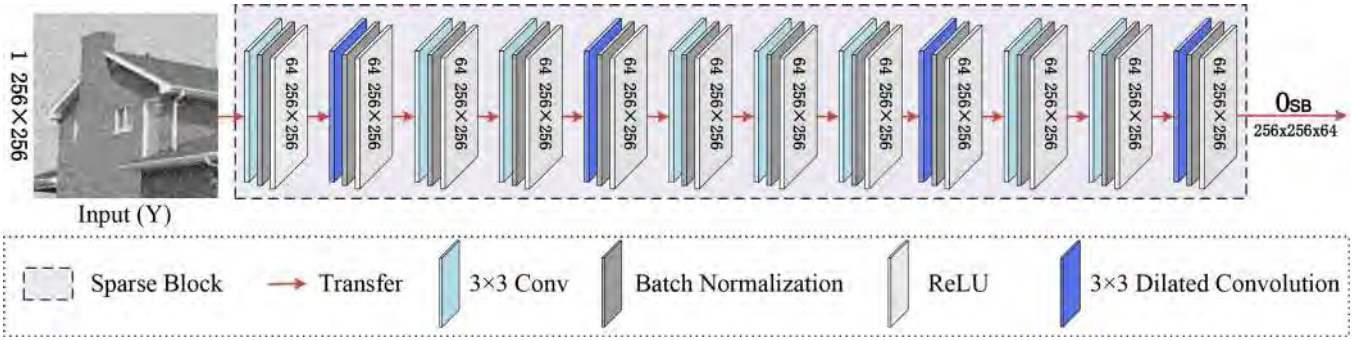


Fig. 3. The architecture of the Sparse Block; This block extracts noise features from the input image Y and feeds the result $O_{SB(256.256.64)}$ into EB.

$O_{SB(256,256,64)}$, represents the output of first 12 layers of the top network, the computing process is expressed as follows:

$$O_{SB(256,256,64)} = \psi^{[12]}\left[\gamma^{[12]}\left(NF_{top}^{[11]} * K_{(3X3,64,D_k=2)}^{[12]}\right)\right] \quad (2)$$

Enhancement Block: This block consists of four layers with first three layers having Conv + BN + ReLU, and the fourth layer with just one Conv operation. The EB achieves feature enhancement through dimensionality reduction. It performs compression on the features extracted by the SB block. Here, the convolution kernel size is 3 × 3, and the output of $O_{SB(256,256,64)}$ serves as its input. Specifically, O_{SB} has an output depth of 64 as it progress through the EB. The output is compressed to a depth of 1 as shown in Fig. 4. Further, the compression achieved through dimensionality reduction preserves the unique characteristic of the features obtained from the SB.

The output of the EB is expressed as $O_{EB(256,256,1)}$, i.e., $O_{EB(256,256,1)} = NF_{top}^{[16]}$. Specially, layer 16 does not use batch normalization and activation functions, so $\psi^{[16]}$=1 and $\gamma^{[16]}$=1 and $O_{EB(256,256,1)}$ is expressed as:

$$O_{EB(256,256,1)} = \psi^{[16]}\left[\gamma^{[16]}\left(NF_{top}^{[15]} * K_{(3X3,1,D_k=1)}^{[16]}\right)\right] \quad (3)$$

Global Feature Block: The GFB is the bottom network in the proposed model. It provides complementary features of the top network. As shown in Fig. 5, the GFB consists of 17 layers and layers 1–16 follow the structure Conv + BN + ReLU with kernel size of 3 × 3 and a depth of 64. The last layer includes just a single convolution. The GFB is void of Dila_Conv therefore enables the framework to extract dense noise features from the noisy input images as shown in Fig. 1. The structure of GFB along with the dimensions of each layer is shown in Fig. 5 and the process of bottom network is expressed as:

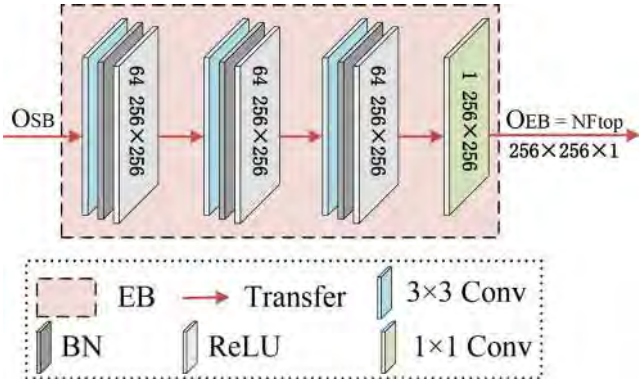


Fig. 4. The architecture of the EB; The EB compresses the output $O_{SB(256,256,64)}$ of the SB and feeds $O_{EB(256,256,1)}$ into the FB.

$$NF_{bottom}^{[l]} = \psi^{[l]}\left[\gamma^{[l]}\left(NF_{bottom}^{[l-1]} * K_{(fXf,c,D_k=1)}^{[l]}\right)\right] \quad (4)$$

where $NF_{bottom}^{[l-1]}$ represents the output of layer $l-1$ of the bottom network and $NF_{bottom}^{[0]}$ being the image in the input. Since dilated convolution is not used, D_k = 1. The output of GFB, $O_{GFB(256,256,1)}$, is expressed as

$$O_{GFB(256,256,1)} = \psi^{[17]}\left[\gamma^{[17]}\left(NF_{bottom}^{[16]} * K_{(3X3,1,D_k=1)}^{[17]}\right)\right] \quad (5)$$

Fusion Block: Generally, deep network can provide greater receptive field enabling extraction of unique features, but increasing depth of the network may not achieve good denoising effect. The available literature suggests that more reliable features can be obtained through the fusion of attributes extracted from various models (Kim et al., 2018; Zhuang et al., 2021). Therefore we employ a FB, shown in Fig. 6 to combine the features extracted from the top and bottom networks. The following operations are performed in FB: (a) addition of features extracted from the top network to the input image, (b) subtraction of features obtained from the bottom network from the input and (c) the fused results from (a) and (b) are then propagated to the activation layer. The fused output is represented by an equation:

$$\begin{aligned} NF_{fused} &= \psi\left[\left(NF_{top}^{[0]} + NF_{top}^{[16]}\right) + \left(NF_{bottom}^{[0]} - NF_{bottom}^{[17]}\right)\right] \\ &= \psi\left(\overline{NF}_{top} + \overline{NF}_{bottom}\right) \end{aligned} \quad (6)$$

Here, the output of FB is be expressed as $O_{FB(256,256,3)} = NF_{fused}$. Where, ψ denotes the activation function Tanh.

As shown in Fig. 1, the noise from both the networks is fused with the input image to enhance the expression ability of the extracted features, which means that a more accurate noise distribution is learnt. Specifically, the top network that employs a sparse mechanism and extracts the key information from the noisy image which is then added to the noisy image to enhance the weight of the key information. The GFB extracts the global information of the noisy image, and noise reduction is carried out by subtracting O_{GFB} from Y. Finally the output is added to the features obtained from the top network to further enhance the weight of the key information. Tanh is used as the activation function within the FB since it reduces the problems due to gradient disappearance and provides derivatives within the range of 0 to 1. The Tanh function maintains nonlinear monotonic ascending and descending relationship, which conforms to the gradient descent solution of back propagation network (Kalman and Kwasny, 1992). The final fusion result includes $O_{EB(256,256,1)}$, $O_{GFB(256,256,1)}$ and $Y_{(256,256,1)}$, i.e. a combination of outputs of the EB, GFB and the noisy input image.

Attention-guided Block: The AGB can locate key information among the local features that help in the decision-making process (Liu et al., 2021b; Borji and Itti, 2013). The challenge of image denoising comes from two aspects: noise extraction and

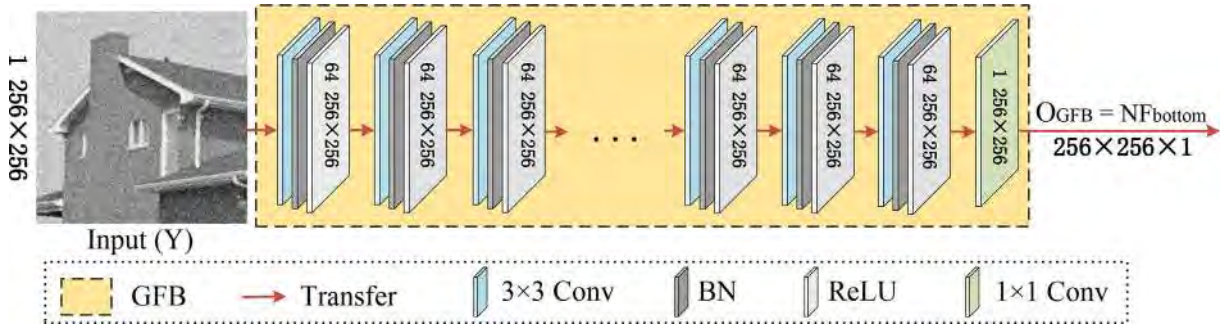


Fig. 5. The architecture of the Global Feature Block.

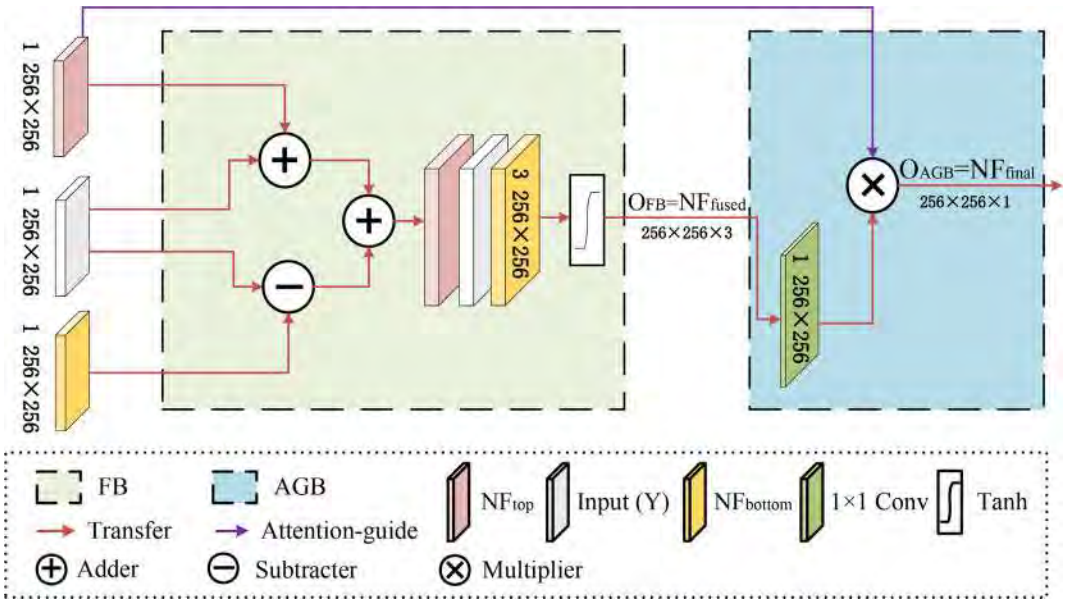


Fig. 6. The architecture of the Fusion Block and Attention-guided Block; FB fuses $O_{EB(256,256,1)}$,$Y_{(256,256,1)}$ and $O_{GFB(256,256,1)}$, and feeds the result $O_{FB(256,256,3)}$ into AGB.

restoration of key details. As mentioned in Section 2.2, we reasoned the choice of dilated convolutions to aggregate the context information with lesser computational cost using shallow network instead of a deep network. Here, we introduce the attention-guiding block to further strengthen the quality of the extracted features from the FB.

From Fig. 6, the designed AGB is composed of a 1×1 convolution kernel and a multiplier. Since $O_{EB(256,256,3)}$ represents the key features, the $O_{EB(256,256,1)}$ will be directly used as the weight of AGB multiplied by $O_{FB(256,256,3)}$. Finally, the noisy map, extracted from the network, indicates the common noise in the noisy image. The output of the AGB will be used as the final feature extracted by the network, i.e., $O_{AGB(256,256,1)} = NF_{final}$. This process can be expressed as

$$NF_{final} = NF_{top}^{[16]}\left(NF_{fused} * K_{(1X1,1,D_k=1)}\right) \tag{7}$$

where $K_{(1X1,1,D_k=1)}$ denotes a convolution kernel of size 1x1. $O_{AGB(256,256,1)}$ and also represents the noise feature extracted from the network. The attention mechanism enables the model to concentrate more on the noise regions while suppressing the irrelevant features i.e. preventing the extraction of features from noise-free regions, so as to obtain a smooth output. As shown in Fig. 1, the top network extracts important information from the noisy image, which plays a key role in detail recovery. Finally, the clean image corresponding to the noisy input image is constructed by residual learning and the operations are illustrated in the equation:

$$X_{(256,256,1)} = Y - NF_{final} \tag{8}$$

where $X_{(256,256,1)}$ represents the final denoised image.

3.2. Loss function

The objective of image denoising is to extract v from the model $y = x + v$, so as to restore clean image x. Therefore, we use the mean square error (MSE) (Ephraim and Malah, 1984) to train ATD-Net. MSE is generally used to find the deviation between the predicted value and the actual value of the model. The training data is presented as, $Train = \{(x_1, y_1), (x_2, y_2), \ldots, (x_n, y_n), \ldots, (x_N, y_N)\}$, where N is the total number of training samples, $n = 1, 2, \ldots, N$. The error generated by the model for the nth time is $\hat{v} = f_{ATDNet}(Y_n)$, and the true error is $v = y_n - x_n$. Therefore, the loss function is expressed as:

$$\begin{aligned} Loss &= \tfrac{1}{N}\sum_{n=1}^{N}(v_n - \hat{v}_n)^2 \\ &= \tfrac{1}{N}\sum_{n=1}^{N}((y_n - x_n) - f_{ATDNet}(x_n))^2 \end{aligned} \tag{9}$$

3.3. Performance evaluation metrics

The performance of image denoising is assessed by PSNR in dB and SSIM. However, both PSNR and SSIM have the following drawback: the values do always realise the correlation with human judgment (Lv et al., 2018), however, for humans, the reconstructed images with higher PSNR and SSIM values do not necessarily look better (Duan et al., 2011). Normally, the characteristic features of the images are distributed unevenly in space and the people with normal visual range can only focus on a small region of the image. The need for a local processing is more inline with the characteristics of human visual system (Hoffman, 1998). Based on the above mentioned points, we employ Intra-SSIM (Li et al., 2021; Wei Li et al., 2022) to evaluate local characteristics of the image. Firstly, the image is divided into subblocks and the SSIM is computed for each subblock. Next, a 3×3 window is placed over the subblocks, and the centre subblock is presented with average SSIM value, referred to as Intra-SSIM. It is computed from nine subblocks consisting centre block along with eight surrounding neighbours. The 3×3 window is slided over the other subblocks and the Intra-SSIM values for the entire image is computed. The calculation process is shown in Fig. 7.

Let $b = \{b_{i,j}|(i,j) = (1,1),(1,2),\ldots,(1,32),\ldots,(32,32)\}$ be the subblock and $B = \{B(p,q)|(p,q) = (1,1),(1,2),\ldots,(1,30),\ldots,(30,30)\}$ be the 3×3 window. Each window contains 9 subblocks. As shown in Fig. 7, where $B(1,1)$ represents the first 3×3 window. We divide an image of 256×256 pixels into 32×32 subblocks, where 8×8 represents the spatial resolution of the subblock. $b(2,2)$ is considered with each of the surrounding 8 subblocks to calculate the SSIM, and then averaged to get the Intra-ssim of the

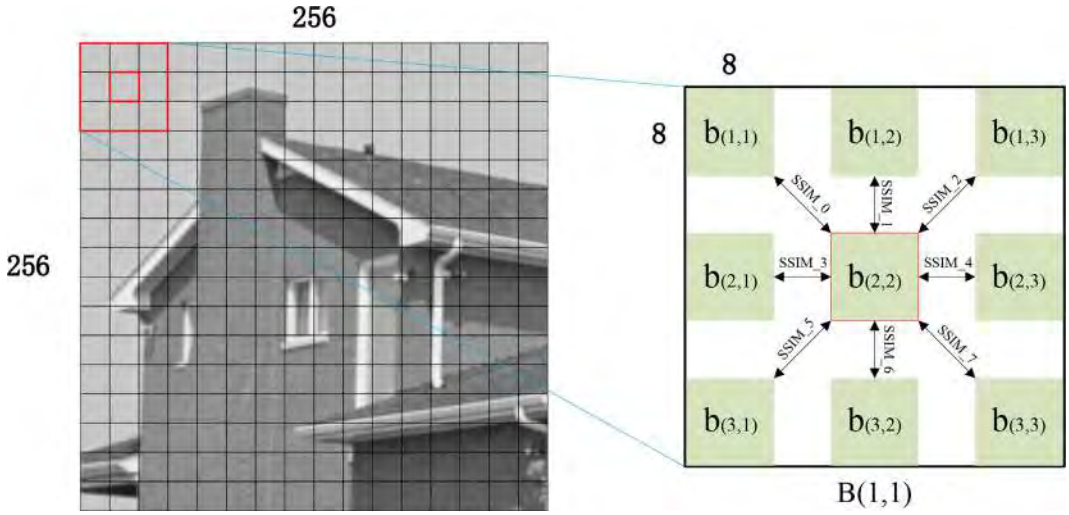


Fig. 7. Left shows the image with size of 256×256 pixels and divided into 32×32 subblocks. At each step of the sliding window, the mean SSIM is computed for the central block by considering the eight neighbouring subblocks as shown on right.

first window $B(1,1)$. This process results in a map known as the Intra-SSIM image map.

Proceeding further, a 30×30 Intra-SSIM image map is obtained. The contour representation of the Intra-SSIM values is useful in understanding the pixel intensity reconstruction in the denoised images, since the contour colors will enable to determine whether the denoised image is closer to the original image. Therefore the Intra-SSIM contour maps are more consistent with human visual characteristics for comparison.

4. Experiments, results and discussion

This section elaborates the dataset used for experiments, objective performance analysis using PSNR, SSIM and Intra-SSIM followed by comparative analysis with other denoising methods.

4.1. Dataset

To train our model, we used 2,000 images from the Waterloo Exploration Database and each image has a different spatial resolution. We employ the following three strategies to augment the training data in order to achieve sufficient training samples: (1) The input image is cropped by applying a sliding window of size 50×50 pixels with a overlap of 10, 20, 30 and 40 pixels (2) The cropped 4 different subimages have the spatial resolution of 50×50. Each of the sub-image is subjected to a random flip along horizontal and vertical direction. This is followed by the rotation operation of 90, 180 and 270 degrees. For comparison, we use Set12, BSD68 (Roth and Black, 2005) and Kodak24 (Franzen, 1999) as test sets. The noisy images are obtained by adding Gaussian noise to the inputs with noise levels(σ) of 15, 25 and 50 db. Since the amplitude of the Gaussian white noise obeys the Gaussian distribution, its power spectral density is uniformly distributed and, therefore, larger variance σ, depicts more noise. Fig. 8, illustrates the three levels of noise that generally represent the impact of noise on the image from low to high.

4.2. Implementation details

For training, we use Adam optimizer (Kingma and Ba, 2014) and train the model for 50 epochs. To achieve faster convergence, the learning rate was set as 0.01 for the first 20 epochs and later decreased to 0.001 and 0.0001 for epochs 21 to 40 and 41 to 50 respectively.

We have implemented the ATDNet model on PyTorch version 1.8.1 with Python3.7 executing at the background. The experiments were conducted on Work Station equipped with a CPU of Intel Xeon Gold 5118 and RAM of 128G. Also, the batch size was set to 128 and NVIDIA Quadro P5000 GPU was used to train the model. The trained model along with results are available at https://github.com/jalexnoel/ATDN.

4.3. Feature analysis

The main objective of the proposed ATDNet model is to reduce the network depth with an increase in the receptive field so as to

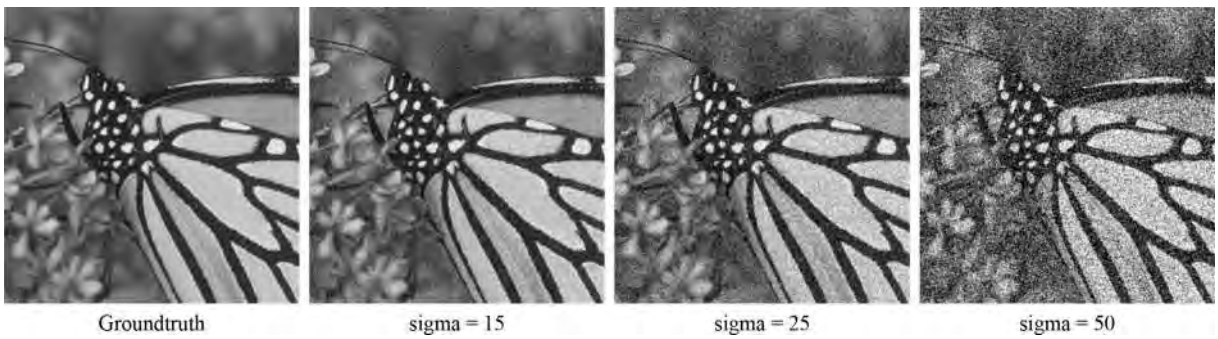


Fig. 8. Original Image (leftmost) and the original image with three different noise levels 15, 25 and 50 db.

improve the denoising effect at a low computational cost. In general, a *fxf* kernel engages a receptive field of *lxl* pixel while performing convolution operation. The receptive field is expressed in Eq. 10.

$$l_k = l_{k-1} + D_k * \left[(f_k - 1) * \prod_{i=1}^{k-1} S_i \right] \tag{10}$$

where l_{k-1} is the receptive field of layer $k-1$, f_k is the filter size of layer k and S_i is the sliding step of convolution kernel at layer i. In the network we designed, f_k=3 and S_i=1. When $k = 0, l_k = 1$. D_k represents the dilation factor of the k-th layer dilated convolution. $D_k = 1$ represents the normal convolution process When dilated convolution is not used.

Specifically, the top network has 16 layers employing 3 × 3 kernels with $D_k = 2$. Substituting the relevant values in Eq. (10) provides a receptive field of 41 × 41 pixels at the final layer. Similarly, the bottom network of 17 layers with 3×3 kernels has a receptive field of 35 × 35 pixels (substitute $D_k = 1$ in Eq. 10). Together, the average receptive field is (41 + 35)/ 2 = 38 × 38 pixels. On contrast, a single network with 19 layers using 3 × 3 kernel of traditional convolution provides a receptive field of only 38 × 38 pixels at the final layer. Thus, the proposed two fold ATDNet model uses less number of layers and contributes to a receptive field comparable to a single layer networks having more layers. Further, the dilation convolution used in the top framework enables to build a spare network that promotes randomly distributed features that improves diversity and the robustness to the ATDNet architecture.

Also, the fused features contains both global and local features, that accurately estimate the noise distribution without smoothing much of the image details. This aspect of the network is especially important in medical image denoising, where accurate recovery of details from the pathological areas is a important constraint. As the process shown in Algorithm 1, the network performs multiple feature fusion and enhancement in extracting the noise features in the image to avoid losing more details in the image during noise reduction.

Algorithm 1. Algorithm flow of ATDNet

Algorithm 1: Algorithm flow of ATDNet

Input: Noise image X

Output: Denoised image Y

The images are input to two networks separately for feature extraction and then feature fusion. The layers 2, 5, 9, and 12 of the top layer network contain the dilated convolution;

n and m represent the number of layers of the bottom and top networks, respectively;

$X_{bottom} = X$ ▷ $X \Rightarrow$ Bottom Network;

for $i \leftarrow 1$ **to** n **do**

  if $i != n$ **then**

    $NF_{bottom} = ReLU(BN(Conv(X_{bottom})))$;

    $X_{bottom} = NF_{bottom}$;

  else

    $NF_{bottom} = Conv(NF_{bottom})$;

  end

end

$X_{top} = X$ ▷ $X \Rightarrow$ Top Network;

for $j \leftarrow 1$ **to** m **do**

  if j *in* $[2, 5, 9, 12]$ **then**

    $NF_{top} = ReLU(BN(Dila\_Conv(NF_{top})))$;

  else

    if $j = m$ **then**

      $NF_{top} = Conv(NF_{top})$;

      break;

    end

    $NF_{top} = ReLU(BN(Conv(X_{top})))$;

  end

  $X_{top} = NF_{top}$;

end

$NF_{fused} = X - NF_{bottom} + X + NF_{top}$;

$NF_{final} = NF_{top}$ x $Tanh(Conv(NF_{fused}))$;

$Y = X$ - NF_{final} ;

Return Y

Initially to accelerate training, we use ReLU as the activation function at each node of the two-layer network that improves the convergence of the loss function and prevents the gradient from vanishing. However, the sparse processing forced by ReLU reduces the performance of the model (Ang-bo and Wei-wei, 2018), i.e., too much feature masking, results in the model not being able to learn the features effectively. In Fig. 9, we compare ATDNet with ReLU, Sigmoid and Tanh activation functions. The results demonstrate that using Tanh as the activation function in the FB allows the model to impart better denoising performance with a comparable converging speed during training.

The ability of the network to learn the noise distribution accurately from the images will determine the denoising performance of the model. The explicit goal of the proposed model is to employ deep learning to dissociate noise from the input data. However, raw data is usually intertwined within highly dense features which requires to be untangled and converted to sparse features that are stripped of extraneous noise (Tong et al., 2016). Therefore, the aggregate of sparse features of the top network, the features obtained from the bottom network and the inclusion of the attention-guided approach help in realizing and preserving key attributes to improve the denoising performance of the ATDnet model. In the following sections, we illustrate the denoising performance through (a) ablation study explaining the importance of the fused features of top and bottom networks, (b) denoising ability of the model for various noise levels and (c) further comparison of the ATDnet with other state-of-the art denoising methods.

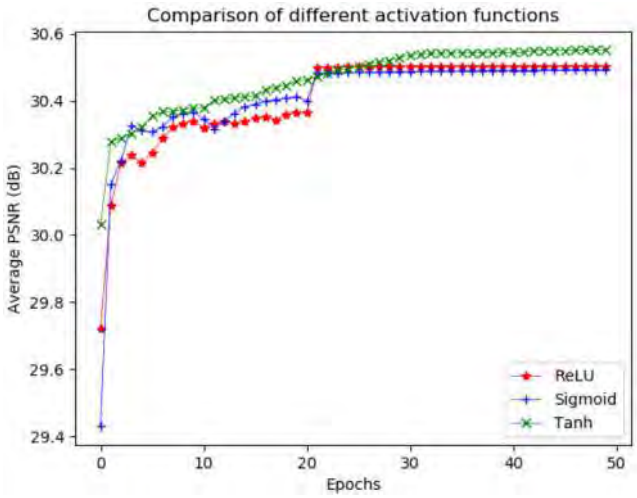


Fig. 9. Gaussian denoising results of three specific models. In the Fusion Block, ReLU, Sigmoid and Tanh were used as activation functions respectively. Results from 12 images of the Set12 dataset are presented.

4.4. Structure analysis

As mentioned above, ATDNet employs fused noise features obtained from both networks to enhance denoising performance of the model. To verify the effectiveness of this technique, we compare the denoising effect of following frameworks, i.e., (a) top network, (b) bottom network and (c) ATDNet: fused top and bottom networks.

Table 1 shows the denoising results of different networks on the same dataset, where the bold values indicates the highest value of PSNR and SSIM. From the results, it can be inferred that the fused networks perform better than the individual networks in terms of denoising performance in most cases. Though the maximum improvement of PSNR and SSIM when compared to single network is 0.05 and 0.01 respectively. The use of both features extracted from top and bottom network, as shown in Fig. 1, predicts the noise distribution which is closer to that of the input images.

Finally the denoised image is obtained by removing the noise from the input image using residual learning. Fig. 10 shows the difference between the denoising results of different networks and the original image. Experiments demonstrate that the fused network reduces MSE and improves the values of PSNR and SSIM. In the subsequent section, the denoising performance of ATDNet of various noise levels is presented.

4.5. Comparisons with the state-of-the-art

In this section, the proposed network is compared with ten state-of-the-art denoising techniques: BM3D (Dabov et al., 2007), WNNM (Gu et al., 2014), MLP (Burger et al., 2012), EPLL (Zoran and Weiss, 2011), TNRD (Chen and Pock, 2016), DnCNN (Zhang et al., 2018), IRCNN (Zhang et al., 2017), ECNDNet (Tian et al., 2019), SANet (Zhang et al., 2019) and PSN-U (Aljadaany et al., 2019). In most cases the performance metrics were either obtained directly from the paper or their publicly accessible codes were used to obtain the data.

Table 2 illustrates the PSNR values for Set12 image denoising dataset. Set12 is the grayscale image dataset, consisting of eight 256×256 images and four 512×512 images. For the noise level (σ) of 15 dB, the ATDNet outperforms the traditional methods with

Table 1
Denoising results of Set12 dataset with noise level 20 dB using for different methods.

Image	Top network		Bottom network		ATDnet	
	PSNR	SSIM	PSNR	SSIM	PSNR	SSIM
C.man	31.05	0.887	31.10	0.887	**31.14**	**0.889**
House	33.64	0.862	33.59	0.860	**33.68**	**0.863**
Peppers	31.80	0.894	31.81	0.889	**31.90**	**0.895**
Starfish	30.28	0.843	30.34	**0.853**	**30.35**	0.850
Monarch	31.09	**0.919**	31.11	0.907	**31.12**	0.902
Airplane	30.18	**0.884**	30.19	0.883	**30.20**	**0.884**
Parrot	**30.36**	0.877	30.31	0.876	**30.36**	**0.878**
Lena	33.01	**0.865**	32.98	0.859	**33.04**	**0.865**
Barbara	30.28	0.858	30.48	**0.872**	**30.55**	0.863
Boat	30.89	0.825	30.89	**0.826**	**30.94**	0.823
Man	30.85	0.801	30.86	0.800	**30.92**	**0.816**
Couple	30.83	0.844	30.84	**0.847**	**30.87**	0.844
Average	31.09	0.863	31.21	0.863	31.26	0.864

• The best results are marked in bold.

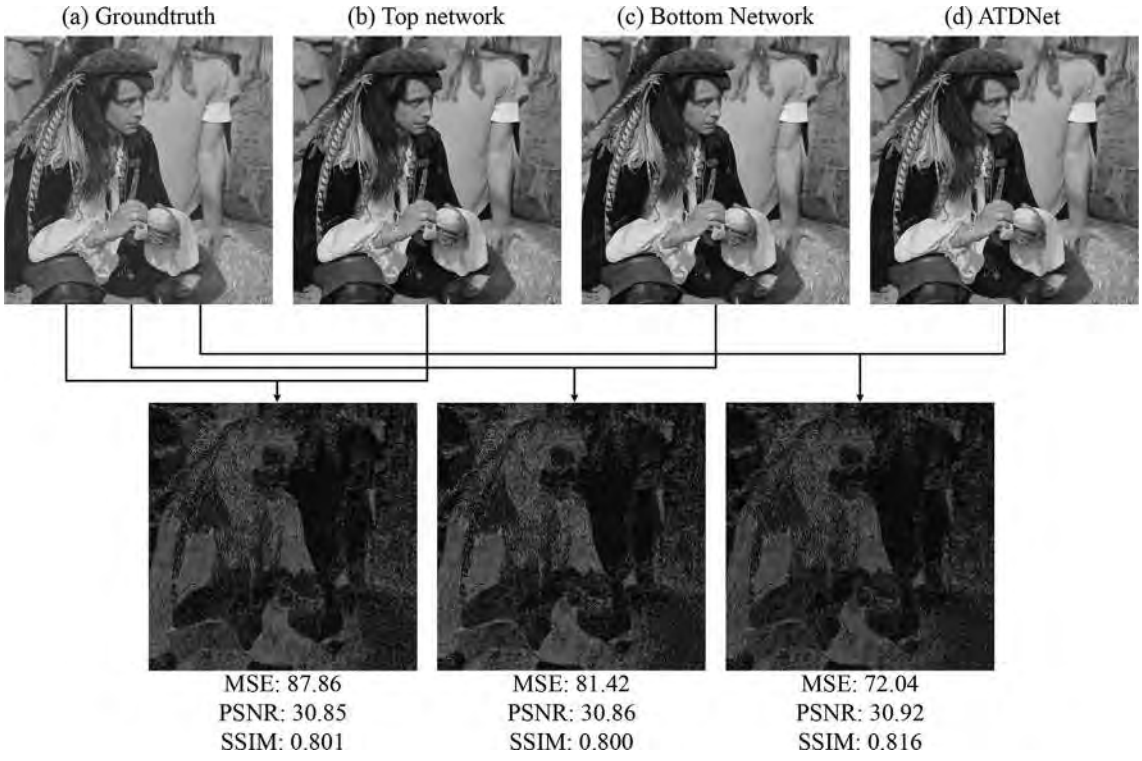


Fig. 10. Differences between the results of other methods and the relevant groundtruth.

Table 2
Average PSNR (dB) for different methods on Set12.

	Method	C.man	House	Peppers	Starfish	Monarch	Airplane	Parrot	Lena	Barbara	Boat	Man	Couple	Average
σ = 15	BM3D (Dabov et al., 2007)	31.91	34.93	32.69	31.14	31.85	31.07	31.37	34.26	33.10	32.13	31.92	32.10	32.37
	WNNM (Gu et al., 2014)	32.17	35.13	32.99	31.82	32.71	31.39	31.62	34.27	**33.60**	32.27	32.11	32.17	32.70
	EPLL (Zoran and Weiss, 2011)	31.85	34.17	32.64	31.13	32.10	31.19	31.42	33.92	31.38	31.93	32.00	31.93	32.14
	CSF (Schmidt and Roth, 2014)	31.95	34.39	32.85	31.55	32.33	31.33	31.37	34.06	31.92	32.01	32.08	31.98	32.32
	TNRD (Chen and Pock, 2016)	32.19	34.53	33.04	31.75	32.56	31.46	31.63	34.24	32.13	32.14	32.23	32.11	32.50
	DnCNN (Zhang et al., 2018)	32.61	34.97	33.30	**32.20**	33.09	31.70	31.83	34.62	32.64	32.42	**32.46**	32.47	32.86
	IRCNN (Zhang et al., 2017)	32.55	34.89	33.31	32.02	32.82	31.70	31.84	34.53	32.43	32.34	32.40	32.40	32.77
	FFDNet (Zhang et al., 2018)	32.43	35.07	33.25	31.99	32.66	31.57	31.81	34.62	32.54	32.38	32.41	32.46	32.77
	ECNDNet (Tian et al., 2019)	32.56	34.97	33.25	32.17	33.11	31.70	31.82	32.52	32.41	32.37	32.39	32.39	32.81
	SANet (Zhang et al., 2019)	32.38	35.03	33.18	32.14	**33.20**	31.71	31.89	34.54	32.61	32.36	32.38	32.41	32.82
	PSN-U (Aljadaany et al., 2019)	32.04	35.03	33.21	31.94	32.93	31.61	31.62	34.56	32.49	32.41	32.37	32.43	32.72
	ATDNet	**32.79**	**35.16**	**33.47**	32.14	33.17	**31.83**	**31.92**	**34.69**	32.74	**32.56**	**32.47**	**32.56**	**32.95**
σ = 25	BM3D (Dabov et al., 2007)	29.45	32.85	30.16	28.56	29.25	28.42	28.93	32.07	30.71	29.90	29.61	29.71	29.97
	WNNM (Gu et al., 2014)	29.64	33.22	30.42	29.03	29.84	28.69	29.15	32.24	**31.24**	30.03	29.76	29.82	30.26
	EPLL (Zoran and Weiss, 2011)	29.26	32.17	30.17	28.51	29.39	28.61	28.95	31.73	28.61	29.74	29.66	29.53	29.69
	CSF (Schmidt and Roth, 2014)	29.48	32.39	30.32	28.80	29.62	28.72	28.90	31.79	29.03	29.76	29.71	29.53	29.84
	MLP (Burger et al., 2012)	29.61	32.56	30.30	28.82	29.61	28.82	29.25	32.25	29.54	29.97	29.88	29.73	30.03
	CSF (Schmidt and Roth, 2014)	29.48	32.39	30.32	28.80	29.62	28.72	28.90	31.79	29.03	29.76	29.71	29.53	29.84
	TNRD (Chen and Pock, 2016)	29.72	32.53	30.57	29.02	29.85	28.88	29.18	32.00	29.41	29.91	29.87	29.71	30.06
	DnCNN (Zhang et al., 2018)	30.18	33.06	30.87	29.41	30.28	29.13	29.43	32.44	30.00	30.21	30.10	30.12	30.43
	IRCNN (Zhang et al., 2017)	30.08	33.06	30.88	29.27	30.09	29.12	29.47	32.43	29.92	30.17	30.04	30.08	30.38
	FFDNet (Zhang et al., 2018)	30.10	33.28	30.93	29.32	30.08	29.04	29.44	32.57	30.01	30.25	**30.11**	30.20	30.44
	ECNDNet (Tian et al., 2019)	30.11	33.08	30.85	**29.43**	30.30	29.07	29.38	32.38	29.84	30.14	30.03	30.03	30.39
	SANet (Zhang et al., 2019)	30.04	33.05	30.83	29.31	30.27	29.08	29.34	32.35	30.00	30.12	30.00	30.05	30.37
	PSN-U (Aljadaany et al., 2019)	29.79	33.23	30.90	29.30	30.17	29.06	29.25	32.45	29.94	30.25	30.05	30.12	30.37
	ATDNet	**30.34**	**33.39**	**31.11**	29.39	**30.33**	**29.15**	**29.49**	**32.59**	30.16	**30.34**	30.08	**30.23**	**30.55**
	BM3D (Dabov et al., 2007)	26.13	29.69	26.68	25.04	25.82	25.10	25.90	29.05	27.22	26.78	26.81	26.46	26.72
	WNNM (Gu et al., 2014)	26.45	30.33	26.95	25.44	26.32	25.42	26.14	29.25	**27.79**	26.97	26.94	26.64	27.05
	EPLL (Zoran and Weiss, 2011)	26.10	29.12	26.80	25.12	25.94	25.31	25.95	28.68	24.83	26.74	26.79	26.30	26.47
	MLP (Burger et al., 2012)	26.37	29.64	26.68	25.43	26.26	25.56	26.12	29.32	25.24	27.03	27.06	26.67	26.78
	TNRD (Chen and Pock, 2016)	26.62	29.48	27.10	25.42	26.31	25.59	26.16	28.93	25.70	26.94	26.98	26.50	26.81
	DnCNN (Zhang et al., 2018)	27.03	30.00	27.32	25.70	26.78	25.87	26.48	29.39	26.22	27.20	**27.24**	26.90	27.18
	IRCNN (Zhang et al., 2017)	26.88	29.96	27.33	25.57	26.61	25.89	26.55	29.40	26.24	27.17	27.17	26.88	27.14
	ECNDNet (Tian et al., 2019)	27.07	30.12	27.30	**25.72**	26.82	25.79	26.32	29.29	26.26	27.16	27.11	26.84	27.15
	SANet (Zhang et al., 2019)	26.92	29.93	27.27	25.52	26.64	25.71	26.18	29.22	26.37	27.20	27.11	26.80	27.09
	PSN-U (Aljadaany et al., 2019)	27.21	30.21	27.53	25.63	**26.93**	**25.89**	**26.62**	29.54	26.56	27.27	27.23	**27.04**	27.30
	ATDNet	**27.24**	**30.46**	**27.68**	25.65	26.85	25.85	26.52	**29.56**	26.40	**27.33**	27.15	**27.04**	**27.31**

• σ= 15, 25, 50 represent noise level in dB on Set12.
• The best results are marked in bold.

a highest improvement of 32.95 dB when compared to EPLL (Zoran and Weiss, 2011). When the noise level is increased to 25 dB and 50 dB, again ATDNet presents the average PSNR of 30.55 dB and 27.31 dB respectively which is better than other denoising methods. In Table 2, for each noise level, the best results are shown in blod. Table 3 shows the SSIM scores for ATDNet and other methods tested on Set12 dataset. Here, for three different noise levels, ATDNet has presented an improvement over BM3D by 0.05, 0.03 and 0.06 respectively.

For datasets such as BSD68 (Roth and Black, 2005) and Kodak24 (Franzen, 1999), most researchers have provided an average performance score for the complete dataset. In Table 4 and Table 5, the average PSNR and SSIM values of the dataset along with ten random images are shown. BSD68 has 68 grayscale images of different sizes while Kodak24 is composed of 24 grayscale images of 500 $\times$ 500 size.

In Table 6, we show the overview of the performance metrics, average PSNR and SSIM, of ATDNet on Set12, BSD68 and Kodak24, along with other state-of-the-art deep learning based denoising methods. Our experiments indicate that ATDNet achieved the highest PSNR and SSIM values for noise levels 15, 25 and 50 dB. In most cases, the results demonstrate that the proposed ATDNet either provides comparable results or surpasses the other denoising techniques. Fig. 11 and Fig. 12 show the denoising effect of a sample image obtained from BSD68 and Kodak24 dataset for noise levels 50 dB and 25 dB respectively. The green box represents the magnification of the red box. From the results it is eveident that ATDNet has better denoising effect and and presents clear details regardless of noise levels. Compared with the most commonly used conventional methods such as BM3D and DnCNN, the PSNR in dB has improved by a factor of 3.54% and 0.66% respectively. Also a good denoising algorithm should achieve a trade-off between denoising effectiveness and computation time, and Table 7 compares the execution time and the complexity (total number of parameters) of different image denoising methods. Although ATDNet employs a twofold network structure that increases the number of parameters of the models, the execution time is comparable to other algorithms. (For an 512×512 image, the compuation time is only 0.02 s longer when compared to ECNDNet).

Table 3
Average SSIM for different methods on Set12.

Noise Level	Method	C.man	House	Peppers	Starfish	Monar	Airplan	Parrot	Lena	Barbara	Boat	Man	Couple	Average
σ = 15	BM3D (Dabov et al., 2007)	0.8957	0.8853	0.9096	0.9065	0.9383	0.9003	0.9007	0.8963	0.9230	0.8603	0.8726	0.8802	0.8794
	DnCNN (Zhang et al., 2018)	0.9220	0.8986	0.9395	0.9422	0.967	0.9265	0.9268	0.9172	0.9361	0.8902	0.9052	0.9068	0.9232
	SANet (Zhang et al., 2019)	0.9078	0.8920	0.9149	0.9094	0.9474	0.9091	0.9065	0.8996	0.9190	0.8588	0.8789	0.8802	0.9020
	ECNDNet (Tian et al., 2019)	0.9242	0.8992	0.94	**0.9438**	0.9667	0.9254	0.9261	0.9175	0.9352	0.891	0.9059	0.9083	0.9236
	ATDNet	**0.9259**	**0.9005**	**0.9417**	0.9435	**0.9677**	**0.9269**	**0.9278**	**0.9207**	**0.9386**	**0.8943**	**0.907**	**0.9119**	**0.9256**
σ = 25	BM3D (Dabov et al., 2007)	0.8477	0.8504	0.8701	0.8529	0.8996	0.8546	0.8549	0.8586	0.8850	0.8019	0.8089	0.8206	0.8504
	DnCNN (Zhang et al., 2018)	0.8873	0.8736	0.9131	0.9051	0.9462	**0.8922**	0.8892	0.8862	0.8992	0.8449	0.8516	0.8607	0.8874
	SANet (Zhang et al., 2019)	0.8665	0.8642	0.8848	0.8602	0.9159	0.8737	0.8637	0.8679	0.8758	0.8070	0.8178	0.8249	0.8602
	ECNDNet (Tian et al., 2019)	0.8884	0.8736	0.9112	**0.9067**	0.9449	0.8907	0.8879	0.8873	0.8995	0.8433	0.8509	0.8600	0.8870
	ATDNet	**0.8906**	**0.8767**	**0.9146**	0.9052	**0.9465**	0.8909	**0.8904**	**0.8925**	**0.9043**	**0.8478**	**0.8526**	**0.8662**	**0.8898**
σ = 50	BM3D (Dabov et al., 2007)	0.7625	0.7945	0.7792	0.7441	0.8071	0.7562	0.7697	0.7855	0.7847	0.6957	0.6985	0.7063	0.7570
	DnCNN (Zhang et al., 2018)	0.8197	0.8344	0.8371	0.8201	0.8824	0.8208	0.8180	0.8232	0.7962	0.7421	0.7427	0.7552	0.8076
	SANet (Zhang et al., 2019)	0.7856	0.8139	0.8100	0.7568	0.8382	0.7928	0.7884	0.8024	0.7713	0.7143	0.7115	0.7173	0.7752
	ECNDNet (Tian et al., 2019)	0.8167	0.8346	0.8456	**0.8245**	0.8868	0.8227	0.8182	0.8243	0.8046	0.7459	0.7440	0.7588	0.8106
	ATDNet	**0.8200**	**0.8414**	**0.8544**	0.8220	**0.8893**	**0.8236**	**0.8238**	**0.8340**	**0.8091**	**0.7504**	**0.7455**	**0.7681**	**0.8151**

• σ= 15,25,50 represent noise level in dB on Set12.
• The best results are marked in bold.

Table 4
Denoising results of ATDNet on 10 images from BSD68 dataset wih different noise levels.

Noise Level	Metric	Img1	Img2	Img3	Img4	Img5	Img6	Img7	Img8	Img9	Img10	Average
σ = 15	PSNR	28.40	32.80	32.70	33.34	30.19	35.38	31.37	29.06	31.62	32.24	31.72
	SSIM	0.8806	0.9385	0.9188	0.9300	0.9231	0.9447	0.9212	0.8940	0.9237	0.9101	0.9136
σ = 25	PSNR	25.95	30.33	30.23	30.73	27.58	33.22	28.75	26.28	29.13	29.81	29.24
	SSIM	0.7998	0.8990	0.8852	0.8772	0.8627	0.9129	0.8620	0.8107	0.8712	0.8502	0.8581
σ = 50	PSNR	23.17	27.48	27.00	27.61	24.68	30.47	25.47	23.12	26.11	26.83	26.26
	SSIM	0.6593	0.8213	0.8157	0.7621	0.7338	0.8472	0.7385	0.6429	0.7673	0.7348	0.7512

• σ= 15,25,50 represent noise level in dB on BSD68.

Table 5
Denoising results of ATDNet on 10 images from Kodak24 dataset wih different noise levels.

Noise Level	Metric	Img1	Img2	Img3	Img4	Img5	Img6	Img7	Img8	Img9	Img10	Average
σ = 15	PSNR	29.51	35.20	31.64	34.15	28.37	30.95	32.89	32.80	33.09	30.97	32.77
	SSIM	0.9004	0.9135	0.8876	0.8745	0.8927	0.8961	0.8681	0.8831	0.9052	0.8974	0.9056
σ = 25	PSNR	27.07	32.64	29.18	32.15	25.68	28.41	30.86	30.57	30.68	28.33	30.32
	SSIM	0.8288	0.8774	0.8230	0.8215	0.8061	0.8314	0.8071	0.8132	0.8559	0.8282	0.8535
σ = 50	PSNR	24.37	29.44	26.25	29.81	22.63	25.42	28.65	28.08	27.58	25.15	27.36
	SSIM	0.6845	0.7945	0.7080	0.7387	0.6273	0.6985	0.7212	0.6878	0.7552	0.6863	0.7536

• σ= 15,25,50 represent noise level in dB on Kodak24.

Table 6
PSNR (dB)/ SSIM comparison of methods on gray-scale image denoising. Dataset Set12, BSD68 and Kodak24 are used for noise levels 15, 25 and 50 dB.

Method	Set12			BSD68			Kodak24		
Noise Level	15	25	50	15	25	50	15	25	50
BM3D (Dabov et al., 2007)	32.37/ 0.8950	29.97/ 0.8507	26.72/ 0.7670	31.07/ 0.8721	28.57/ 0.8013	25.62/ 0.6862	–/–	–/–	–/–
WNNM (Gu et al., 2014)	32.70/ 0.8978	30.26/ 0.8558	27.05/ 0.7780	31.37/ 0.8766	28.83/ 0.8086	25.87/ 0.6981	–/–	–/–	–/–
EPLL (Zoran and Weiss, 2011)	32.14/ 0.8937	29.69/ 0.8446	26.47/ 0.7487	31.21/ 0.8823	28.68/ 0.8121	25.67/ 0.6875	–/–	–/–	–/–
CSF (Schmidt and Roth, 2014)	32.32/ 0.8914	29.84/ 0.8432	–/–	31.24/ 0.8730	28.74/ 0.8030	–/–	–/–	–/–	–/–
DNCNN (Chen and Pock, 2016)	32.86/ 0.8231	30.43/ 0.8874	27.18/ 0.8076	**31.73**/ 0.9129	29.23/ 0.8572	26.23/ 0.7483	32.77/ 0.9056	30.32/ 0.8535	27.36/ 0.7536
IRCNN (Zhang et al., 2017)	32.77/–	30.38/–	27.14/–	31.63/–	29.23/–	26.23/–	–/–	–/–	–/–
SANet (Zhang et al., 2019)	32.82/ 0.9020	30.37/ 0.8602	27.09/ 0.7752	31.68/ 0.8896	29.13/ 0.8244	26.10/ 0.7100	–/–	–/–	–/–
PSA-U (Aljadaany et al., 2019)	32.43/–	30.12/–	27.04/–	31.70/–	29.27/–	26.32/–	–/–	–/–	–/–
ECNDNet (Tian et al., 2019)	32.81/ 0.9236	30.39/ 0.8870	27.15/ 0.8106	31.71/ 0.9134	29.22/ **0.8581**	26.23/ 0.7511	32.74/ 0.9063	30.33/ 0.8534	27.36/ 0.7558
ATDNet	**32.95**/ **0.9256**	**30.55**/ **0.8898**	**27.31**/ **0.8151**	31.72/ **0.9136**	**29.24**/ **0.8581**	**26.26**/ **0.7512**	**32.81**/ **0.9075**	**30.41**/ **0.8553**	**27.44**/ **0.7583**

• – indicates no reference data
• The best results are marked in bold.
• 15,25,50 dB indicates the noise level in dB on datasets.

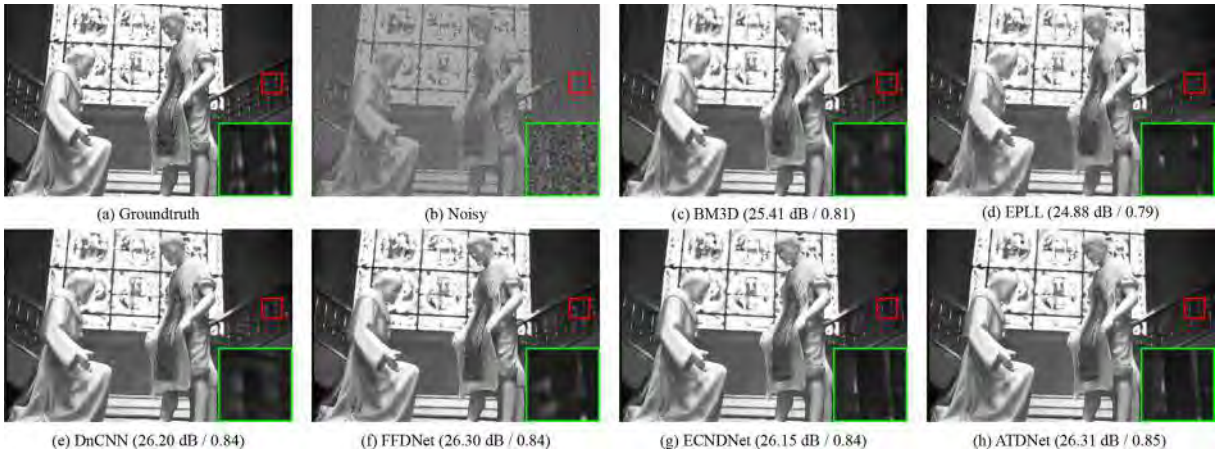


Fig. 11. Denoising results of a gray image from BSD68 dataset with noise level 50 dB.

Fig. 12. Denoising results of a gray image from Kodak24 dataset with noise level 25 dB.

Table 7
Complexity and Computation time for different denoising methods.

Methods	Device	Parameters	Computation time (s)	
			256×256	512×512
BM3D	CPU	-	0.59	2.52
WNNM	CPU	-	203.1	773.2
EPLL	CPU	-	25.4	45.5
ECNDNet	GPU	0.52 M	0.0030	0.0040
ATDNet	GPU	1.08 M	0.0070	0.0060

• – indicates no parameters

Furthermore in Fig. 13, we qualitatively compare the denoised images obtained from BM3D, EPLL, DnCNN, FFDNet, ECNDNet and ATDNet on Set12. As discussed in Section 3.3, we employ the contour map of Intra-SSIM, computed from the 3×3 window placed over the SSIM subblocks. The reconstruction ability of the model is measured from the intensity of the pixels that represents the structural properties within the image. Specifically, when the pixel intensities are similar, the contour maps alike (yellow as in our case). whereas, if the intensities differ, the contour maps appear in various shades of blue as shown in Fig. 13. For example, in Fig. 13, the green box indicates the enlarged region of the denoising image. We can visually see the pixel distribution of the ATDNet output within this region which is significantly better than other methods, such as BM3D and DnCNN.

Finally, the proposed ATDNet was subjected to three experiments and the results illustrate that the denoising ability of ATDNet outperforms other methods. The noise extracted from both the top and bottom network complement each other in the recovery of images that were subjected to noise levels of a maximum 50 dB. The Intra-SSIM contour plots clearly illustrate that the pixel intensities from the denoised image are visually closer to the ATDNet model outputs. For the noise level of 50 dB, ATDNet presents the highest PSNR and SSIM on Set12 of 27.31 and 0.8151 when compared to the other traditional techniques.

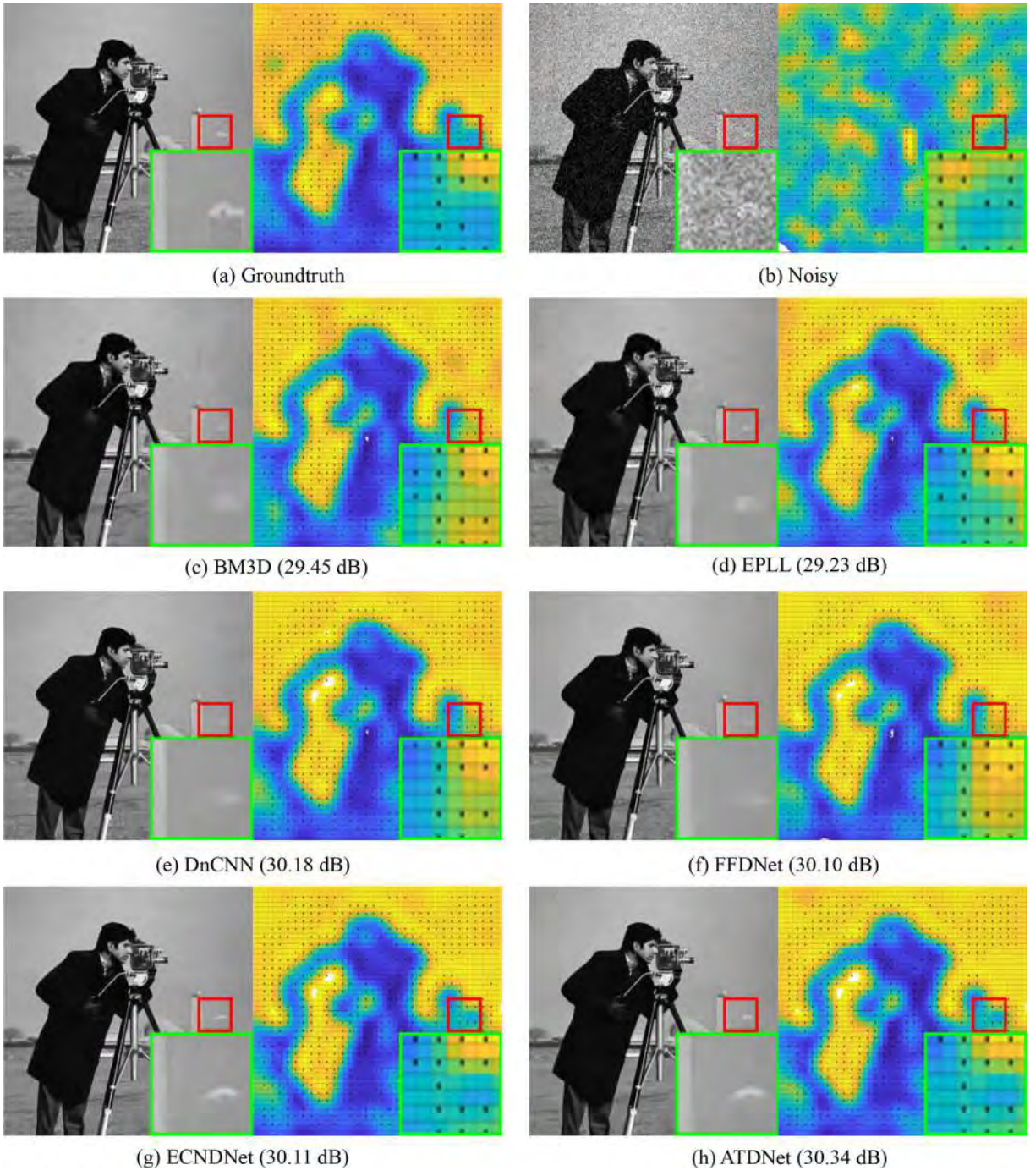


Fig. 13. Denoising results and contour reconstruction of a gray image from Set12 dataset with noise level 25.

5. Conclusion

In this paper, we proposed a denoising method based on deep learning. In order to obtain more noise information from the noisy images, we created a twofold convolutional neural network to complement each other's feature information and used an attention mechanism for reinforcement. A structural analysis of the proposed network is performed to illustrate the role played by each block in the denoising task, and experimentally demonstrated that the denoising performance of the twofold network outperforms that of the single network. For the activation function, Tanh is used to process the fused noise information and demonstrate its superiority experimentally. In addition, the performance of the proposed method is analyzed and compared with the state-of-the-art methods on datasets such as Set12, BSD68 and kodak24 with the noise levels of 15, 25 and 50 db using PSNR and SSIM. In most cases, the proposed model achieves good denoising results. In particular, the metric Intra-ssim is used to plot the contour maps of the different denoising results, which makes it easier to directly observe that the recovered details are closer to the original image. In the future, it is planned to embed the proposed network into grayscale image optimization methods for real scenes, such as CT images and MRI

images, to further develop the capability of our twofold network. Among the conducted experiments, ablation experiments proved that features extracted from dual networks are richer, and different activation functions were used in the experiments to demonstrate which activation function is more suitable for feature fusion.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgment

This work was supported by the Scientific research Grant of Shantou University, China, Grant (No. NTF17016), Basic and Applied Basic Research Foundation of Guangdong Province (No. 2020B1515120061) and National Natural Science Foundation of China (No. 82071992).

References

Aljadaany, R., Pal, D.K., Savvides, M., 2019. Proximal splitting networks for image restoration. In: International Conference on Image Analysis and Recognition. Springer, pp. 3–17.

Alsaiari, A., Rustagi, R., Thomas, M.M., Forbes, A.G., et al., 2019. Image denoising using a generative adversarial network. In: 2019 IEEE 2nd International Conference on Information and Computer Technologies (ICICT). IEEE, pp. 126–132.

Ang-bo, J., Wei-wei, W., 2018. Research on optimization of relu activation function. Transducer Microsyst. Technol. 2.

Arbelaez, P., Maire, M., Fowlkes, C., Malik, J., 2010. Contour detection and hierarchical image segmentation. IEEE Trans. Pattern Anal. Machine Intell. 33 (5), 898–916.

Borji, A., Itti, L., 2013. State-of-the-art in visual attention modeling. IEEE Trans. Pattern Anal. Mach. Intell. 35 (1), 185–207.

Burger, H.C., Schuler, C.J., Harmeling, S., 2012. Image denoising: Can plain neural networks compete with bm3d? 2012 IEEE Conference on Computer Vision and Pattern Recognition. IEEE, pp. 2392–2399.

Chen, Y., Pock, T., 2016. Trainable nonlinear reaction diffusion: A flexible framework for fast and effective image restoration. IEEE Trans. Pattern Anal. Machine Intell. 39 (6), 1256–1272.

Chen, L.-C., Yang, Y., Wang, J., Xu, W., Yuille, A.L., 2016. Attention to scale: Scale-aware semantic image segmentation. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3640–3649.

Chen, D., He, Z., Cao, Y., Yang, J., Cao, Y., Yang, M.Y., Tang, S., Zhuang, Y., 2019. Deep neural network for fast and accurate single image super-resolution via channel-attention-based fusion of orientation-aware features, arXiv preprint arXiv:1912.04016.

Dabov, K., Foi, A., Katkovnik, V., Egiazarian, K., 2007. Image denoising by sparse 3-d transform-domain collaborative filtering. IEEE Trans. Image Process. 16 (8), 2080–2095.

Duan, Y., Chen, W., Feng, Q., Ma, J., 2011. Gradient-weighted ssim based medical image quality assessment. Jisuanji Gongcheng yu Yingyong(Comput. Eng. Appl.) 47 (24), 205–210.

Elad, M., Aharon, M., 2006. Image denoising via sparse and redundant representations over learned dictionaries. IEEE Trans. Image Process. 15 (12), 3736–3745.

Ephraim, Y., Malah, D., 1984. Speech enhancement using a minimum-mean square error short-time spectral amplitude estimator. IEEE Trans. Acoust. Speech Signal Process. 32 (6), 1109–1121.

Fan, C.-M., Liu, T.-J., Liu, K.-H., 2022. Sunet: Swin transformer unet for image denoising, arXiv preprint arXiv:2202.14009.

Franzen, R., 1999. Kodak lossless true color image suite, source: http://r0k.us/graphics/kodak, pp. vol. 4, 1999.

Fukui, H., Hirakawa, T., Yamashita, T., Fujiyoshi, H., 2019. Attention branch network: Learning of attention mechanism for visual explanation. In: Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition, pp. 10705–10714.

Geng, T., Liu, X.-Y., Wang, X., Sun, G., 2021. Deep shearlet residual learning network for single image super-resolution. IEEE Trans. Image Process. 30, 4129–4142. https://doi.org/10.1109/TIP.2021.3069317.

Gu, S., Zhang, L., Zuo, W., Feng, X., 2014. Weighted nuclear norm minimization with application to image denoising. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2862–2869.

He, K., Zhang, X., Ren, S., Sun, J., 2016. Deep residual Learning for Image Recognition. IEEE.

He, W., Zhang, H., Shen, H., Zhang, L., 2018. Hyperspectral image denoising using local low-rank matrix recovery and global spatial–spectral total variation. IEEE J. Sel. Top. Appl. Earth Obser. Remote Sens. 11 (3), 713–729.

Hoffman, J.E., 1998. Visual attention and eye movements. Attention 31 (2), 119–153.

Ioffe, S., Szegedy, C., 2015a. Batch normalization: Accelerating deep network training by reducing internal covariate shift. In: International Conference on Machine Learning, PMLR, pp. 448–456.

Ioffe, S., Szegedy, C., 2015b. Batch normalization: Accelerating deep network training by reducing internal covariate shift. JMLR.org.

Jain, V., Seung, H.S., 2008. Natural image denoising with convolutional networks. In: International Conference on Neural Information Processing Systems.

Jifara, W., Jiang, F., Rho, S., Cheng, M., Liu, S., 2019. Medical image denoising using convolutional neural network: a residual learning approach. J. Supercomput. 75 (2), 704–718.

Kai, Z., Zuo, W., Gu, S., Lei, Z., 2017. Learning deep cnn denoiser prior for image restoration. In: 2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR).

Kalchbrenner, N., Espeholt, L., Simonyan, K., Oord, A.V.D., Graves, A., Kavukcuoglu, K., 2016. Neural machine translation in linear time. arXiv preprint arXiv:1610.10099.

Kalman, B.L., Kwasny, S.C., 1992. Why tanh: choosing a sigmoidal function. [Proceedings 1992] IJCNN International Joint Conference on Neural Networks, vol. 4. IEEE, pp. 578–581.

Kim, S.-W., Kook, H.-K., Sun, J.-Y., Kang, M.-C., Ko, S.-J., 2018. Parallel feature pyramid network for object detection. In: Proceedings of the European Conference on Computer Vision (ECCV), pp. 234–250.

Kingma, D.P., Ba, J., 2014. Adam: A method for stochastic optimization. arXiv preprint arXiv:1412.6980.

Li, W., Raj, A.N.J., Tjahjadi, T., Zhuang, Z., 2021. Digital hair removal by deep learning for skin lesion segmentation. Pattern Recogn. 117, 107994.

Liu, Z., Lin, Y., Cao, Y., Hu, H., Wei, Y., Zhang, Z., Lin, S., Guo, B., 2021a. Swin transformer: Hierarchical vision transformer using shifted windows. In: Proceedings of the IEEE/CVF International Conference on Computer Vision, pp. 10012–10022.

Liu, X., Zhang, L., Li, T., Wang, D., Wang, Z., 2021b. Dual attention guided multi-scale cnn for fine-grained image classification. Inf. Sci. 573, 37–45.

Lu, H., Li, Y., Chen, M., Kim, H., Serikawa, S., 2018. Brain intelligence: go beyond artificial intelligence. Mobile Networks Appl. 23 (2), 368–375.

Lv, B., Liu, Y., Zhang, S., Zeng, H., Zhu, G., 2018. Super resolution with generative adversarial networks, ArXiv preprint.

Mao, X., Shen, C., Yang, Y.-B., 2016. Image restoration using very deep convolutional encoder-decoder networks with symmetric skip connections. Adv. Neural Informat. Process. Syst. 29, 2802–2810.

Oord, A.V.D., Dieleman, S., Zen, H., Simonyan, K., Vinyals, O., Graves, A., Kalchbrenner, N., Senior, A., Kavukcuoglu, K., 2016. Wavenet: A generative model for raw audio, arXiv preprint arXiv:1609.03499.

Pan, J., Liu, S., Sun, D., Zhang, J., Liu, Y., Ren, J., Li, Z., Tang, J., Lu, H., Tai, Y.W., 2018. Learning Dual Convolutional Neural Networks for Low-level Vision. IEEE.

Rawat, S., Rana, K., Kumar, V., 2021. A novel complex-valued convolutional neural network for medical image denoising. Biomed. Signal Process. Control 69, 102859.

Redmon, J., Divvala, S., Girshick, R., Farhadi, A., 2016. You only look once: Unified, real-time object detection. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 779–788.

Romano, Y., Elad, M., Milanfar, P., 2017. The little engine that could: Regularization by denoising (red). SIAM J. Imag. Sci. 10 (4), 1804–1844.

Roth, S., Black, M.J., 2005. Fields of experts: A framework for learning image priors. 2005 IEEE Computer Society Conference on Computer Vision and Pattern Recognition (CVPR'05), vol. 2. IEEE, pp. 860–867.

Schmidt, U., Roth, S., 2014. Shrinkage fields for effective image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 2774–2781.

Tai, Y., Yang, J., Liu, X., Xu, C., 2017. Memnet: A persistent memory network for image restoration. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 4539–4547.

Tian, C., Xu, Y., Fei, L., Wang, J., Wen, J., Luo, N., 2019. Enhanced cnn for image denoising. CAAI Trans. Intell. Technol. 4 (1), 17–23.

Tong, S., Gu, H., Yu, K., 2016. A comparative study of robustness of deep learning approaches for vad. In: 2016 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP). IEEE, pp. 5695–5699.

Wang, Z., Liu, D., Yang, J., Han, W., Huang, T., 2015. Deep networks for image super-resolution with sparse prior. In: Proceedings of the IEEE International Conference on Computer Vision, pp. 370–378.

Wei Li, T.T., Raj, Alex Noel Joseph, Zhuang, Z., 2022. Fusion of anns as decoder of retinal spike trains for scene reconstruction. Appl. Intell. https://doi.org/10.1007/s10489-022-03402-w.

Xie, J., Xu, L., Chen, E., 2012. Image denoising and inpainting with deep neural networks. Adv. Neural Informat. Process. Syst. 1.

Yamanaka, J., Kuwashima, S., Kurita, T., 2017. Fast and accurate image super resolution by deep cnn with skip connection and network in network. In: International Conference on Neural Information Processing. Springer, pp. 217–225.

Yu, F., Koltun, V., 2015. Multi-scale context aggregation by dilated convolutions, arXiv preprint arXiv:1511.07122.

Zhang, K., Zuo, W., Chen, Y., Meng, D., Zhang, L., 2017. Beyond a gaussian denoiser: Residual learning of deep cnn for image denoising. IEEE Trans. Image Process. 26 (7), 3142–3155.

Zhang, K., Zuo, W., Gu, S., Zhang, L., 2017. Learning deep cnn denoiser prior for image restoration. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 3929–3938.
Zhang, K., Zuo, W., Zhang, L., 2018. Ffdnet: Toward a fast and flexible solution for cnn-based image denoising. IEEE Trans. Image Process. 27 (9), 4608–4622.
Zhang, Y., Tian, Y., Kong, Y., Zhong, B., Fu, Y., 2018. Residual Dense Network for Image Super-resolution. IEEE.
Zhang, L., Li, Y., Wang, P., Wei, W., Xu, S., Zhang, Y., 2019. A separation–aggregation network for image denoising. Appl. Soft Comput. 83, 105603.
Zhang, K., Li, Y., Zuo, W., Zhang, L., Van Gool, L., Timofte, R., 2021. Plug-and-play image restoration with deep denoiser prior. IEEE Trans. Pattern Anal. Mach. Intell.
Zheng, H., Wang, X., Gao, X., 2018. Fast and accurate single image super-resolution via information distillation network. In: 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR).
Zhuang, Z., Yang, Z., Raj, A.N.J., Wei, C., Jin, P., Zhuang, S., 2021. Breast ultrasound tumor image classification using image decomposition and fusion based on adaptive multi-model spatial feature fusion. Comput. Methods Programs Biomed. 208, 106221.
Zhu, Z., Wu, W., Zou, W., Yan, J., 2018. End-to-end flow correlation tracking with spatial-temporal attention. In: Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pp. 548–557.
Zoran, D., Weiss, Y., 2011. From learning models of natural image patches to whole image restoration. In: 2011 International Conference on Computer Vision. IEEE, pp. 479–486.

International Journal of Electronics

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tetn20

Optimal design of improved H-infinity controller for MMC

Rashmi N

To cite this article: Rashmi N (2023): Optimal design of improved H-infinity controller for MMC, International Journal of Electronics, DOI: 10.1080/00207217.2023.2173807

To link to this article: https://doi.org/10.1080/00207217.2023.2173807

INTERNATIONAL JOURNAL OF ELECTRONICS
https://doi.org/10.1080/00207217.2023.2173807

Taylor & Francis
Taylor & Francis Group

Optimal design of improved H-infinity controller for MMC

Rashmi N

Assistant Professor, Electronics and Communication Engineering, BMS Institute of Technology and Management, Bengaluru, India

ABSTRACT

In MMC, the variations in capacitor voltage of sub-modules produce the circulating current. This research work presented intends to develop the MMC-HVDC using an the most effective 3 H-infinity regulator for solar energy systems. Initially, the unstable currents are observed using the new sensor and are compared to the current controller, which produces an erroneous current flowing. The current controller is set to zero. Since variations amongst error circulating current adds reference mistakes was expressed as a non-linear meaningful purpose, it is thought to be the primary factor in lowering the total harmonic distortion (THD). Using the Self-Adaptive Cuckoo search algorithm (SA-CS) for signal control, the gain is maximised in the H-infinity controller. Regulate evaluation, conversion performance monitoring, and battery voltage analysis are also used in the performance calculation of the proposed model in addition to THD and circulating current analysis. The results of the simulation show how well the proposed control strategy operates.

ARTICLE HISTORY
Received 3 January 2022
Accepted 22 January 2023

KEYWORDS
MMC; H-infinity controller; self-adaptive cuckoo search; circulating current; total harmonic distortion

1. Introduction

Modular Multilevel Converter (MMC) topology includes rising significance owing to its extraordinary features such as the unproblematic substitution of fault SMs, modular structure and also reduced values of rating (Babikir Bashir & Beig, 2018; Beddard & Barnes, 2015; Haifeng et al., 2015; Jinke et al., 2017; Shi et al., 2018; Wang et al., 2017). The MMCs exploit Pulse Width Modulation (PWM) regulating methods for a quicker response. Along with PWM regulating approaches, MMCs accomplish increased control and speed of reactive and active powers by sustaining a constant frequency and voltage that consecutively assists in the noteworthy minimisation of changing costs. The exploitation of MMC-dependent converters assists in mitigating voltage stress (Božiček et al., 2017; Vidal-Albalate et al., 2016; Guo et al., 2018; Satyanarayana & Lakshmi Ganesh, 2015). As a result, it helps to reduce overall footprint and expense. The typical indicators of the MMC's operation are (i) line current ratings, (ii) transformer ratings and (iii) capacitor voltage ripples (Guoqing et al., 2018; Jianzhong et al., 2016; Raymundo & Garces, 2017). Voltage source converters (VSC)-based high voltage direct current (HVDC) technologies are being used in a number of applications, such as the combination of offshore wind generation and long-distance transmission system (Tavakoli et al., 2021). MMC-oriented

CONTACT Rashmi N nrashmi814@gmail.com BMS Institute of Technology and Management, Bengaluru, Karnataka 560064, India

HVDC, which can handle inactive devices and has a hardware platform, has minimal overtones, variable stability, and is becoming a popular choice for incorporating marine renewable power (Ji et al., 2021).

MMC exploitation is raising in High-Voltage Direct Current (HVDC) properties (Guo et al., 2018; Mehrasa et al., 2018), owing to its significance over Line Commutated Converters (LCCs) and Voltage Source Converters (VSC). The MMC-HVDC includes the noteworthy improvement in the area of RES (Renewable Energy Systems), the major development of the VSC-HVDC transmission system (Rui & Fletcher, 2017; Venkat Ram Reddy et al., 2017; Xia et al., 2016). Accordingly, sinusoidal channels can produce the necessary control and voltage for MMC. It improves the harmonic features and could be carried out in DCs. Consequently, it is extensively deployed in the balancing of MMCs (Cretnik et al., 1992; Liu et al., 2014). In addition, the grid settings affect the powerful abilities and functioning region of the MMC. Predictive current control (PCC) has risen as a well-known Model Predictive Control (MPC) method in recent times that increasing in MMC appliances, as the execution of PCC in MMC turns out to be more demanding as the amount of SMs raises (Guo et al., 2018; Ramirez et al., 2017; Zhang & Zhao, 2016). The best possible environment that may be achieved for an aim under the limitations that are present is known as global optimisation (Madichetty et al., 2016; Mohana et al., 2016; Yang, 2014).

This paper develops the MMC-HVDC system that incorporates the logical concept of optimisation. For a point, this ideal H-infinity regulator is appropriate. At first, the control signal is reset to zero, and the current sensor detects instability currents at each stage, leading to the generation of an erroneous current. Next to this, the proposed SA-CS is used to control the signal, which optimises the gain of the H-infinity microcontroller. Finally, the performance is compared to other standard approaches like the PI controller, H-infinity-FF, H-infinity-PSO (Particle Swarm Optimisation) and H-infinity-CS in terms of regulation evaluation, conversion efficiency and voltages analysation (Cuckoo Search).

The structure of the essay is as follows. The associated research and reviews are shown in Section 2. A modular multilevel converter is depicted in Section 3, and the suggested current controller is discussed in Sections 4 and 5 that also illustrate the H-infinity controller including optimisation in action. Section 6 summarises results, and Section 7 wraps up the report.

2. Literature review

2.1 Related works

In 2017, Eduardo *et al*. (Prieto-Araujo et al., 2017) implemented a control design MMC which is possible to function under varied conditions. Initially, a steady-state investigation was carried out to recognise the deployment of present elements contained by the control scheme. Accordingly, a thorough modelling process was involved for the current and energy regulators, for making a certain better response. At last, hypothetical improvements were authenticated through simulation outcomes using a comprehensive representation under varied conditions.

In 2017, Venkat *et al*. (Venkat Ram Reddy et al., 2017) adopted a technique based on counting sort in combination with CPD PWM. The increased harmonics in current

waveforms might be accredited to the random capacitors. This whole design was deployed for HVDC models for justifying the THD in voltage and current correspondingly. In addition, a constancy investigation for introduced controllers was carried out to establish their robust function, and at last, the entire results were surveyed.

In 2018, Majid *et al*. (Mehrasa et al., 2018) suggested a dynamic six-order representation of MMC for applications in HVDC. For the primary examination of MMC function, a curve depending on reactive and active power of MMC was accomplished through a measurement of the implemented formulation. Finally, Matlab/Simulink setting was deployed to systematically confirm the capability of the introduced method for regulating MMC beneath both MMC's constraint variations and load.

In 2015, Fernando *et al*. (Martinez-Rodrigo & Santiago de Pablo, 2015) established two novel schemes for current control where a band was described regarding the three-phase current. In the initial approach, voltage levels nearby to the grid voltage level were selected and were known as 'constant excitation'. This was suitable while the count of (Sub Modules) SM for each arm was diminutive. The subsequent scheme was a most suitable one when the amount of SMs for each arm was enormous. The hypothetical establishment of the approaches was also offered, and at last, the current control approaches were described in a better way.

In 2017, Ardavan *et al*. (Mohammad Hassani et al., 2017) provided a method for achieving balance and consistent MMC control while decreasing the ac signal of a harmonic currents. A suggested method has exploited GA to generate the optimum lower- and upper-arm manifestations. The scheme subsequently generates the equivalent patterns of switching by the exploitation of SM categorisation approach. The weighting factors were chosen depending on simulation outcomes. Moreover, the simulations were performed with probable expansion to advanced levels and stages. Finally, hypothetical examination and outcomes were offered and described that substantiate the efficiency of the suggested scheme.

In 2018, Lee and Heng (Sing Lee & En Heng, 2018) established a novel scheme based on a tuning-less model of MMC for a three-phase MMC. Accordingly, 3 control phases were modelled to provide for several control intentions within MMC devoid of requirement for any adjustment in tuning the weights. A theoretically uncomplicated power control scheme was introduced into the suggested MPC. Finally, the obtainable simulation outcomes substantiated the accuracy and probability of the presented control approach.

In 2017, Ahmed *et al*. (Elserougi et al., 2017) proposed the new fusion-based MMC that facilitates the functioning with a minimised amount of SMs with a similar boosting ability; as a result, it facilitates transformer-less function. Accordingly, a comprehensive demonstration of the established STATCOM formation together with its controllers was offered. Simulation resultants also demonstrate capable and acceptable performance.

In 2016, Gebreel and Xu (Almula Gebreel & Longya, 2016) suggested a scheme that influences output current and voltage THD response depending a tight circuit (Proportional-Integral-Controller) PID controller representation. Accordingly, the association is sandwiched between the maximum resultant outcomes and arm inductance as described. At last, the implemented models of the proposed work were offered to authenticate the value exploratory testing of the elbow isolator using MMC. Finally, the investigational outcomes were also described, and the superiority of the introduced scheme was verified.

In 2021, Dadjo (Tavakoli et al., 2021) presented using multivariable H optimum control approaches to handle such a complicated dynamic system guarantee stability and VSC-HVDC optimal performance link. The network actual performance is initially measured using a filled, centralised, multi-input multi-output (MIMO) controller that is likewise derived based on H optimisation. The transition from a high-order, central system to a low-order, distributed controller results in efficiency and reliability loss, which is evaluated towards the end through simulations.

2.2 Review

The methodologies, characteristics and difficulties of traditional techniques based on controls for MMC systems are shown in Table 1. HVDC was initially used in Prieto-Araujo et al. (2017) because it provides more regulation and lessens energy deviations. However,

Table 1. Review of conventional models of controllers for MMC.

Author [citation]	Adopted methodology	Features	Challenges
Salonitis *et al.* (Prieto-Araujo et al., 2017)	HVDC	• Offers increased regulation • Reduces the energy deviations	• Possibilities for loss in converter
Venkat *et al.* (Venkat Ram Reddy et al., 2017)	CPD PWM.	• Increased robustness. • Reduced delay.	• Problems emerge if grid voltage is considered as a reference.
T. Ganesan *et al.* (Mehrasa et al., 2018)	PCT	• Offers stable operation • Provides accurate functions	• Increase in cost owing to circulating currents
Shifeng Wen *et al.* (Martinez-Rodrigo & Santiago de Pablo, 2015)	HVDC	• Offers minimized error • Reduces the energy deviations	• Possibilities for loss in converter
Himanshu and Ravi (Mohammad Hassani et al., 2017)	GA	• Minimized cost function • Offers constant frequency.	• Increased internal loss
T. Ganesan *et al.* (Sing Lee & En Heng, 2018)	Voltage balancing algorithm	• Offers greater flexibility • Reduced computation time	• Challenging real-time implementation.
Kumaravadivel and Natarajan (Elserougi et al., 2017)	PWM scheme	• Includes boosting capacity • Limited fault occurrence	• No DC fault blocking ability.
B. Surekha *et al.* (Almula Gebreel & Longya, 2016)	Tuning-less algorithm	• Minimizes the loss of converter • Offers large motor control	• High switching frequency increases the loss.
Dadjo *et al.* (Ji et al., 2021)	MIMO	• Simple design • The process is balanced one.	• At the design stage limited amount of controller is used.

there may be possibilities for loss in the converter. In addition, CPD PWM was deployed in Venkat Ram Reddy et al. (2017) which provides increased robustness and reduced delay, but problems could emerge if grid voltage is considered as a reference. PCT was suggested in Mehrasa et al. (2018) that provides stable operation along with accurate functions. Anyhow, there is an increase in cost owing to circulating currents. And then, HVDC employed in Martinez-Rodrigo & Santiago de Pablo (2015) which offers minimised error combine reduced energy deviations, but there may be possibilities for loss in the converter. GA was introduced in Mohammad Hassani et al. (2017) which offer minimised cost function and constant frequency. However, there was an increase in internal loss. In addition, a voltage balancing algorithm was deployed in Sing Lee & En Heng (2018) that offers greater flexibility with reduced computation time, but it includes challenging real-time implementation. PWM algorithm was suggested in Elserougi et al. (2017) that includes boosting capacity along with limited fault occurrence. Anyhow, there was no DC fault blocking ability. The tuning-less algorithm was employed in Almula Gebreel & Longya (2016) that minimises the loss of the converter, and offers large motor control, however, a rise in switching frequency increases the loss. The MIMO method is implemented in Tavakoli et al. (2021); the main advantage is that the design is a simple one and a limited amount of controller is used at the initial level.

2.3 Objectives

- To minimise the distortion.
- To improve the gain.
- To reduce the computation time.

3. Modular multilevel converter

Figure 1 illustrates the fundamental three-phase and single-phase MMC designs. A lot of SMs, each of which includes two valves coupled in anti-parallel mode by diode, make up an MMC (Zhang & Zhao, 2016). Two windings, high and low, are required for each phase of MMC and are connected by an arm inductance L_0 that is chiefly exploited to control the fault current taking place due to converter leg with a path fault. SM switches were indicated as S_1 and S_2 also with D1 and D2 diodes. Similarly, SMs might be set up as open systems as a whole in Figure 1. $V_{sm}.I_{s1}$, I_{s2}, I_{cu}, I_{cir} and V_{cu} indicate the voltage across SMs, the current flowing via the SM switch 1, SM switch 2, and SM capacitor and SM capacitor respectively.

As stated previously, capacitor dynamics are chiefly accountable for the voltage imbalance of SMs. While the switch of SM1 is kept, the current flows via capacitor, which was known as capacitor current, i_{c1}. Similarly, i_{c-n} denotes the SM current of n^{th}capacitor. From the switching arrangement of SM capacitors, it is clear that the SM switch controls how much current flows through the capacitor. For example, a time-altering current could be represented as given by Eq. (1), which $S_1(t)$ denotes the switch's status.

$$i_{c_1}(t) = S_1(t).i_a(t) \tag{1}$$

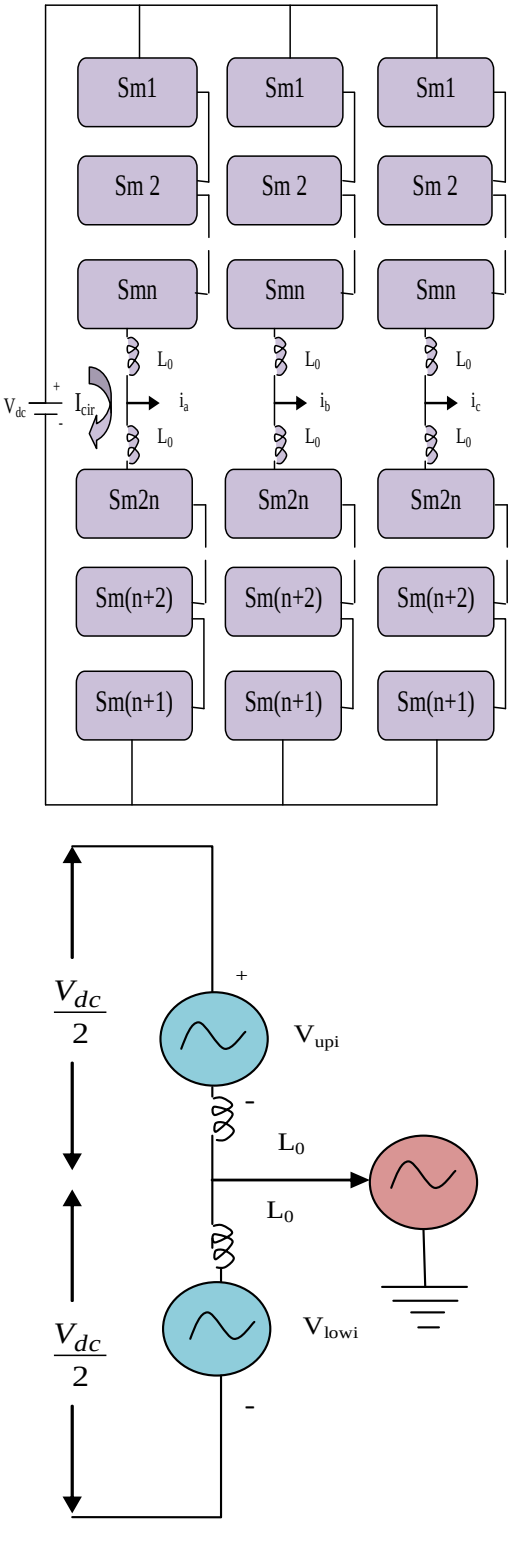


Figure 1. Basic three-phase MMC with circulating currents.

When status of $S_1(t)$ is on, it can be taken as '1' or else it is indicated by '0'. Moreover, the SM capacitor voltage has to be examined. Therefore, the SM capacitor ripple voltage V_{cu-1} could be evaluated as given in Eq. (2), where n and$f$ are the n^{th} current harmonic component and frequency correspondingly.

$$V_{cu1} = \frac{i_{c1}(t)}{(2\pi f)(n).C_{s1}} \tag{2}$$

On the other hand, owing to the switching behaviour of PWM, the capacitor acquires a DC component. Therefore, the entire SM capacitor voltage V_{cSM1} can be indicated as given in Eq. (3).

$$V_{cSM1} = S_1(t).\left(\frac{V_{dc}}{2}\right) + S_1(t).V_{cu_1}(t) \tag{3}$$

The lower arm voltage $V_{al}(t)$ with upper arm voltage $V_{au}(t)$ of phase 'a' is given by Eq. (4) and Eq. (5) respectively.

$$V_{al}(t) = \sum_{i=n+1}^{2n} S_i(t)\left(\frac{V_{dc}}{2}\right) + \sum_{i=n+1}^{2n} S_i(t).V_{cli}(t) \tag{4}$$

$$V_{au}(t) = \sum_{i=1}^{n} S_i(t)\left(\frac{V_{dc}}{2}\right) + \sum_{i=1}^{n} S_i(t).V_{cui}(t) \tag{5}$$

The phase 'a' voltage $V_{pha}(t)$ is given by Eq. (6) which is formulated from Eq. (4) and Eq. (5).

$$V_{pha}(t) = \sum_{i=1}^{n} S_i(t)\left(\frac{V_{dc}}{2}\right) + \sum_{i=1}^{n} S_i(t).V_{cui}(t) + \sum_{i=n+1}^{2n} S_i(t)\left(\frac{V_{dc}}{2}\right) + \sum_{i=n+1}^{2n} S_i(t).V_{cli}(t) \tag{6}$$

The circulating current can be obtained from Eq. (6) and it is given in Eq. (7).

$$i_{cir} = \frac{V_{DC} - V_{pha}(t)}{(2\pi f)(n)L_0} \tag{7}$$

From Equation 7, it is obvious that the circulating currents are available in converter stages owing to capacitor instability. Therefore, a stable approach with improved speed (Sreedhar & Mishra, 2014) is necessary to change over the SM switches according to necessity.

4. H-infinity controller with optimisation

4.1 Proposed H-infinity controller

H-infinity models are exploited in control theory to combine controllers for attaining stabilisation with assured performance. By using the H-infinity notion, a control designer can represent the control scheme as a problem with mathematical programming and then find the device that takes care of this improvement. H-infinity techniques include the benefit over traditional control models in that they are appropriate to issues linking cross-coupling with multivariate systems among channels. The expression of the H-infinity is

given by Equation 8, where A, B, C and D indicate the matrix and K_1, K_2, K_3 and K_4 denote the gain factor. The formulations of A, B, C and D are given by Equation 14 - Equation 17.

U is the manipulated variable/controller output, andV_1 is the measured variable. We are taking the error Z into account together with the measured variable.

$$U = K(S_1) * (V_1 * z) \tag{8}$$

where K is a matrix,

$$U = V_1 * Z \tag{9}$$

Here,

$$z = F(L, M) * e \tag{10}$$

$$F(L, M) = DK_4 + BK_2(1 - AK_1)^T CK_3 \tag{11}$$

$$\left(\left[BK_2\left[(I - AK_1)^T[CK_3]\right]\right] + DK_4\right).e\left(\frac{u}{||u^2||}\right).\frac{u}{||u^2||} = U \tag{12}$$

$$e = \frac{U}{(BK_2(I - AK_1)^T CK_3 + DK_4)} \times \frac{||u^2||}{u}\left(\frac{||u^2||}{u}\right) \tag{13}$$

So that the $e$is determined by using Eq.(13)

$$A = \left[-1\ 1\ \ -1\ \frac{1}{L_0} - \frac{1}{L_0}\ \frac{1}{L_0}\ \frac{1}{C_{SM_{reg}}} - \frac{1}{C_{SM_{reg}}}\ 0\right] \tag{14}$$

$$\begin{aligned} B_1 &= \left[1\ \ 0\ \ -\tfrac{1}{L_0}\ 0\ \ 0\ 0\right] \\ B_2 &= [1\ \ 0\ \ 0\ ] \end{aligned} \tag{15}$$

$$\begin{aligned} C_1 &= [\ 0\ \ -1\ \ 0] \\ C_2 &= [\ 0\ \ 1\ \ ] \end{aligned} \tag{16}$$

$$D_2 = 0 \tag{17}$$

4.2 Gain encoding and objective function

The elements of gain encoding such as K_1, K_2, K_3 and K_4 are given as an approach to enhancement, which is given in Figure 2. The gain factors are randomly chosen. With improved gain factors, the fundamental purpose of the suggested model is to minimise

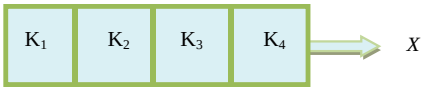


Figure 2. Solution encoding.

THD. The gain factors are optimally chosen by using the SA-CS Algorithm. Here X represents the proposed algorithm.

4.3 Conventional CS algorithm

An technique that was developed based on cuckoos' reproduction is the CS algorithm. In order to reduce THD, the CS algorithm accepts data from rise parameters including K_1,K_2, K_3 and K_4 (Mareli & Twala, 2017). Usually, cuckoos would lay their eggs in the nests of other cuckoos in the hopes that other parents would rear the young. In these situations, the unusual eggs are ejected from the clutches. The following three conditions must be met for the CS strategy to work. The best nests containing eggs are passed down to the next. A host cuckoo may also find another egg also with probability $P\varepsilon[0, 1]$. A parent bird has two options: throw the egg or leave the home and create a fresh ones.

A portion of host nests c in comparison to other ones could be used to estimate the end state. A fitness Q_i of solution combining matched to optimal value. Every eggshell in a nest defines an answer, and every cuckoo might only have one solution, according to the example that was used. The goal here is to substitute the worst solution already in the nest with a fresh, effective one. As a result, Equations (18) and (19) provide the balance between locally and globally random walks, respectively. Also in Eq. (18), $X_i^t$and X_k^t represent the positions selected through random permutation, β showing the favourable variable step scaling factor, X_i^{t+1} identifies the following place, s shows the magnitude of the step, $\otimes$ is used to describe the product of 2 vectors. F represents the heavy side of the function. P signifies constraints; this is used to alternate between global and random walks. ε relates to the uniform distribution's random parameter. At the same time Eq. (19), $M(s,\tau)$ shows that the the arbitrary walk's step size was determined by the Levy distribution.

$$X_i^{t+1} = X_i^t + \beta\, s \otimes F(P-\varepsilon) \otimes \left(X_j^t - X_k^t\right) \tag{18}$$

$$X_i^{t+1} = X_i^t + \beta\, M(s,\tau) \tag{19}$$

The pseudo-code of the existing CS algorithm is given by Algorithm 1.

Algorithm 1: Conventional CS Algorithm (Mareli & Twala, 2017)
Objective function, $f(X) = X = (X_1, X_2 \ldots X_D)^T$
Produce initial population of n host nests $Xi(i = 1, 2 \ldots .n)$
While $(t < Max\ Generation)$or terminate the process
arbitrarily identify a cuckoo using the Lévy distribution;
Compute its quality/fitness; Q_i
Choose a nest among n nests arbitrarily;
Evaluate its quality/fitness; Q_j
If $(Q_i > Q_j)$
Replace j by novel solution;
End
A fraction of (P) of worse nests are left, and novel nest are constructed at novel locations;
keep the finest solutions in place;
Sort answers to find the best (current);
End while
Post processing

4.4 Proposed SA-CS algorithm

Notwithstanding the CS algorithm's intriguing facts, it is simpler to settle for the local border optimal value. To avoid this shortcoming, SA-CS approach is proposed here, which modifies the scaling factor, βas shown by Eq. (20), which $\tau$is chosen based on Eq. (21). Moreover, Eq. (10) varies based on Eq. (20) for each iteration. In Eq. (21), $f(t-1)$ relates old iteration and $f(t)$ depicts a new step.

$$\beta = \frac{3}{2}\tau \tag{20}$$

$$\tau = \frac{f(t-1) - f(t)}{f(t-1)} \tag{21}$$

Algorithm 21: Proposed SA-CS-based Gain Optimisation

Objective function, $f(X) = X = (X_1, X_2 \ldots X_D)^T$
Produce initial population of n host nests $Xi(i = 1, 2 \ldots .n)$
While $(t < Max\ Generation)$or terminate the process
arbitrarily identify a cuckoo using the Lévy distribution;
Evaluate its quality/fitness; Q_i
pick a nest among n (assume j) arbitrarily;
Compute its quality/fitness; Q_j
Find βusing Equation 15
Determine τusing Equation 16
If $(Q_i > Q_j)$
Change j by novel solution;
End
A fraction of (P) of worse nests are left and novel nest are constructed at novel locations;
keep the finest solutions in place;
Sort answers to find the best (current);
End while
Post-processing

5. Suggested CIRCULATING current controller

The suggested current controller is made up of a number of blocks that are built using the H-infinity controller, whose output is represented by U_{ref}. The control method operator's numerical solution is provided below. Figure 3 shows the proposed current controller design.

$$H_1 = (I_d - I_{dref})\left[\left([BK_2[(I - AK_1)[CK_3]]]^T + DK_4\right).e\left(\frac{u}{||u^2||}\right).\frac{u}{||u^2||} = U\right] \tag{22}$$

$$Q = V_d + R_{ff}I_{d_{ref}} + (I_{q_{ref}}.L_{ff}.w) \tag{23}$$

$$H_2 = (I_q - I_{q_{ref}})\left[\left([BK_2[(I - AK_1)[CK_3]]]^T + DK_4\right).e\left(\frac{u}{||u^2||}\right).\frac{u}{||u^2||} = U\right] \tag{24}$$

$$R = V_q + R_{ff}I_{q_{ref}} + (I_{d_{ref}}.L_{ff}.w) \tag{25}$$

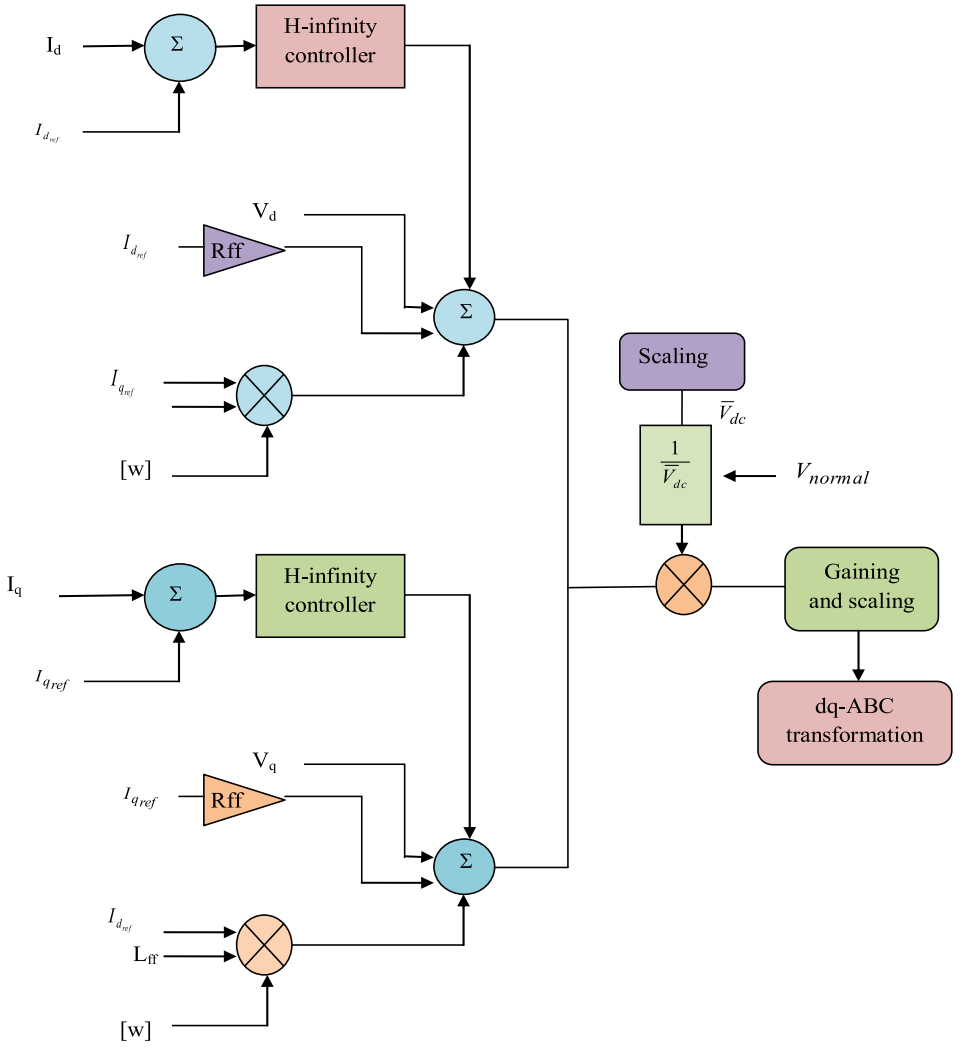


Figure 3. Proposed current controller design.

$$M = H_1 + Q \tag{26}$$

$$N = H_2 + R \tag{27}$$

$$U_{ref} = \begin{bmatrix} M \\ N \end{bmatrix} \left(\frac{1}{V_{dc}} \right) \tag{28}$$

As a result, Equation (28) generates the current operator's signal.

Table 2. Simulation parameters.

Parameters	Values
Fnom, nominal system frequency	50
Pnom, Converter 3-phase rated power	1000e6
Vnom-prim, nominal primary voltage	400e3
Vnom-sec, nominal secondary voltage	333e3
Nb-PM, number of power module per arm	36
Vnom-dc, DC nominal volage	640e3
C-PM, power module capacitor	1.758e-3
VC0-PM	0
Tbrk 1, closing time of breaker 1	0.1
Tbrk 2, closing time of breaker 2	1.0
Tdcblock, converter dc-block time	1.5
Tramping vdc ref,start time vdc-ref ramping to nominal	2
Slope-vdc-ref, reference damping	dc/5
Ton-PQ reg, regulators turn on time	4
Tramping pref, start time pref ramping	Ton-PQreg+0.2
Slope-Qref, sloge pref ramping	0.5
Tramping Qref, start time pref ramping	Ton-PQreg+3.5
Slope-Qref, sloge pref ramping	0.5
Ton-converter 2	4
p	0
Select pulses	pp
Select pulse	pp+36
Control, control system tinme step	40e-6
Lxfo,total leakage inductance	0.12
Rxfo,total winding resistance	0.003
Fc carriers frequency	Fnom*3.37
Fn-filter	1000
Zeta-filter	1
Kp-preg,Proportional gain	0.5/3
Ki-preg, integral	1.0
Psc, short circuit power	Pnom*20
X-R; X/R ratio	7
p-ld1, load	Psc/30
R-startup, startup resistance	400

5. Results and discussions

5.1. Simulation procedure

In MATLAB, the suggested SA-CS-based was simulated, and the desired results were obtained. A standard HVDC strategy was downloaded from the URL https://in.mathworks.com/matlabcentral/fileexchange/54506-simpowersystems-model-of-a-1000-mw-hvdc-mmc-interconnection?s_tid=FX_rc2_behav. Moreover, the proposed SA-CS was

Table 3. Analysis of circulating current for proposed and conventional models.

Methods	Circulating current (kA)
PI controller	5.2377×10^5
H-infinity controller (Tavakoli et al., 2021)	5.2969×10^5
H-infinity-GA (Vrionis et al., 2014)	5.2627×10^5
H-infinity-FF (Fister et al., 2013)	5.3575×10^5
H-infinity-PSO (Zhang & Xia, 2017)	5.4535×10^5
H-infinity-CS (Mareli & Twala, 2017)	5.3227×10^5
H-infinity -NSGA (Ji et al., 2021)	5.2326×10^5
H-infinity-HHO (Diab et al., 2020)	5.4426×10^5
H-infinity-SA-CS	5.1546×10^5

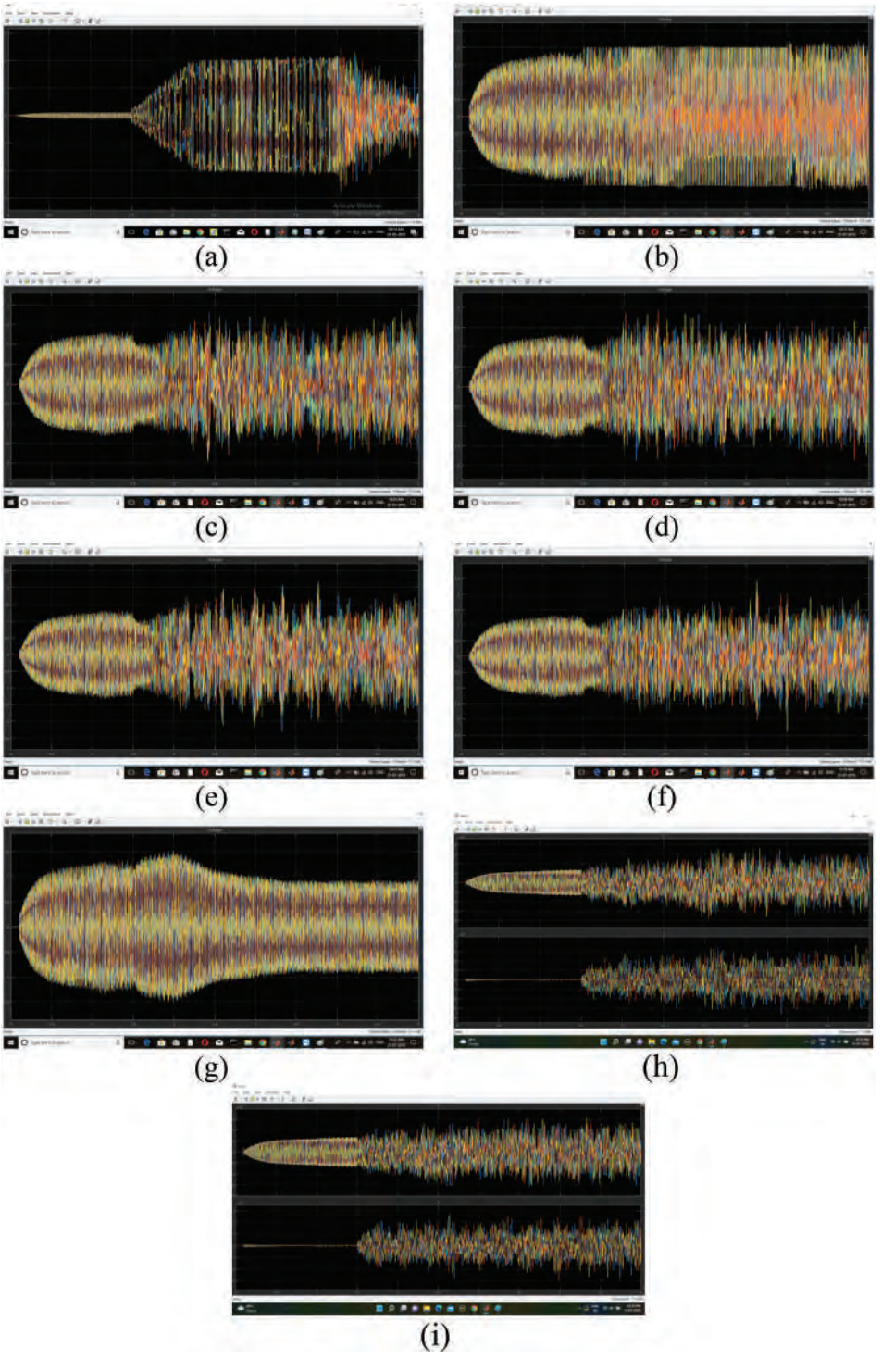


Figure 4. Control analysis of the suggested model (a) PI controller (b) H-infinity controller (c) H-infinity-GA (d) H-infinity-FF (e) H-infinity-PSO (f) H-infinity-CS (g) H-infinity-SA-CS(h) H-infinity-NSGA-III algorithm and (i) H-infinity-HHO optimization.

Table 4. Analysis of THD for proposed method with other conventional methods.

THD analysis			
Methods	1st phase	2nd phase	3rd phase
PI controller (Sreedhar & Mishra, 2014)	−34.652	−35.09	−34.216
H-infinity controller (Tavakoli et al., 2021)	−29.526	−32.407	−29.67
H-infinity-GA (Vrionis et al., 2014)	−24.458	−26.438	−26.611
H-infinity-FF (Fister et al., 2013)	−30.815	−31.899	−34.026
H-infinity-PSO (Zhang & Xia, 2017)	−36.666	−35.628	−33.694
H-infinity-CS (Mareli & Twala, 2017)	−35.343	−32.575	−29.985
H-infinity -NSGA (Ji et al., 2021)	−32.112	−31.321	−36.323
H-infinity-HHO (Diab et al., 2020)	−34.411	−32.112	−33.211
H-infinity-SA-CS	−41.867	−41.534	−41.133

analysed in terms of control analysis, converter analysis and capacitor power voltage. Accordingly, the suggested SA-CS approach was compared with PI (Proportional – Integral) controller, H-infinity controller, H-infinity-GA (Genetic Algorithm) (Vrionis et al., 2014), H-infinity-FF(Firefly Algorithm) (Fister et al., 2013), H-infinity-PSO (Particle Swarm Optimisation) (Zhang & Xia, 2017), H-infinity-CS (Mareli & Twala, 2017) H-infinity-NSGA-III Algorithm (Ji et al., 2021) and H-infinity-HHO (Diab et al., 2020) optimisation. When compared to other models, the simulation results show that the SA-CS technique can obtain superior outputs. Table 2 depicts the summary of simulation variables.

5.2. Control analysis

Figure 4 illustrates the control evaluation of both the proposed and traditional models. The current regulator pulses is used as the signal. The goal is to regulate the voltages and currents in the arms, although an misconstrued linear MMC method is employed. Additionally, only linearisation mistakes are studied in terms of the controller's ability to reject disturbances; no actual disturbances (such as variations in power or voltage) are taken into account. The proposed H-infinity controller minimises the overshoot and error. The attained control signal is clearly illustrated in Figure 4(g), which shows better control over other models.

5.3. Converter performance

The converter performance in terms of voltage as well as current is displayed in Figures 5 and 6. The voltage used here is converter DC voltage. The proposed model concerning DC voltage is given in Figure 5(g), which shows that the suggested strategy performs better than other methods. Similarly, converter performance of presented model in terms of DC current is given in Figure 6(g). The converter DC current is used in Figure 6.

5.4. Capacitor voltage

Figure 7 presents the capacitor voltage assessment for the suggested and traditional methods. The capacitor voltage used here is MMC output. The resultant signal of the proposed H-infinity based on SA-CS is given in Figure 7(g).

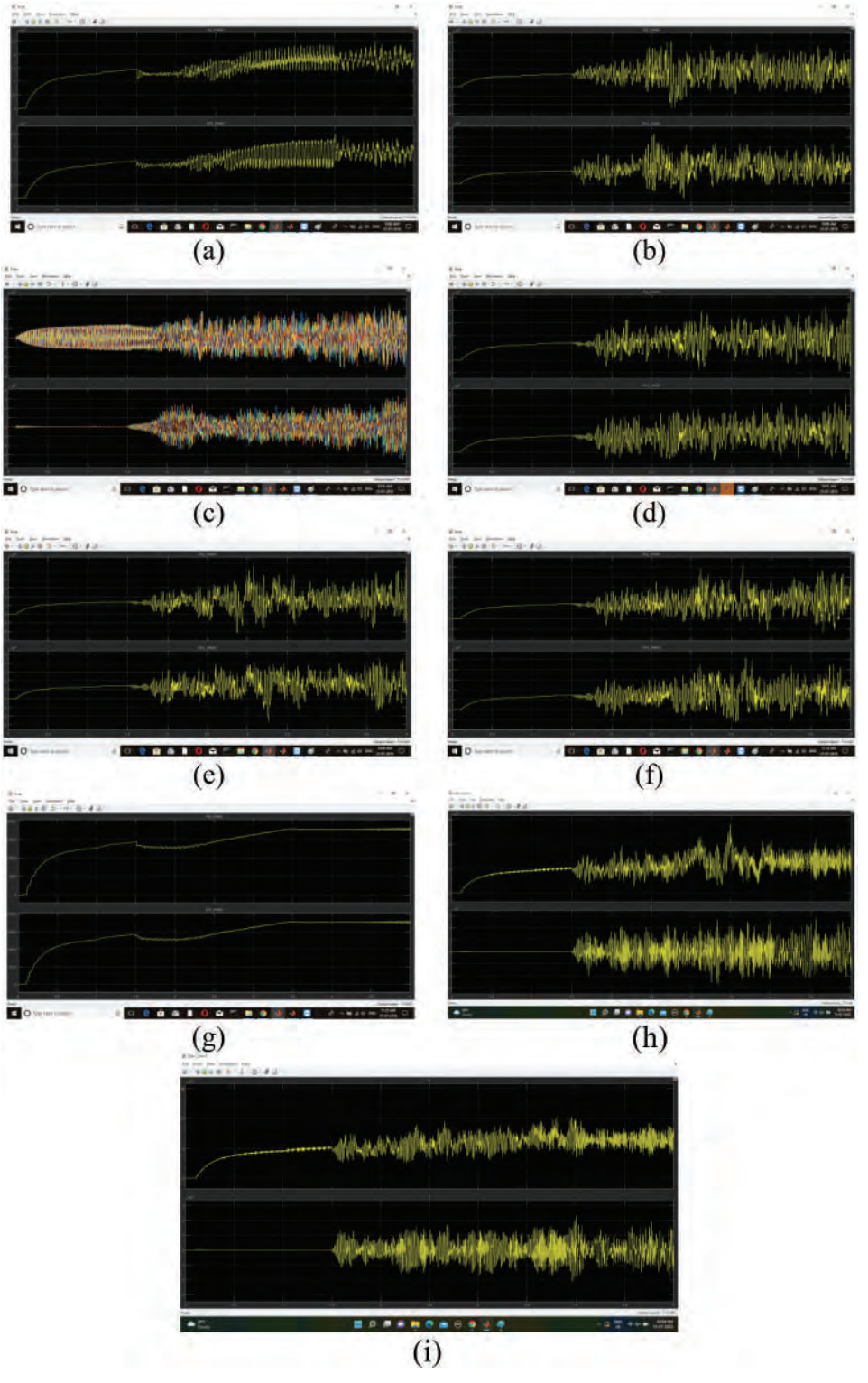


Figure 5. Converter performance analysis of suggested strategy in terms of voltage for (a) PI controller (b) H-infinity controller (c) H-infinity-GA (d) H-infinity-FF (e) H-infinity-PSO (f) H-infinity-CS (g) H-infinity-SA-CS (h) H-infinity- NSGA-III algorithm and (i) H-infinity-HHO optimization.

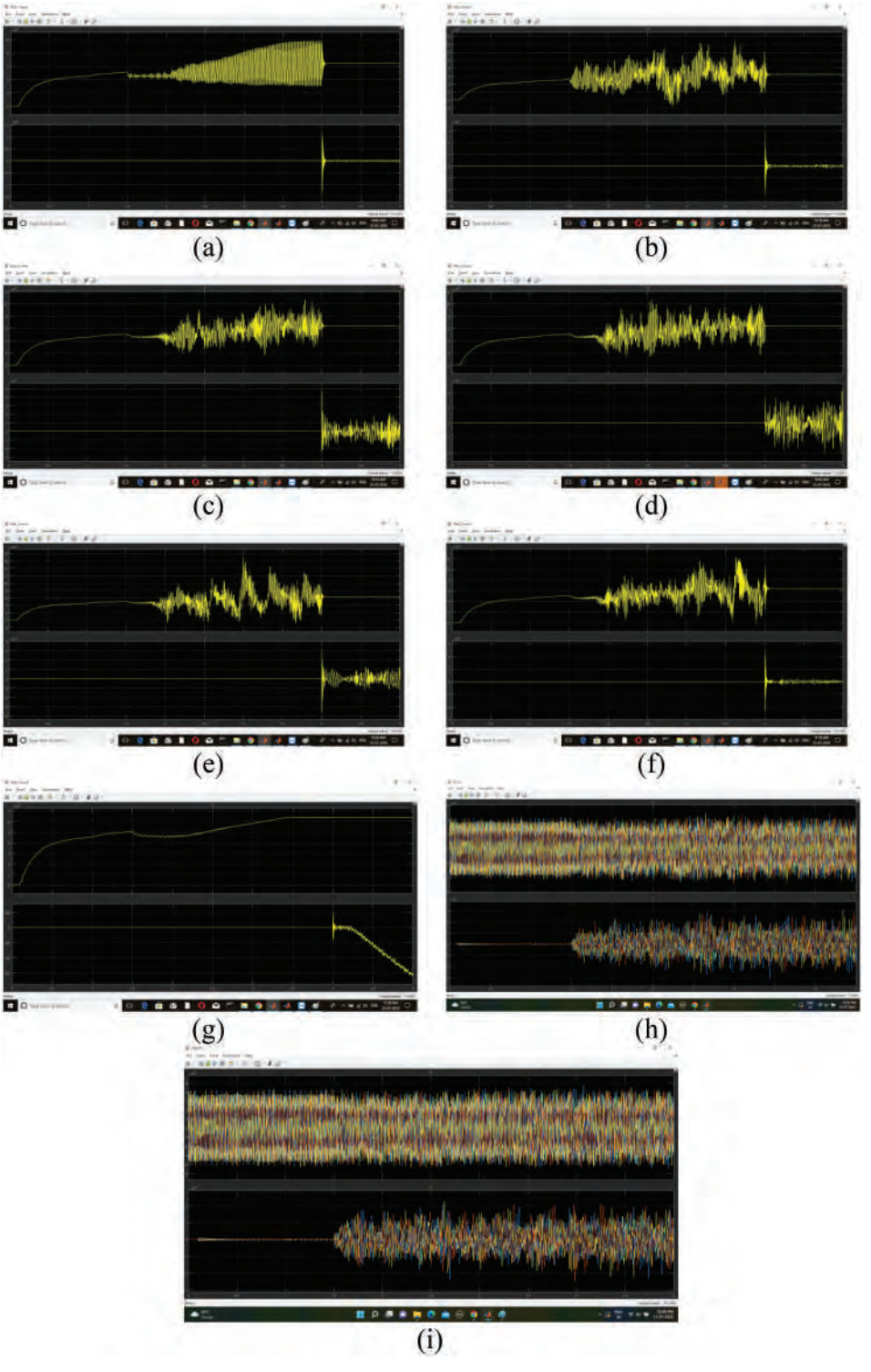


Figure 6. Converter performance analysis of suggested model in terms of current for (a) PI controller (b) H-infinity controller (c) H-infinity-GA (d) H-infinity-FF (e) H-infinity-PSO (f) H-infinity-CS (g) H-infinity-SA-CS (h) H-infinity- NSGA-III algorithm and (i) H-infinity-HHO optimization.

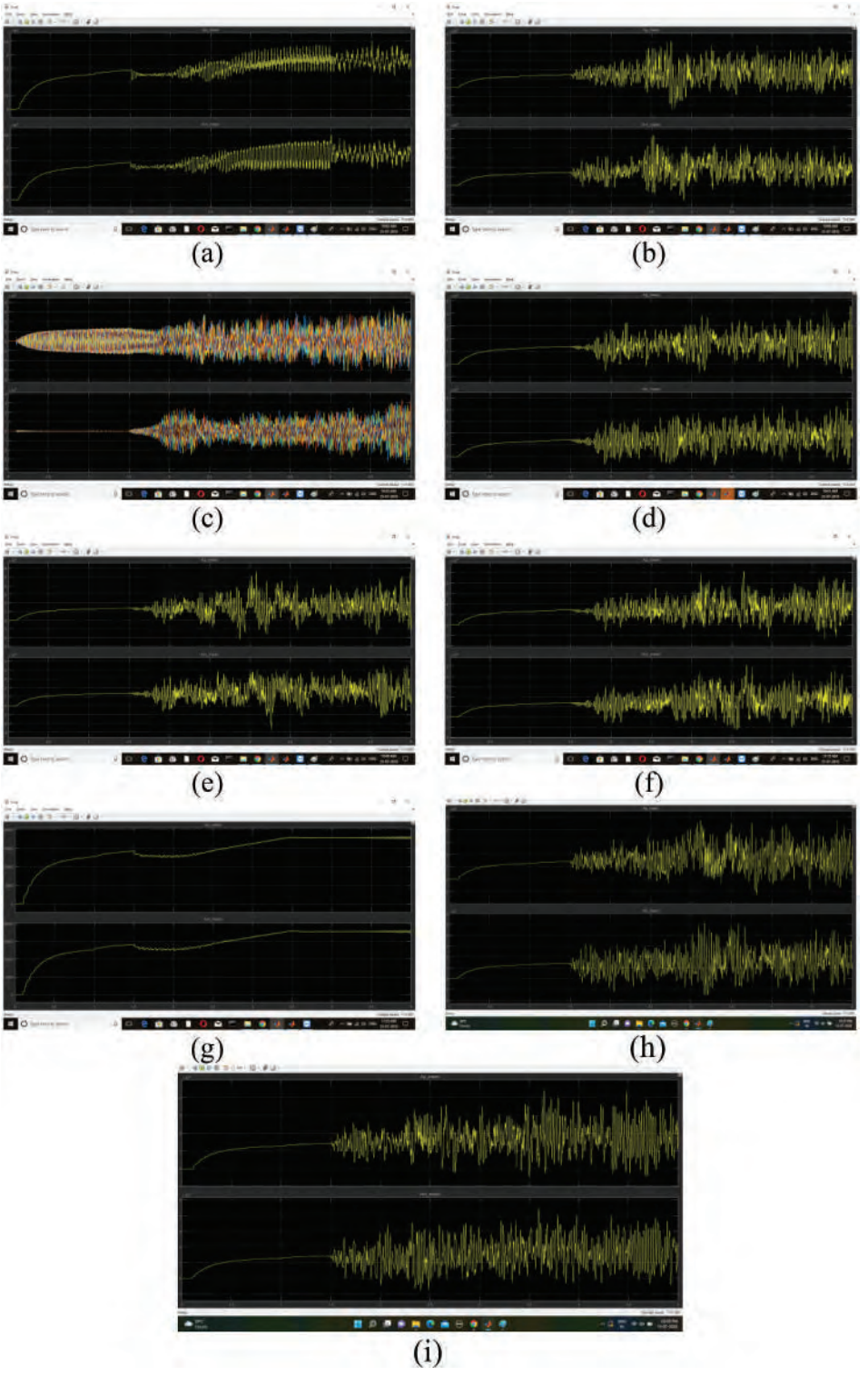


Figure 7. Capacitor voltage assessment of suggested model (a) PI controller (b) H-infinity controller (c) H-infinity-GA (d) H-infinity-FF (e) H-infinity-PSO (f) H-infinity-CS (g) H-infinity-SA-CS (h) H-infinity-NSGA-III algorithm and (i) H-infinity-HHO optimization.

5.5. Analysis of circulating current

Table 3 compares the assessment of overcurrent for the suggested protocol to other traditional methods. The RMS current passing through the power devices would increase due to the current flowing in the MMC, which would lead to an increase in converter power loss. Because of its low frequency, arm inductors can be employed to suppress the circulating current, although a very large inductance is required. From the analysis, the presented H-infinity-SA-CS is 1.61% better than PI-controller, 2.76% better than the H-infinity controller, 2.09% better than H-infinity-GA, 3.93% better than H-infinity-FF algorithms, 5.79% H-infinity – PSO, 3.26% better than H-infinity CS, 1.51% better than NSGA and 5.58% better than H-infinity-HHO. The circulating current is unavoidable, and its unit is measured in (kA). From this, comparing the suggested model to existing models, it is clear that it achieves less circulating current.

5.6. Analysis of total harmonic distortion

Table 4 compares a THD analysis of the proposed scheme to other established techniques. From the table, for the 1st phase, the adopted H-infinity-SA-CS is 20.82% better than the PI controller, 41.8% better than the H-infinity controller, 71.18% better than H-infinity-GA, 35.87% better than H-infinity-FF, 14.18% better than H-infinity-PSO, 18.46% better than H-infinity-CS, 30.37% better than H-infinity-NGSA and 21.66% better than H-infinity HHO algorithms. Similarly, the adopted approach for 2nd phase is 18.36% superior to the PI controller, 28.16% superior to the H-infinity controller, 57.09% superior to H-infinity-GA, 30.2% superior to H-infinity-FF, 16.58% superior to H-infinity-PSO, 27.5% superior to H-infinity-CS, 32.70% better than H-infinity-NGSA and 29.34% better than H-infinity HHO algorithms. Thus from the analysis, comparing the proposed approach to other models, it is clear that it achieves fewer THD.

5.7 Analysis in terms of fitness

Table 5 depicts the fitness function. The fitness of this method is higher when compared to the other traditional method like H-infinity GA, H-infinity FF, H-infinity PSO, H-infinity CS, H-infinity NSGA-III and H-infinity HHO. The resultant outcomes, it is demonstrated that the new schemes is better to the other old method.

Table 5. Fitness analysis.

Methods	Fitness
H-infinity-GA (Vrionis et al., 2014)	−62.525
H-infinity-FF (Fister et al., 2013)	−61.255
H-infinity-PSO (Zhang & Xia, 2017)	−61.157
H-infinity-CS (Mareli & Twala, 2017)	−61.164
H-infinity NSGA-III (Ji et al., 2021)	−61.134
H-infinity HHO (Diab et al., 2020)	−61.204
H-infinity-SA-CS	−61.351

Table 6. Analysis in terms of computational time.

Methods	Fitness
H-infinity-GA (Vrionis et al., 2014)	19.574
H-infinity-FF (Fister et al., 2013)	19.416
H-infinity-PSO (Zhang & Xia, 2017)	19.222
H-infinity-CS (Mareli & Twala, 2017)	19.393
H-infinity NSGA-III (Ji et al., 2021)	19.121
H-infinity HHO (Diab et al., 2020)	19.026
H-infinity-SA-CS	19.045

5.8 Analysis in terms of computational time

The computational analysis of a suggested methodology compared to the traditional method is shown in Table 6. The suggested technique outperforms other conventional methods as H-infinity-GA, H-infinity-FF, H-infinity-PSO, H-infinity-CS, H-infinity NSG III and H-infinity-HHO in terms of computing time with 2.77%, 1.94%, 0.92%, 1.81%, 0.39% and 0.09%, respectively. As a result, the suggested scheme's superiority over other options is established.

6. Conclusion

This paper has developed an MMC for HVDC using optimal H-infinity controller for PV systems that are applicable for three-phase systems. Firstly, the sensor measured the unstable currents in each phase and compared them to the reference signal, which was set to zero. The gain was optimised using SA-CS to control the signal in H-infinity controller. At last, the performance analysis was formed by differentiating suggested methodology over other conventional models in terms of analysis of systems, conversion efficiency and voltage gain, circulating current and THD. From analysis, a circulating current of the presented H-infinity-SA-CS was 6.64% better than the H-infinity controller, 0.16% better than H-infinity-GA and 5.62% better than H-infinity-FF algorithms. Similarly, the THD of presented H-infinity-SA-CS was 20.82% better than PI controller, 41.8% better than H-infinity controller, 71.18% better than H-infinity-GA, 35.87% better than H-infinity-FF, 14.18% better than H-infinity-PSO and 18.46% better than H-infinity-CS strategies. Advantages of suggested methodology are its simple design and the circulating current is also reduced. The limitation of this study includes the imbalance is occurred due to the peak arm current. Future research would be fascinating if it included a thorough investigation of MMC performance factors like lower circulation power, balanced power, little voltage deviation, control of frequency, common-mode voltage regulation and avoidance of narrow pulses are all desirable.

Disclosure statement

No potential conflict of interest was reported by the author.

Nomenclature

Acronyms	**Description**
MMC	Modular Multilevel Converter

THD	Total Harmonic Distortion
SA-CS	Self-Adaptive Cuckoo search algorithm
SMs	sub-modules
HVDC	High-Voltage Direct Current
PWM	Pulse Width Modulation
LCCs	Line Commutated Converters
VSCs	Source Converters
PCC	Predictive current control
MPC	Model Predictive Control
H-infinity-FF	H-infinity with Firefly
RCS	Repetitive Control System
H-infinity-PSO	H-infinity with Particle Swarm Optimisation
H-infinity-CS	H-infinity with Cuckoo Search
HHO	Harris Hawks Optimisation
CPD PWM	Carrier Phase Disposition PWM Technique
DC	Digital Controllers
RES	Renewable Energy Source
GA	Genetic Algorithm
PCT	Passivity Control Theory

References

Almula Gebreel, A., & Longya, X. (2016, April). Power quality and total harmonic distortion response for MMC with increasing arm inductance based on closed loop-needless PID controller. *Electric Power Systems Research*, *133*, 281–291. https://doi.org/10.1016/j.epsr.2015.12.024

Babikir Bashir, S., & Beig, A. R. (2018, February). An improved voltage balancing algorithm for grid connected MMC for medium voltage energy conversion. *International Journal of Electrical Power & Energy Systems*, *95*, 550–560. https://doi.org/10.1016/j.ijepes.2017.09.002

Beddard, A., & Barnes, M. (2015). Modelling of MMC-HVDC systems – an overview. *Energy Procedia*, *80*, 201–212. https://doi.org/10.1016/j.egypro.2015.11.423

Božiček, A., Papič, I., & Blaži, B. (2017, October). Performance evaluation of the DSP-based improved time-optimal current controller for STATCOM. *International Journal of Electrical Power & Energy Systems*, *91*, 209–221. https://doi.org/10.1016/j.ijepes.2017.03.015

Cretnik, J., Petrovčič, J., & Bitenc, A. (1992, March). A design and implementation of a combustion controller in microcomputer-based multiloop controller MMC-90. *IFAC Proceedings Volumes*, *25* (1), 123–128. https://doi.org/10.1016/S1474-6670(17)50440-1

Diab, A. A. Z., Ebraheem, T., Aljendy, R., Sultan, H. M., & Ali, Z. M. (2020). Optimal design and control of MMC STATCOM for improving power quality indicators. *Applied Sciences*, *10*(7), 2490. https://doi.org/10.3390/app10072490

Elserougi, A. A., Massoud, A. M., & Ahmed, S. (2017, May). A transformerless STATCOM based on a hybrid boost modular multilevel converter with reduced number of switches. *Electric Power Systems Research*, *146*, 341–348. https://doi.org/10.1016/j.epsr.2017.02.014

Fister, I., Fister, I., Yang, X.S., & Brest, J. (2013, December). A comprehensive review of firefly algorithms. *Swarm and Evolutionary Computation*, *13*, 34–46. https://doi.org/10.1016/j.swevo.2013.06.001

Guoqing, L., Jiang, S., Xin, Y., Wang, Z., & Xue, L. (2018, July). An improved DIM interface algorithm for the MMC-HVDC power hardware-in-the-loop simulation system. *International Journal of Electrical Power & Energy Systems*, *99*, 69–78. https://doi.org/10.1016/j.ijepes.2017.12.013

Guo, C., Yin, Z., Zhao, C., & Iravani, R. (2018, June). Small-signal dynamics of hybrid LCC-VSC HVDC systems. *International Journal of Electrical Power & Energy Systems*, *98*, 362–372. https://doi.org/10.1016/j.ijepes.2017.12.010

Haifeng, L., Liu, C., Gengyin, L., Guo, Q., & Luo, H. (2015, July). A start strategy for synchronized connection of MMCs to an AC system. *International Journal of Electrical Power & Energy Systems*, *69*, 380–390. https://doi.org/10.1016/j.ijepes.2015.01.016

Jianzhong, X., Ding, H., Fan, S., Gole, A. M., & Zhao, C. (2016, December). Enhanced high-speed electromagnetic transient simulation of MMC-MTdc grid. *International Journal of Electrical Power & Energy Systems*, *83*, 7–14. https://doi.org/10.1016/j.ijepes.2016.03.062

Ji, K., Chen, W., Wu, X., Pang, H., Hu, J., Liu, S., Cheng, F., & Tang, G. (2021). High frequency stability constraints based MMC controller design applying NSGA-III algorithm. *CSEE Journal of Power and Energy Systems*. 1–9. https://doi.org/10.17775/CSEEJPES.2020.03810

Jinke, L., Konstantinou, G., Wickramasinghe, H. R., Pou, J., & Jin, X. (2017, November). Investigation of MMC-HVDC operating region by circulating current control under grid imbalances. *Electric Power Systems Research*, *152*, 211–222. https://doi.org/10.1016/j.epsr.2017.07.003

Liu, C., Lin, X., Haifeng, L., Cheng, X., & Gengyin, L. (2014, March). Sub-module component developed in CBuilder for MMC control and protection test in RTDS. *International Journal of Electrical Power & Energy Systems*, *56*, 198–208. https://doi.org/10.1016/j.ijepes.2013.11.010

Madichetty, S., Dasgupta, A., Mishra, S., Panigrahi, C. K., & Basha, G. (2016). Application of an advanced repetitive controller to mitigate harmonics in MMC with APOD scheme. *IEEE Transactions on Power Electronics*, *31*(9), 6112–6121. https://doi.org/10.1109/TPEL.2015.2501314

Mareli, M., & Twala, B. (6 September. 2017). An adaptive Cuckoo search algorithm for optimisation. *Applied Computing and Informatics*, *14*(2), 107–115. https://doi.org/10.1016/j.aci.2017.09.001

Martinez-Rodrigo, F., & Santiago de Pablo, L.C.H.D.L. (2015, November). Current control of a modular multilevel converter for HVDC applications. *Renewable Energy*, *83*, 318–331. https://doi.org/10.1016/j.renene.2015.04.037

Mehrasa, M., Pouresmaeil, E., Zabihi, S., Vechiu, I., & Catalão, J. P. S. (2018, March). A multi-loop control technique for the stable operation of modular multilevel converters in HVDC transmission systems. *International Journal of Electrical Power & Energy Systems*, *96*, 194–207. https://doi.org/10.1016/j.ijepes.2017.10.006

Mohammad Hassani, A., Bektas, S. I., & Hossein Hosseini, S. (2017). Modular multilevel converter circulating current control using model predictive control combined with genetic algorithm. *Procedia computer science*, *120*, 780–787. https://doi.org/10.1016/j.procs.2017.11.308

Mohana, S., Sahaaya, S. A., & Mary, A. (2016). A comparitive framework for feature selction in privacy preserving data mining techniques using pso and k-anonumization. *Iioab Journal*, *7*(9), 804–811.

Prieto-Araujo, E., Junyent-Ferré, A., Collados-Rodríguez, C., Clariana-Colet, G., & Gomis-Bellmunt, O. (2017, November). Control design of modular multilevel converters in normal and AC fault conditions for HVDC grids. *Electric Power Systems Research*, *152*, 424–437. https://doi.org/10.1016/j.epsr.2017.06.020

Ramirez, D., Martinez-Rodrigo, F., de Pablo, S., & Carlos Herrero de Lucas, L. (2017, February). Assessment of a non linear current control technique applied to MMC-HVDC during grid disturbances. *Renewable Energy*, *101*, 945–963. https://doi.org/10.1016/j.renene.2016.09.050

Raymundo, E.T.O., & Garces, A. (2017, October). Grid integration of offshore wind farms using a hybrid HVDC composed by an MMC with an LCC-based transmission system. *Energy Procedia*, *137*, 391–400. https://doi.org/10.1016/j.egypro.2017.10.363

Rui, L., & Fletcher, J. E. (2017, February). A novel MMC control scheme to increase the DC voltage in HVDC transmission systems. *Electric Power Systems Research*, *143*, 544–553. https://doi.org/10.1016/j.epsr.2016.11.004

Satyanarayana, G., & Lakshmi Ganesh, K. (2015). Incorporate of FB-MMCs converter topologies for hybrid PV/FC based EV applications. *Procedia Technology*, *21*, 271–278. https://doi.org/10.1016/j.protcy.2015.10.028

Shi, K., Yuan, X., & Liu, L. (2018, January). Model predictive controller-based multi-model control system for longitudinal stability of distributed drive electric vehicle. *ISA transactions*, *72*, 44–55. https://doi.org/10.1016/j.isatra.2017.10.013

Sing Lee, S., & En Heng, Y. (2018, January). A tuning-less model predictive control for modular multilevel converter capable of unbalanced grid fault. *International Journal of Electrical Power & Energy Systems*, *94*, 213–224. https://doi.org/10.1016/j.ijepes.2017.07.011

Sreedhar, M., & Mishra, S. (2014, July). Voltage balancing scheme in MMC – a new approach. *International Journal of Emerging Electric Power Systems*, *15*(4), 2194–5756. https://doi.org/10.1515/ijeeps-2014-0096

Tavakoli, S. D., Fekriasl, S., Prieto-Araujo, E., Beerten, J., & Gomis-Bellmunt, O. (2021). Optimal $ \mathcal {H} _ {\infty} $ control design for MMC-based HVDC links. *IEEE Transactions on Power Delivery*, *37*(2), 786–797.

Venkat Ram Reddy, K., Gowri Manohar, T., & Sreedhar, M. (2017, February). Circulating current mitigating scheme in MMC based HVDC system with H∞ repetitive controllers. *International Journal of Electrical Power & Energy Systems*, *85*, 143–152. https://doi.org/10.1016/j.ijepes.2016.08.007

Vidal-Albalate, R., Belenguer, E., Beltran, H., & Blasco-Gimenez, R. (2016, December). Efficient model for modular multi-level converter simulation. *Mathematics and Computers in Simulation*, *130*, 167–180. https://doi.org/10.1016/j.matcom.2015.10.001

Vrionis, T. D., Koutiva, X. I., & Vovos, N. A. (2014, May). A genetic algorithm-based low voltage ride-through control strategy for grid connected doubly fed induction wind generators. *IEEE Transactions on Power Systems*, *29*(3), 1325–1334. https://doi.org/10.1109/TPWRS.2013.2290622

Wang, Y., Yuan, Z., Jiao, F., Yan, L., & Zhao, Y. (2017, September). A feasible coordination protection strategy for MMC-MTDC systems under DC faults. *International Journal of Electrical Power & Energy Systems*, *90*, 103–111. https://doi.org/10.1016/j.ijepes.2017.02.005

Xia, X., Zhou, Y., Chunhui, F., Zhou, Z., & Yunjiu, H. (2016, November). Research on high voltage DC transmission system optimal control based on MMC. *International Journal of Electrical Power & Energy Systems*, *82*, 207–212. https://doi.org/10.1016/j.ijepes.2016.02.050

Yang, X. -S. (2014). *Nature -inspired optimization algorithms* (First ed.). Elsevier.

Zhang, J., & Xia, P. (17 February. 2017). An improved PSO algorithm for parameter identification of nonlinear dynamic hysteretic models. *Journal of Sound and Vibration*, *389*, 153–167. https://doi.org/10.1016/j.jsv.2016.11.006

Zhang, J., & Zhao, C. (2016, November). Analysis and control of MMC-HVDC under unbalanced voltage conditions. *Electric Power Systems Research*, *140*, 528–538. https://doi.org/10.1016/j.epsr.2016.05.021

Measurement: Sensors 27 (2023) 100751

Contents lists available at ScienceDirect

Measurement: Sensors

journal homepage: www.sciencedirect.com/journal/measurement-sensors

Computer vision based healthcare system for identification of diabetes & its types using AI

Avinash Sharma [a], K.D.V. Prasad [b,c], Sadashiva V. Chakrasali [d], Dankan Gowda V [e,*], Chanakya Kumar [f], Abhay Chaturvedi [g], A. Azhagu Jaisudhan Pazhani [h]

[a] Computer Science & Engineering Department, M. M. Engineering College, Maharishi Markandeshwar (Deemed To Be) University, Mullana, Ambala, India
[b] Symbiosis Institute of Business Management, Hyderabad, India
[c] Symbiosis International (Deemed University), Pune, India
[d] Department of Electronics and Communication, M S Ramaiah Institute of Technology, MSR Nagar, MSRIT Post, Bangalore, Karnataka, India
[e] Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India
[f] Faculty of Marketing, Indira School of Business Studies PGDM, Pune, Maharashtra, India
[g] Department of Electronics and Communication Engineering, GLA University, Mathura, Uttar Pradesh, India
[h] Department of Electronics and Communication Engineering, Ramco Institute of Technology, Rajapalayam, Tamilnadu, India

ARTICLE INFO

Keywords:
Diabetes
Computer vision
Healthcare system
Artificial intelligence
Machine leaning
Prediction

ABSTRACT

Diabetes mellitus, often known as diabetes, is an endocrine disorder that has a wide global impact today. Here is a requirement for an effective model that able prognoses diabetes and its types with more accurateness as early. Given the breadth and depth of existing studies, there is a pressing need for accurate and timely illness forecasting in the healthcare sector. Current circumstances need the creation and design of systems that are quicker to respond, more accurate, more durable, and more generalizable. For increasing the accurateness of prediction with best effectiveness innovative Artificial Intelligence and Machine Learning Model is proposed. This model predicts the diabetes class using the symptoms located into the data-set which is having the row as one rule of the system & this rule are need to understand and compile using feature.

1. Introduction

It is one of the most significant problems facing public health on a global scale today. The International Diabetes Federation (IDF) estimates that by 2030, as many as 550 million people would have diabetes, up from an estimated 350 million in 2011. 80% of all diabetes-related deaths happen in third world countries like India. In addition to putting a significant strain on the global healthcare system and the global budget, diabetes mellitus is a major economic burden worldwide [1]. All current medical services and administrations have access to high-quality monitoring and a plethora of information collection technologies, providing very inexpensive ways to capture and store data. Because of the sheer volume of data stored in healthcare databases, specialised tools are needed for data retrieval, data storage and loading, information examination, and operational use of data. Fig. 1 depicts the Healthcare AI and DM Applications.

Data mining is prevalently known as KDD Knowledge Discovery in Databases and reason for this KDD or DM is to extricate valuable data from enormous or tremendous databases or information distribution centres. Mining applications are utilized for business just as logical sides [2]. DM is characterized as the way toward finding already obscure examples, drifts in databases and utilizing that data which will assist with building prescient models. Else, It is characterized as the methodology of information determination and investigation and building models utilizing colossal information stores to reveal beforehand unidentified, obscure examples.

In healthcare domain, mining a data is becoming progressively popular, where the massive volumes of information produced, produced by clinical dealings are much multifaceted as well as large for managing and analyzing using old approaches. Data _mining can discover designs, pattern as well as tendencies in enormous quantities of multifaceted dataset which improve decision-making [3]. The healthcare industry will take lots of aids, benefits from data mining applications. Data Mining is shown in Fig. 2 as a stage of the knowledge discovery process.

* Corresponding author.
E-mail addresses: asharma@mmumullana.org (A. Sharma), Kdv.prasad@sibmhyd.edu.in (K.D.V. Prasad), sadashiva.c@msrit.edu (S.V. Chakrasali), dankan.v@bmsit.in (D. Gowda V), chanakya.kumar@indiraisbs.ac.in (C. Kumar), abhaychat@gmail.com (A. Chaturvedi), alagujaisudhan@gmail.com (A.A.J. Pazhani).

https://doi.org/10.1016/j.measen.2023.100751
Received 23 January 2023; Accepted 2 February 2023
Available online 22 March 2023

Fig. 1. Healthcare AI and DM applications.

Following is the study which researcher has continued to explore the data science impact on healthcare, its spread, and characteristics [4]. The focus of this research is an analysis of data science application in the field of healthcare application. Before going for Data Science one need to understand the concept of Data Mining. This Data Mining denotes to mining either extracting knowledge or information as of enormous quantities of records. Data Mining stands for 'knowledge mining from data', or 'mining of Knowledge" a shorter term. Knowledge mining from records or datasets, knowledge extraction, pattern analysis, data archaeology, and data dredging are other terms carry DM. Data mining as a synonym for Knowledge Discovery from Data (KDD). Otherwise, others vision Data Mining as simply a necessary phase in processing of knowledge discovery. Altered procedures or form of data pre-processing as well as the datasets are made ready for mining. This mining stage might cooperate with knowledge base else with the customer [5]. The applicable designs are accessible to again customer. This deposited as novel information in the knowledge base [6]. Thus, it can be correctly said that data mining is a perfect merger of various technological flavours, which are master in their own specific areas, as shown in Fig. 3.

2. Literature review

The aim of this section is to give an insight towards the requirements of healthcare domain and about suitable choice of solution using available most suitable technique. This research review will help for assessment of the current state of research, and help for identification of key problem about a topic that needs further research. From last few decades, with a beginning of computer technology, many researchers had contributed significantly in the area of data mining [7]. To understand the problem domain in terms of its applicability, variety of techniques, nature of data, an extensive review had been conducted. The primary objective of this effort is to determine how artificial intelligence and machine learning may be used mathematically in the clinical and healthcare sector to develop a model that will aid in making sound choices.

In beginning, the survey was started with a broader vision by including the research papers from all different domains. Later, the survey was narrowed to one application type i.e. AI and ML application healthcare. Within, healthcare mining after surveying the literature for different data sets; finally, this present study was focused on the diabetes mellitus. The present research has a road map as depicted in the Fig. 4. In Ref. [8] author's are planned to structure a framework that would help specialists in clinical determination. They introduced a symptomatic FCM grouping just as SVM utilizing SMO and choose which procedure helps in finding of diabetes. The FCM bunching with a precision of 94.3%. It has a positive prescient worth which is 88.57%. While SVM has a low precision of 59.5%. These outcomes are generally acceptable, in light of the fact that distinguishing the Diabetes is an extremely mind boggling issue.

In [9] author's proposed using support vector machines (SVMs) for the determination of diabetes. Creator likewise utilized an extra clarification module "blackbox model", this transforms the model of a SVM into a comprehensible portrayal of the SVM demonstrative order choice. SVMs give a promising apparatus to the forecast of diabetes with expectation exactness of 94%, affectability of 93%, and particularity of 94%, approved with Real-life diabetes dataset. This proposed work brings about keeping diabetic people from being influenced by coronary illness. 'SVM' was end up being a best grouping strategy with superb prescient execution. Which additionally been researched with the assistance of 'ROC' bend for both preparing and testing information. Henceforth researcher suggested 'SVM' model for the grouping of the diabetic dataset.

In [10] author's proposed system with multilayer characterization utilizing upgraded sacking and advanced weighting. This structure was assessed on five diverse data sets where two diabetes data sets are assessed. The examination of the yield demonstrates that 'HM-BagMoov' gathering structure accomplished the most elevated exactness, affectability and 'F-Measure' when contrasted and singular classifiers [11]. Furthermore, this structure additionally accomplished the most noteworthy exactness contrasted and the cutting edge methods. Proposed model assistance specialists and patients for sickness grouping and

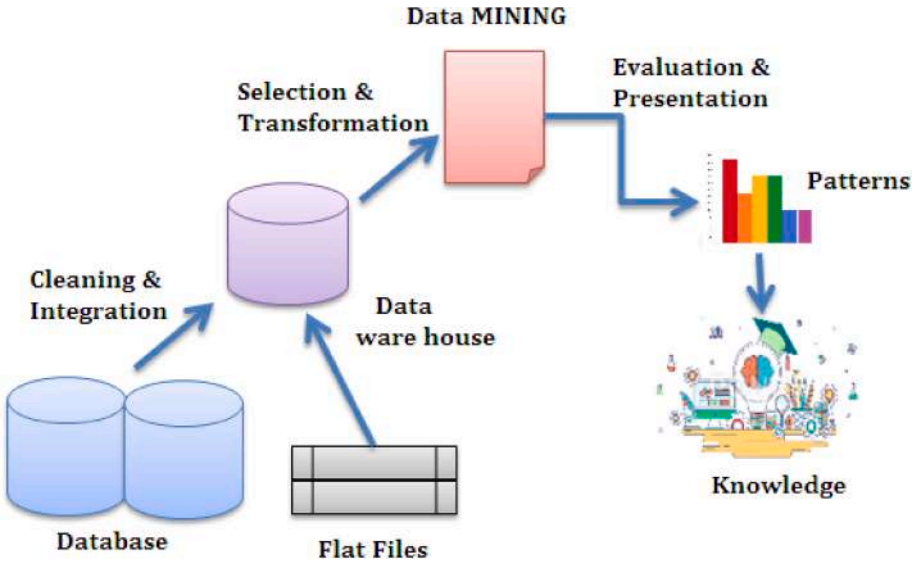


Fig. 2. Knowledge discovery using Data Mining.

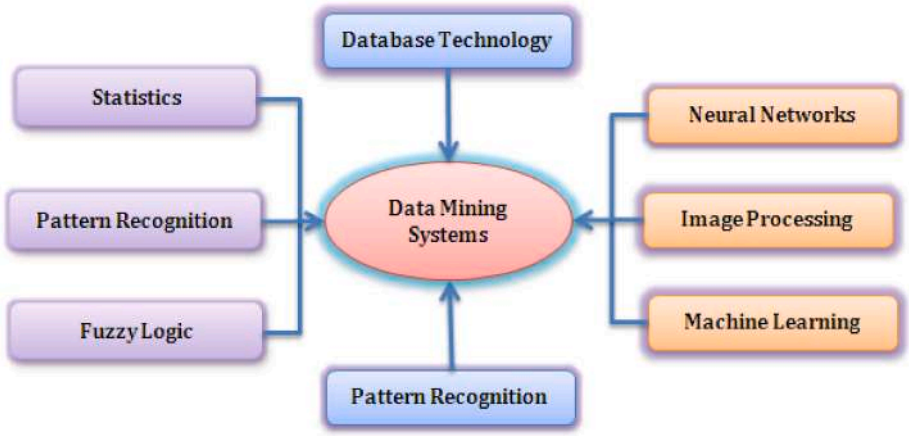


Fig. 3. Different technologies available I data mining systems.

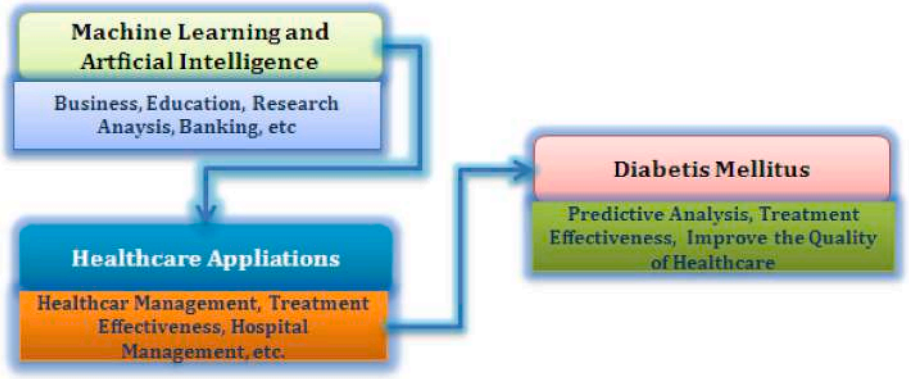


Fig. 4. Methodology of the Proposed work.

forecast dependent on diabetes side effects.

Fig. 5 shows that it comprises of Knowledge Base, Inference Engine and User Interface. These contain heuristic just as verifiable information [12]. Inference also worked as reasoning has types forward chaining and backward chaining in forward chaining it begins with available data and extents a objective and backward chaining begins with objective, exertion backward to data if available allowing for making conclusion of objectives.

3. Healthcare expert systems

The main objective of healthcare expert system is the analysis, prognosis, diagnosis and disease treatment. The healthcare DSS and intelligent system are designed by programs and healthcare knowledge base which generated by computer [13]. Healthcare information essentials to be timely updated. The knowledge for the expert system is gathered into two steps. In first step, knowledge is collected form physician and patient as medical circumstances of diseases are verified through the creation of individual convention whereas, in the second step rules was arrived. In the rule 'if' part have symptoms or termed as

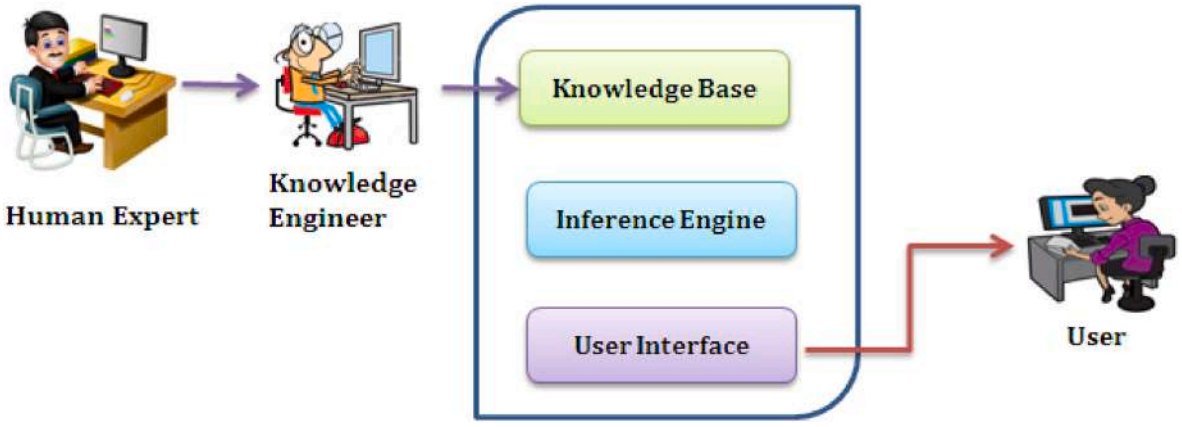


Fig. 5. Components of expert system.

condition and 'then' part with disease which need to be diagnosed. Performance of expert system is depend upon knowledge engineering which done by knowledge engineer. The healthcare expert systems early diagnosis will help effective treatment of patient (Fig. 6.)

In [14] author's made survey on several approaches of Soft Computing such as Support vector machines, Ant Colony Optimization, Neural Network, Fuzzy Logic, Genetic Algorithm, Rough Sets, etc. which diagnosis of diabetes mellitus. The researcher developed web based expert_system for accelerating prognosis of diabetes. The web app's database of information includes "if-then" logic. From the user's name and address to a list of symptoms, the application may take it all. The major benefit of an intelligent system is that it can function in an optimal state indefinitely. Since there may be numerous suitable choices, the optimization approach was not intended to take the position of an expert, but rather to provide additional assurance when making a selection. In Ref. [15] author's designed fuzzy expert system provides a better flexibility for the diagnosis of diabetes. They used PIDD dataset was processed and the crisp-values form dataset transformed for fuzzification the interface. The correlation determination tool evaluates effect of fuzzy operator with rules to make a decision for medical practitioner, the numbers of membership functions for presenting the knowledge-descriptions. Lastly de-fuzzification interface were adopted for converting the fuzzy-output set to the crisp-output. Experimental outcome shows that this tool able to evaluate information well compared to earlier tools.

In [16] author's structured new fuzzy model particularly for type 2 diabetes patients for assessment and decision of ideal remedial technique on singular level. This fuzzy multi criteria tool, decides a number of thought about helpful systems for each treated patient. This model was adaptable as per the chance of change in number, sort of advancement standards change and significance of enhancement criteria change.

In [17] author's structured and executed proposed intelligent system for diet in diabetes. This framework furnishes the patients with clinical suggestions and essential information of diet for diabetic patients. The improvement of the proposed master framework experienced various stages from issue and need recognizable proof, prerequisites investigation, information procurement, formalization, plan and usage [18,19]. For structuring the GUI and the usage of the framework parts visual prolong was utilized. The principle target of the exploration was to build up a clinical master framework for the type_2_diabetes.

Artificial intelligence research is needed not only for diabetes diagnosis, but also for diabetes risk prediction, which would benefit both doctors and patients [20,21]. Developing expert system for diabetes diagnosis is time saving process. Applications of AI, Expert System and ML techniques in the healthcare are hopefully grow in the coming future.

4. Insulin and glucagon: Important roles

Pancreatic endocrine hormones namely glucagon and insulin are the two hormones will help the human body for maintain the blood glucose or sugar. They referred as endocrine hormones because pancreas secreting glucagon and insulin. The Fig. 7. Shows connection among glucagon and insulin [22]. The pancreas glands works as the dominant actor in this system. The cell from the pancreas called islet cell secret Insulin and glucagon hormones within the pancreas. This secretion is in the response to blood sugar levels. The type of islet cell those are the beta cells of the pancreas secretes Insulin hormone. . The Fig. 7 is self-explanatory for the same.

Glucagon hormone stimulates the liver cells to secrete the glucose which is deposited on cells to the blood_streams; the resultant effect is increasing blood glucose level. From Fig. 8 the glucagon also convinces by the liver to create glucose from foundations of energy gained using nutrients_protien consumed human body.

5. Transmission diabetes healthcare through artificial intelligence

Many studies represent bunch of original methods that are purpose to transforming diabetes healthcare into 4 major domains. Artificial intelligence (AI)-powered optical imaging system is a system for diabetic retinopathy detection, computational modeling programmes for mellitus diagnosis, glucose detectors for regulatory oversight, insulin pumps for glucose control, Mobile apps as an essential aid, and other judgement helping mankind in many ways [23,24]. Machine leaning and Artificial intelligence methodologies has the prospective to change diabetes healthcare and help thousands of the diabetes patients for achieving well blood glucose control, for reducing hypoglycaemic occurrences, as well as complication associated with diabetes. Artificial intelligence and machine learning applications proposed superior accurateness, effectiveness, usability, and gratification for patients, their family. Natural language processing (NLP) has capacity of computer systems for analysis, understanding as well as generation of social language, with communication. NLP need to progress through natural language interaction [25]. This agrees performing the tasks by humans being to communicate with computer systems using regular, everyday language. Computer vision (CV) dependent onto deep learning and pattern recognition. This helps for identify the instances present in video/-picture. Computer systems able to process understand and analyze images. In real time with their surroundings this computer vision able to detect pictures or videos. Motivated by the mechanisms of the human brain and nerve cell the neural network (NN) is a type of machine learning approach. This method is collection and interconnection neurons similar units which progression the data by replying to external

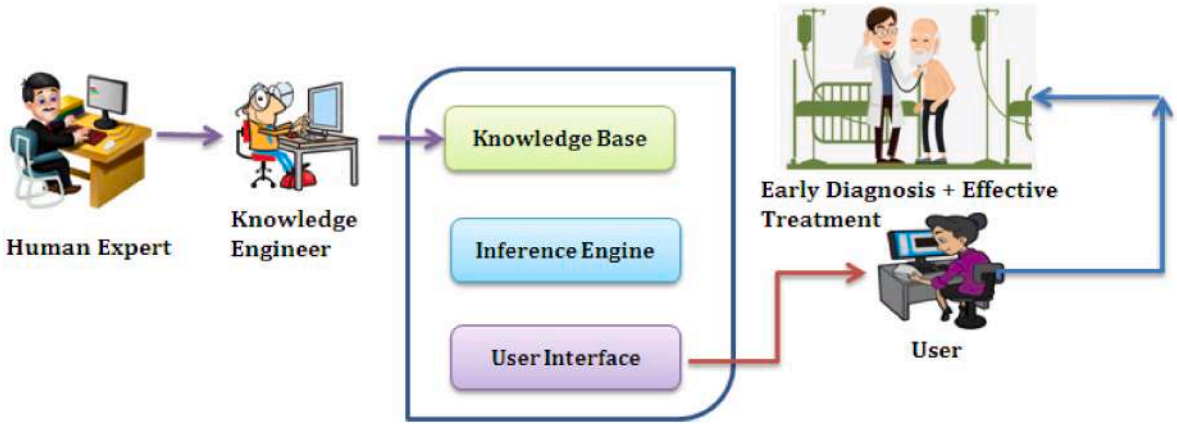


Fig. 6. Expert systems in healthcare.

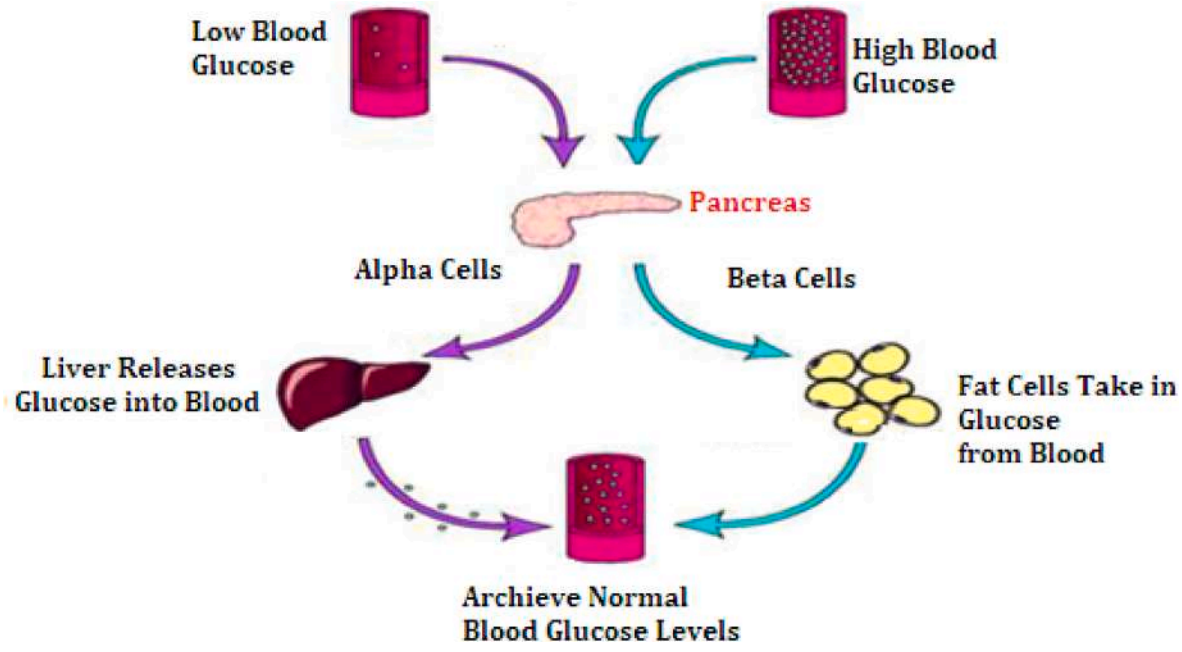


Fig. 7. Pancreas role.

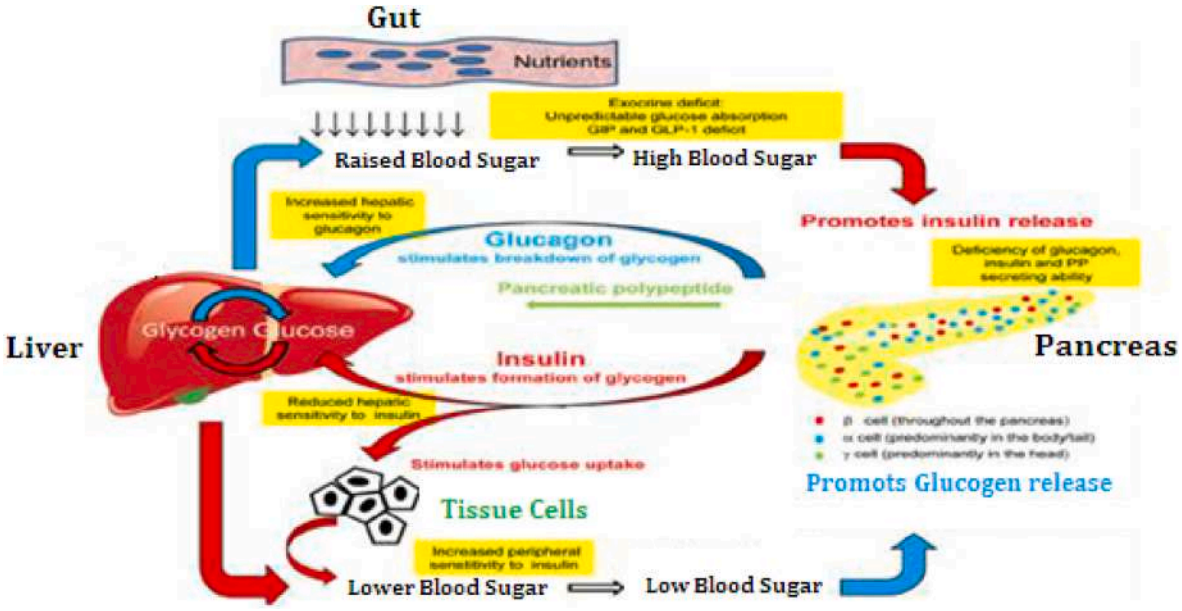


Fig. 8. Function of liver and pancreas.

inputs, communicating information with every node. Deep learning (DL) consumes huge neural networks and many units of processing nodes. This takes benefit superiority in computational power as well as enhanced, better training methods for learning difficult outlines in outsized quantities of records. Common uses comprise speech and image identification. Machine learning is mechanizes methodical model construction and usages approaches like statistics, operations research, physics etc., to discover unknown visions in records deprived of clearly encoded for required production.

6. AI and ML model for prediction

For multiclass classification the use of decision tree classifier is the best and systematic methodology. This positions list makes queries to the dataset. The best visualization of decision tree classification onto a binary tree is possible. At each step of separation in the decision tree the queries are posed as well as then the node is next separated to information with the have dissimilar features. When dataset is separated or split the leave node are making references from the class of belonging. Libraries Dataset Import For proper execution of program code researcher needs respective functional libraries, following code helps us for libraries importing (Fig. 9).

By saving the labels into variable get the data frame from dataset. ScikitLearn has inbuilt function train-test-split ( ) from the library of model selection which randomly separates data set in class train test and test set. The last stage where algorithm able to make predication hence the data must separate into test as well as train mode.

```
import numpy as np
import pandas as pd
import matplotlib.pyplot as plt
import sklearn

df=pd.read_csv("Diadnosis of Diabetes types ML Final.csv")
```

Fig. 9. Importing libraries and dataset.

7. Results and discussion

Researcher defines accuracy in the form of the images of retina with a accurate classification. This final designed network succeeded with 80% of accuracy. With the use of Festal libraries it able to categorize any number of image. The trained convolution neural network makes a rapid prediction a well as immediate reply to user. Following are some of test cases researcher has given. Predict the class of the images for which we created the model. Cloud services offers by Google with virtual machines where graphical processing units prepared for more than nine hundred and sixty tera flops, execution for each event. Despite the dimension of your remarkable job needing to be done, anyone can get faultless GPU given by Ref. [5]. Figs. 10 and 11 shows the Prediction of Categories of normal retina versus diabretina.

8. Test cases

Case 1 In this scenario, the researchers took a picture from the dataset that was intended to be used for testing purposes in order to assess the class of an image. The CNN model provided the proper interpretation, which is that the classification is diabetic retina.

Case 2 Researchers selected one of the images from the dataset that was intended for testing purposes in order to verify the class of the picture. They found that the CNN model had correctly interpreted the category as diabetic retina.

Case 3 In this particular instance, the researchers selected one of the images from the dataset that was intended for testing purposes in order to validate the class of the picture. The CNN model provided the proper interpretation, which is that the class is diabetic retina.

Case 4 Researchers utilized one of the images from the dataset that is used for testing purposes to examine the class of the picture. The CNN model provided the proper interpretation, which is that the classification is diabetic retina.

Case 5 The researchers used one of the images from the dataset that was intended to be used for testing purposes in order to check the classification of the image where the CNN model had provided an incorrect interpretation, namely that the category should have been normal retina but should have been diabetic retina.

Case 6 Researchers used one of the images from the dataset that is used for testing purposes to check the class of the portrait where CNN model gave the correct interpretation, which is that the category is regular retina. This image was used because it was part of the dataset that was used for testing purposes.

Case 7 Researchers used one of the images from the dataset that is used for testing purposes to check the class of the portrait where

Fig. 10. Prediction of Categories case 1–6 (normal retina Vs diabretina).

Fig. 11. Prediction of Categories case 7–10 (normal retina Vs diabretina).

CNN model gave the correct interpretation, which is that the classification is normal retina. This image was used because it was part of the dataset that was used for testing purposes.

Case 8: In this particular instance, the researchers made use of one of the images from the dataset that is intended for testing purposes in order to validate the class of the picture and ensure that the CNN model had correctly interpreted the category as "normal retina."

Case 9 Researchers used one of the images from the dataset that is used for testing purposes to check the class of the portrait where CNN model gave the correct interpretation, which is that the category is regular retina. This image was used because it was part of the dataset that was used for testing purposes.

Case 10 In this particular instance, the researchers made use of one of the images from the dataset that is intended for testing purposes in order to validate the class of the picture and ensure that the CNN model had correctly interpreted the category as "normal retina."

9. Conclusion

Image processing, disease detection, and risk assessment are just some of the areas where AI and machine learning are proving their worth in the healthcare industry. This work proposes a model for detecting diabetic retinopathy that makes use of the Fast AI toolkit and requires much less code than existing methods. Having the neural network already precompiled will speed up the process and improve accuracy. To differentiate between photographs of the normal retina and those with diabetic retinopathy, this model is suggested. A more refined model, using a more comprehensive dataset, will allow for the early detection of all forms of diabetic retinopathy in the near future. This CNN is assistance for the ophthalmologists for healthcare diagnosis. The proposed convolution neural network model is only restricted to abnormal blood vessels, haemorrhages, micro aneurysms and exudates in the diabetic retinopathy. These cotton wool spots, drusen, macular edema like diseases are not extracted by the proposed model. There is a massive potential to find, discover the unknowns and study more into the healthcare domain to the accessibility of healthcare database, only to expand the superiority of human being by way of prevention and improved controlling of health related issues.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] S. Yamini, .V. Khanaa, Krishna Mohantha, A state of the art review on various data mining techniques, Int. J. Innov. Res. Sci. Eng. Technol. 5 (3) (2018).

[2] J.-J. Yang, J. Li, J. Mulder, Y. Wang, S. Chen, H. Wu, Q. Wang, H. Pan, Emerging information technologies for enhanced healthcare, Comput. Ind. 69 (2019) 3–11.

[3] Han Wu, Shengqi Yang, Zhangqin Huang, Jian He, Xiaoyi Wang, Type 2 diabetes mellitus prediction model based on data mining, Inform. Med. Unlocked 10 (2018) 100–107.

[4] Ioannis Kavakiotis, Olga Tsave, Athanasios Salifoglou, Nicos Maglaveras, Ioannis Vlahavas, Ioanna Chouvarda, Machine learning and data mining methods in diabetes research, Comput. Struct. Biotechnol. J. 15 (2019) 104–116.

[5] Sajida Perveen, Muhammad Shahbaz, Guergachi Aziz, Karim Keshavjee, Performance analysis of data mining classification techniques to predict diabetes, Proc. Comput. Sci. 82 (2019) 115–121.

[6] Saba Bashir, Usman Qamar, Farhan Hassan Khan, IntelliHealth: a medical decision support application using a novel weighted multi-layer classifier ensemble framework, J. Biomed. Inf. 59 (2019) 185–200.

[7] Dilip Kumar Choubey, Sanchita Paul, Bhattachrjee Joy, Soft computing approaches for diabetes disease diagnosis: a survey, Int. J. Appl. Eng. Res. ISSN: 0973-4562 9 (21) (2021) 11715–11726.

[8] Abdullah A. Aljumah, Mohammed Gulam Ahamad, Mohammad Khubeb Siddiqui, Application of data mining: diabetes health care in young and old patients, J. King Saud Univer. Comp. Inform. Sci. 25 (2020) 127–136.

[9] Krati Saxena, Dr. Zubair Khan, Shefali Singh, Diagnosis of diabetes mellitus using K nearest neighbor algorithm, Int. J. Comp. Sci. Trends Technol. 2 (4) (2018) 101–120.

[10] Tawfik Saeed Zeki, Mohammad V. Malakooti, Yousef Ataeipoor, S. Talayeh Tabibi, An expert system for diabetes diagnosis, Am. Acad. Sch. Res. J. 4 (5) (2019) 44–52.

[11] Anindito Yoga Pratama, R. Dewi Agushinta, Remi Senjaya, Design of mobile expert system for diabetes risk diagnosis and information, J. Inf. Syst. 9 (1) (2018) 32–36.

[12] Abdulla Al-Malaise Al-Ghamdi, et al., An expert system of determining diabetes treatment based on cloud computing platforms, Int. J. Comput. Sci. Inf. Technol. 2 (5) (2019) 1982–1987.

[13] G. Naveena Pai, M. Swathi Pai, M. Shruthi, B. Naveen K, Internet of things: a survey on devices, ecosystem, components and communication protocols, in: 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 611–616, https://doi.org/10.1109/ICECA49313.2020.9297458.

[14] S.R. Kawale, S.P. Diwan, D.G. V, Intelligent breast abnormality framework for detection and evaluation of breast abnormal parameters, in: 2022 International Conference on Edge Computing and Applications, (ICECAA), 2022, pp. 1503–1508, https://doi.org/10.1109/ICECAA55415.2022.9936206.

[15] Meysam Rahmani Katigari, Haleh Ayatollahi, Mojtaba Malek, Mehran Kamkar Haghighi, Fuzzy expert system for diagnosing diabetic neuropathy, World J. Diabetes 8 (2) (2019) 80–88.

[16] Ranjeet Suryawanshi, C.R. Revanna, B. Kameswara Rao, Parismita Sarma, Enhanced diagnostic methods for identifying anomalies in imaging of skin lesions, IJEER 10 (4) (2022) 1077–1085, https://doi.org/10.37391/IJEER.100452.

[17] E. Onuiri Ernest, U. Ndukwe Victor, Igwe Nkechinyere, C. Olise Saviour, Simulation of an expert system for diabetes diagnosis, Int. J. Adv. Comp. ISSN: 2051-0845 48 (1) (2021) 1614–1621.

[18] L. Rajesh, Mirzanur Rahman, Ghazaala Yasmin, A. Parismita Sarma, Azhagu Jaisudhan Pazhani, A novel method of data compression using ROI for biomedical 2D images, Measurement: Sensors. ISSN: 2665-9174 24 (2022), 100439, https://doi.org/10.1016/j.measen.2022.100439.

[19] Ursula Schmidt-Erfurth, Amir Sadeghipour, Bianca S. Gerendas1, Sebastian M. Waldstein, Hrvoje Bogunovic, Artificial Intelligence in retina, Prog. Retin. Eye Res. 67 (2020) 1–29.

[20] K.R. Swetha, A.R. Namitha, Y.M. Manu, G.R. Rashmi, Veera Sivakumar Chinamuttevi, IOT based smart health care system to monitor covid-19 patients, IJEER 10 (1) (2022) 36–40, https://doi.org/10.37391/IJEER.100105.

[21] M. Ramesha, K. Jeevan, B.M. Sathisha, Implementation of IoT based wireless electronic stethoscope, in: 2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT), 2020, pp. 103–106, https://doi.org/10.1109/MPCIT51588.2020.9350476.

[22] D.R. M. R. G, H. Anandaram, Extraction of fetal ECG using ANFIS and the undecimated-wavelet transform, in: 2022 IEEE 3rd Global Conference for Advancement in Technology, (GCAT), 2022, pp. 1–5, https://doi.org/10.1109/GCAT55367.2022.9972078.

[23] N. Singh, R. Chandra Tripathi, Automated early detection of diabetic retinopathy using image analysis techniques, Int. J. Comput. Appl. 8 (2) (2019) 18–23.

[24] Y. Zhao, Intensity and compactness enabled saliency estimation for leakage detection in diabetic and malarial retinopathy, IEEE Trans. Med. Imag. 36 (1) (2018) 51–63.

[25] Ravi Sanakal, Smt. T. Jayakumari, Prognosis of diabetes using data mining approach-fuzzy C means clustering and support vector machine, Int. J. Comput. Trends Technol. 11 (2) (2018) 94–110.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

A Novel Method of Enhancing Security Solutions and Energy Efficiency of IoT Protocols

KDV Prasad[1], D Palanikkumar[2], Dankan Gowda V[3], Naziya Hussain[4], A. Azhagu Jaisudhan Pazhani[5], Anil Kumar N[6]
[1]Assistant Professor (Research), Symbiosis Institute of Business Management, Hyderabad, Symbiosis International (Deemed University), Pune, India. e-mail: Kdv.prasad@sibmhyd.edu.in
[2]Professor, Department of Computer Science and Engineering, Dr NGP Institute of Technology, Coimbatore, Tamilnadu, India. e-mail: palanikkumard@gmail.com
[3]Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India. e-mail: dankanies@gmail.com
[4]Associate Professor, School of Computers, IPS Academy, Indore, Madhya Pradesh, India. e-mail: naziyahussain@gmail.com
[5]Associate Professor, Department of Electronics and Communication Engineering, Ramco Institute of Technology, Rajapalayam, Tamilnadu, India. e-mail: alagujaisudhan@gmail.com
[6]Assistant Professor, Department of Electronics & Instrumentation Engineering, School of Engineering, Mohan Babu University (Erst while Sree Vidyanikethan Engineering College), Tirupati, Andhra Pradesh, India. e-mail: anilkumar.n@vidyanikethan.edu
Corresponding author: KDV Prasad, e-mail: Kdv.prasad@sibmhyd.edu.in

Abstract— Mobile Ad-hoc Networks (MANET's) are wireless networks that are capable of operating without any fixed infrastructure. MANET routing protocols must adhere to strict secrecy, integrity, availability and non-repudiation criteria. In MANETs, attacks are roughly categorised into two types: active and passive. An active attack attempts to modify or remove data being transferred across a network. On the other hand, passive attack does not modify or erase the data being sent over the network. The majority of routing protocols for MANETs were built with little regard for security and are therefore susceptible to a variety of assaults. Routing technologies such as AODV and dynamic source routing are quite common. Both however are susceptible to a variety of network layer attacks, including black holes, wormholes, rushing, byzantine, information disclosure. The mobility of the nodes and the open architecture in which the nodes are free to join or leave the network keep changing the topology of the network. The routing in such scenarios becomes a challenging task since it has to take into account the constraints of resources of mobile devices. In this an analysis of these protocols indicates that, though proactive routing protocols maintain a route to every destination and have low latency, they suffer from high routing overheads and inability to keep up with the dynamic topology in a large sized network. The reactive routing protocols in contrast have low routing overheads, better throughput and higher packet delivery ratio. AODVACO-PSO-DHKE Methodology boosts throughput by 10% while reducing routing overhead by 7%, latency by 8% and energy consumption by 5%. To avoid nodes always being on, a duty cycle procedure that's also paired with the hybrid method is used ACO-FDR PSO is applied to a 100-node network and NS-3 is used to measure various metrics such as throughput, latency, overhead, energy consumption and packet delivery ratio.

Keywords- Security, Throughput, Internet of Things, Energy Efficient, Mobile Ad-hoc Network, Pocket Delivery Ratio and reactive protocols.

I. INTRODUCTION

A seamless connectivity to users is provided by wireless networks regardless of their location. There are two kinds of wireless networks; one based on fixed infrastructure and the other is Ad hoc networks independent of a permanent infrastructure of any kind. The mobile devices in infrastructure based wireless networks are connected to fixed routers and gateways through a network of access points/base stations [1]. The mobile device in range of a base station remains connected to the network and on moving out of range is handed over to the next base station. A cellular network or a WiFi Local Area Network (LAN) is the example of infrastructure based network. The MANETs assume a trusted behaviour among nodes and are characterized by lack of centralized infrastructure, dynamic topology, open architecture, processing power, short range, and limited bandwidth, memory for storage and battery power [2]. The characteristics of MANETs have the following implications on the operations of MANETs: Multi-hop relaying: When source node and destination node are out of range, then, source node use their neighbour node as a relay nodes to send the packet from sender to receiver. Dynamic topology: MANET uses dynamic topology. Each node moves at different speeds from one place to other place and they dynamically establish routes among themselves.

Energy management: Nodes in MANET fully depend on battery power. Energy management plays a vital role in MANET. MANETs are finding increasing use in military, disaster management, search and rescue operations, surveillance networks, law enforcement operations, sensor

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

networks, vehicular ad hoc networks, public health, security and many commercial applications. In military the MANETs find extensive use to network various entities in hostile terrain or terrain where no infrastructure exists. It can be used to network devices carried by a small body of troops for surveillance and communication in an area which lacks networking infrastructure or can even network the entities of network-centric warfare in a battlefield in an enemy area where networking infrastructure is not available [3]. The MANETs can also be used to network a flotilla on high seas where no network is available. The disaster management and the emergency relief operations require networking of diverse teams, where existing infrastructure has been made inoperable; MANETs can effectively network all participating entities. The security agencies are often required to establish fast and secure network for law enforcement operations and the requirement can be met through ad hoc networks [4]. The machine-to-machine communication, sensor networks, smart home applications, connected cars, conferences and various business applications are some of the commercial applications of MANETs. The diverse applications of MANETs require quality of service, reliability and secure communication. Security of the networks is a priority, before these networks can be effectively employed in military and security related applications.

On-demand, Table-driven, and hybrid are the three types of routing protocols. The classification is clearly depicted in the Figure.1. Table-driven or PRP keep a routing table on each node. Routing tables are frequently sent across the network to keep the list of destinations and routes up to date. It always keeps a route open between any two hosts. A hybrid routing system combines proactive and reactive features to provide the best of both worlds. For the nodes in the routing zone, proactive routing protocols are employed. Outside of the routing zone, nodes must employ an in-band reactive routing protocol (IRRP). The security of MANETs is very important due to its increasing deployment in various applications and requires constant innovations and up gradations. Security was not the major design objective of most of these protocols and their emphasis was on simplicity and efficiency so that the routing protocols are light on resources [5].

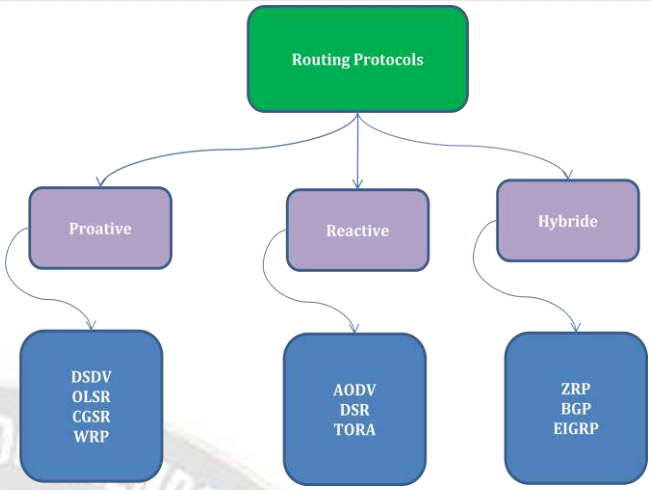


Figure.1: Categorization of protocols for routing.

The protocols were not addressing the issue of provision of security against often unpredictable attacks and making the protocol's reliability doubtful under such circumstances. In MANETs, there is no central mechanism for access control and traffic control and therefore it is hard to discriminate a malicious node from a legitimate node. An attacker therefore can easily gain access and disrupt the network operations. The security solutions for MANETs should therefore meet the following security objectives:

Confidentiality: It ensures that the authorized user only has access to the data and a malicious node is not able to eavesdrop and get the confidential information. The confidentiality is normally achieved by encryption of the data.

Availability: The requisite resources, like bandwidth and connectivity with other nodes of the network are available in timely manner to the legitimate nodes. Integrity: It guarantees that the data being transmitted has not undergone any modification or alteration by a malicious node.

As security becomes more of a concern, a secure communication channel must be established to allow mobile nodes to communicate in a antagonistic environment. A security environment is needed to overcome a set of challenges. Self-communicating devices don't need infrastructure to share information and constitute a MANET. Reducing network inefficiency is a major priority for computer experts who want to improve the central administration's network performance. ACO comes up with a higher-lasting energy-efficient route based on the ants' unique knowledge to fine-tune and is used to help the nodes to have a longer life by making sure that as little energy is wasted during transmission as possible [6]. Other techniques are examined to see whether they produced similar findings and also address the security concerns, including the inherent vulnerability of MANET's It is necessary to analyze network, security and the major attack types in order to provide network and information security. The major contributions of this work are: Enhancing

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

Security Solutions, Energy Efficient Transmission, Throughput enhancement and Identifying identify the vulnerabilities and types of attacks on routing protocols to analyze the impact of attacks and identify techniques for securing the routing protocols.

II. LITERATURE REVIEW

Securing of routing protocols for MANETs is an open research problem, many researchers have worked to identify the vulnerabilities and have proposed various techniques for securing the routing protocols. In [7] author's, compared the proactive and reactive routing protocols and have concluded that, the proactive protocols have lower route acquisition delay, since they maintain routes to all the nodes in their cache, but have requirements of large memory, greater processing power, more bandwidth, higher power and high control overheads. AODV, DSR, Ad-hoc On-demand Multi-path Distance Vector (AOMDV) and the proactive DSDV routing protocols decreased sharply with increasing traffic load. Maximum decrease was observed for the proactive protocols. PDR of all three reactive routing protocols was higher than that of the proactive routing protocol in the changing network topology. Proactive protocol, DSDV was not able to update the routes due to frequent

In [8] author's, observed that the hybrid routing protocols combine the proactive and reactive techniques to reduce the communication overheads and latency. The proactive routing protocol out of the three types had the worst PDR and throughput. The hybrid routing protocols had low communication overheads and latency but these advantages were dependent on the extent of the zone and zone dynamics. However, DSR had lesser NRL than that of AODV under all conditions. DSR showed maximum energy saving; however, the performance of DSR reduced considerably as compared to AODV at higher traffic loads. AODV outperformed DSR in all scenarios, except for small and medium size networks under low load conditions.

The protocols have been compared for the routing overheads, average delay per flow and PLR. DSDV had advantage over OLSR, and the reactive routing protocols in small and medium size networks with static or slow moving nodes. OLSR in contrast performed better in large and high density networks since the routing control traffic was minimized using Multi Point Relay (MPR) nodes. OLSR used smaller updates with an interval value of 1 sec as compared to DSDV, which used larger updates with interval value up to 5 seconds. The average delay per flow in OLSR was better than that of DSDV. DSDV exhibited the worst performance even as compared to reactive routing protocols in large and fast changing network topology. It had high PLR which can be attributed to its inability to update the routes in dynamic topology of the network.

In [9] author's discussed the attacks and their countermeasures across the complete stack of layers of MANETs. At the network layer, the authors identified the vulnerabilities of the routing process which led to attacks on proactive and reactive routing protocols. The defense mechanism suggested against passive attacks involved encryption of data. Customized solutions were proposed in the literature against every attack which may require additional hardware with nodes or modification of the protocols. In [10] author's, concluded that the security of MANETs was still in early stages and was an open challenge. The proposed solutions can typically provide security against few specific attacks and more research in this area was required. In [11] author's suggested the mitigation techniques for the various attacks, also conducted the survey of Intrusion Detection Systems (IDSs) for protecting MANETs from range of attacks.

In [12] author's discussed Flooding, Black-hole, Link-with holding, Link spoofing, Replay attack, Worm-hole and Colluding Miss-relay attacks on AODV and OLSR routing protocols. The authors suggested counter measures against each attack which were based on, monitoring the behaviour of the nodes and assigning trust values for flooding attack, re-confirming the authenticity of routing messages from neighbouring nodes for Black-hole attack, use of special hardware and cryptographic techniques for Link spoofing and Wormhole attack. The relative merits of each technique were compared and the authors concluded that the proposed solutions were not optimal when effectiveness and efficiency was considered. Therefore, the future research should explore the cost effectiveness of the solutions so as to make them suitable for MANETs. The types of attacks and the analysis of the attacks based on the type of protocols were presented by [13]. The authors analysed the effects of various attacks on the proactive and reactive routing protocols to determine if any protocol provided inherent protection against the attacks. The authors classified the attacks into, attack on sequence number, misdirection and DoS attacks. The Sequence number could be manipulated and flooding the network with spoofed packets with very high sequence number would discard the routing request from the legitimate node. The misdirection attacks could be launched by an attacker by various methods such as; impersonation as an existing node or a previously connected node, or through replay attacks or by attracting traffic towards it or deflecting traffic away from it or through Wormhole attack [14]. Different methods were used to exploit the vulnerabilities of routing protocols to launch such types of attacks. The DoS attacks were aimed at consuming the resources of the network by sending fake messages. The authors recommended authentication of originator and integrity check to prevent against most of the attacks and

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

concluded that proactive protocols were inherently more secure than reactive routing protocols.

The routing messages received from malicious nodes were discarded. The simulation results of the suggested technique for detection of Black-hole attack showed an improvement of 20 % in throughput and 40 % improvement in PDR.

The attacker in Sybil attack acquired multiple identities. Author's in [15] in their research proposed that the protection against the Sybil attack can be provided through authentication of a unique and unchangeable identity of the nodes through a Trusted Third Party (TTP). The authors proposed that the identity of the node can be tied down with the hardware of the node and suggested using the serial number of the hard disk as the unchangeable identity. The proposed technique authenticated the identities of the nodes with minimum use of TTP; however, it took much longer time for authentication of legitimate nodes. The authors were of the view that self-Certification Authority and distributed CA, though highly practical could not provide the security against Sybil attacks. A Central CA was required as TTP for issue of certificates for which more research was required for issue of keys and the certificates to the participating nodes.

To detect the malicious node using simultaneous identities attempting to launch a high bit rate attack, author's [16] proposed to track the movement of all the identities. The sender associated its location with every packet which was signed by its private key. The detection algorithm constructed the path each identity travels. The identities travelling similar paths were considered Sybil. The proposed technique required additional hardware in terms of directional antennas or GPS for determining the location of the nodes. The experimental results indicate 80% accuracy in detecting Sybil identities at slow speed of 5 mps; however, the accuracy dropped as the speed of the identities was increased.

The protocols in these broad categories even differ in their operation and have different characteristics. An in-depth study of routing protocols and the performance analysis of their characteristics are required to be undertaken[17]. The AODV and DSR are extensively used reactive routing protocols and the performance of these protocols has been compared in the literature. However, the performance analysis of these two protocols is not conclusive and requires further investigation under different scenarios. The vulnerabilities and attack on routing protocols were studied in the literature in piecemeal manner. A comprehensive study of the vulnerabilities of the routing protocols, the counter measures against the attacks exploiting the vulnerabilities and the techniques adopted by the existing secure routing protocols are important and is required to be studied. Security to the routing protocols is provided by a number of cryptographic techniques [18]. A detailed study of various techniques and a comparative analysis of various algorithms is required to be done.

To find the network's Black Hole, most current approaches look at the network from the perspective of its neighbours. However, this is less effective than dealing directly with the final node, since the journey from one source to another may not always be safe. The limited wireless range of the wireless node becomes yet another factor when considering this technique of neighbour negotiation since it is not a definite condition that nodes remain in range of each other.

III. SECURE AND EFFICIENT DATA TRANSMISSION

MANETs include several Sensor nodes and the placement of Sensor Nodes determines where information is sent. The nodes may move at will. ACO methods are used to optimize the efficient route construction utilizing the AODV routing protocol. Network sensors are clustered using Particle Swarm Optimization (PSO) algorithms. DHKE algorithms are used in sensor networks to protect individual packets. Due to their unique capacity to interact outside the node's immediate vicinity, MANETs are different from conventional wireless networks. When the destination is beyond a node's range of communication, the nodes work together to send the data there instead. As shown in Figure.2, the most extensively used and prevalent procedures in each of these groups are highlighted. A quick overview of these procedures is given, in this paper is for you. It has been decided to compare and contrast distinct types of protocols as well as the protocols within each category. RLEACH approached, a safe solution for LEACH clustering formed dynamically and regularly were developed [19]. A random pair-wish key orphan node was a big issue in RLEACH. Practical cryptography approach was proposed to solve the problem of orphan nodes. PSO and DHKE algorithms were used for clustering and security respectively. The protocols in this category differ from each other on the basis of number of tables maintained by each node and the algorithms for computing the routes. The changes in the network topology are communicated through routing update messages, which help in computing the latest routes.

IV. AODVACO-PSO-DHKE SYSTEM

In every ad-hoc wireless network, security is the primary issue. An innovative new solution for sensor network communication security, the AODVACO-PSO-DHKE is described in this research. In addition, the AODVACO-PSO-DHKE system will protect the data from a wide range of threats. As a result, DHKE requires less processing power while using less storage space. It's a Dynamic Optimization Problem for MANET routing since the search area varies over time. The routing policy determines which node must connect with which other on the route to get to the final node. In all, there are eight

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

critical phases to this system: Development of mobile nodes, clustering with PSO, selection of cluster heads based on node weight, an estimate of routers with AODVACO, the transmission of data to the cluster head, encryption with DHKE, receipt of data successfully, and finally, decryption [20]. Using the PSO-based clustering technique, the three steps need to be followed are: initializing the particles, calculating the fitness function, and updating their positions. PSO integrates local and global fitness in its search method. Predicting the current location of the legitimate data using particle changes and the associated data of nearby particles yields the ideal answer via iteration.

Two greatest esteems, for example, the original one, is used to update the particle on each iteration [21]. An individual particle can achieve the best Pbst. The global set of Pbst values yields a second-best value, which is denoted by gbst. A random deposition of particles and the use of the K-means clustering technique to get centroids is the first stage in particle initialization. PSO method was used to find the best possible locations for these points. In a fitness function, the particles' clustering scatter is considered. The created cluster scatter used as an input parameter for the fitness functions, provides a numerical value to verify the clustering scatters' accuracy.

$$F = \frac{1}{m}\sum_{i=1}^{m}\left\|O_j, D_j{}^{C_j}\right\| \quad (1)$$

The optimal particle for the cluster head is obtained via PSO. In order to get this information, it is necessary to calculate the distance between each node in the neighbourhood and the optimal centroid. D best is computed and the cluster that has the lowest D best is chosen

The D_best value is recalculated and the particle's location is updated as a result of the foregoing procedure [22]. The distance between the cluster's centre and its surrounding particles is used as a measure of the cluster's fitness. d best is equal to the minimum. The velocity of a particle has an effect on its location as given below:

Let $x_i(t)$ denote the particle position i in the search space at time step t. Then

$$x_i(t+1) = x_i(t) + v_i(t+1) \quad (2)$$

$$v_i(t) = v_i(t-1) + c_1 r_1\big(localbest(t) - x_i(t-1)\big) + c_2 r_2(globalbest(t) - x_i(t-1)) \quad (3)$$

Cluster networks of Sensor Nodes may be seen in Figure.3, clustered together. Each cluster network elects its CH based on the degree value that is the bare minimum. The cluster-role head collects data from nearby nodes and sends it to another CH on another network. Confirmation messages from the CHs give a time slot plan for their members to employ during the steady-state phase of their communication. Clustering changes dynamically in the "AODVACO-PSO-DHKE" approach, which is known as dynamic clustering.

V. ENERGY EFFICIENCY

As security becomes more of a concern, a secure communication channel must be established to allow mobile nodes to communicate in a antagonistic environment. The features of mobile adhoc networks offer a variety of potential security risks. Security features face nontrivial problems, such as open peer-to-peer networks [23]. Shared wireless media with rigorous network architecture restrictions on available resources, together with changing network topology, provide a difficult design challenge. A security environment is needed to overcome a set of challenges. Self-communicating devices don't need infrastructure to share information and constitute a MANET. Reducing network inefficiency is a major priority for computer experts who want to improve the central administration's network performance. A Collaborative Operating Committee (ACO) comes up with a higher-lasting energy-efficient route based on the ants' unique knowledge to fine-tune the nodes in the network. MANETs is a kind of multi-hop network, built from the ground up with no previous infrastructure in place. The major difficulty system security designer’s face is due to the wireless and dispersed nature of the system. The emphasis placed on MANET security issues has dramatically increased in the last few years. Last but not least, by dividing a message into N pieces and spreading them over the routes, a limited degree of transmission redundancy is provided with the goal of enabling. Each bit is protected by an integrity and replay-proof cryptographic header that verifies and tracks where it is coming from, as well as who created it. When one or more items are delivered, the destination includes an acknowledgment which notifies the source of which items arrived undamaged, and thus routes. To promote the strength of the feedback system, the acknowledgement In case the source receives no more than M pieces, the source retransmits the remaining pieces using all the previously established routes. If the protocol recognizes that a small percentage of recognized pieces is insufficient, or if a substantial percentage throughout the path is set, among other things, if this is not the case, the computer will transmit the next message in the queue to the recipient.

VI. ANALYSIS AND PREVENTION

The TCP/IP protocol suite is used by MANETs, and the underlying layers adhere to IEEE 802.11 specifications. MANET-specific routing protocols are used at the network layer, where most of the differences may be found. To perform its functions of transmission of messages, neighboring relationship formation and collection of route information in routing databases, the routing protocols exchange control messages. The routing protocols are subjected to threats at various levels. An attacker may attack the control messages to break the neighboring relationship or may attack the routing

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

control messages to disrupt the routing process. There are a number of ways that an attacker might disrupt adjacent relationships in a routing protocol, such as by modifying 'Hello' messages or by altering Topology Control (TC) messages. The RREQ, RREP, and RERR control messages in the Reactive Routing protocol may also be attacked in the same way. An outsider or a Byzantine node might be the perpetrator. It is possible for an outsider to serve as both a source and an intermediary node. To launch an attack, it may either use its own address, another node's address, or any other address it chooses. It is possible for a malicious node to alter or replay received packets, causing a route interruption or denial of service, if it is functioning as an intermediary node. Byzantine nodes on the other hand are faulty, misconfigured or subverted nodes which are legitimate participants of routing process. The threat consequences due to actions of the attackers compromise the correct behavior of the routing process which can damage the data traffic intended for a particular node. Figure.5. depicts the taxonomy of attacks which are explained as under:

By monitoring/intercepting network traffic, a passive attacker may figure out what is going on. important and critical information without disturbing the routing process. The information could be about the topology, location of important nodes or the identity of the nodes, type of data exchanged on the network. Some of the prominent passive attacks are as follows:

Sniffing Attack. The wireless transmission over the MANET can be heard by every device equipped with a receiver and within a range. An attacker monitors/records the routing exchange control messages on the network. The consequences of sniffing depend on the transmission range of nodes, geographical distribution of nodes, location of the attacker, type of routing protocol and the information.

Traffic Analysis. An attacker can listen to the data traffic that is flowing through the network and can deduce the routing information by analyzing the amount of data transmitted, communication pattern and the type of data transmission. The analysis can be carried out even if the data is encrypted and can reveal the location of important nodes which can be targeted.

Spoofing Attack. In this type of attack, the attacker assumes the identity of a legitimate node. Spoofing in itself does not constitute an attack but can be used to launch other attacks. The consequences of spoofing are that an attacker will be able to know the routing information and also can disrupt the neighboring relationship with other nodes.

Routing protocol flaws, attacks that make use of these vulnerabilities, and methods to prevent these attacks are all covered in this paper. It can be observed that no single method can provide security against all attacks. The vulnerabilities from external attackers and internal attackers have to be addressed in a separate manner. Some of the secure versions of routing protocols have been developed and proposed in the literature. However, none of these protocols has been able to provide protection against all types attacks. Moreover, the secure version of the routing protocols excessively uses the cryptographic techniques which add excessive computation load on the resource constrained nodes. The requirement is therefore to devise a light weight cryptographic technique which meets the goals of security. A Black Hole attack may be more damaging to the AODV protocol than to the OLSR protocol, according to our investigation.

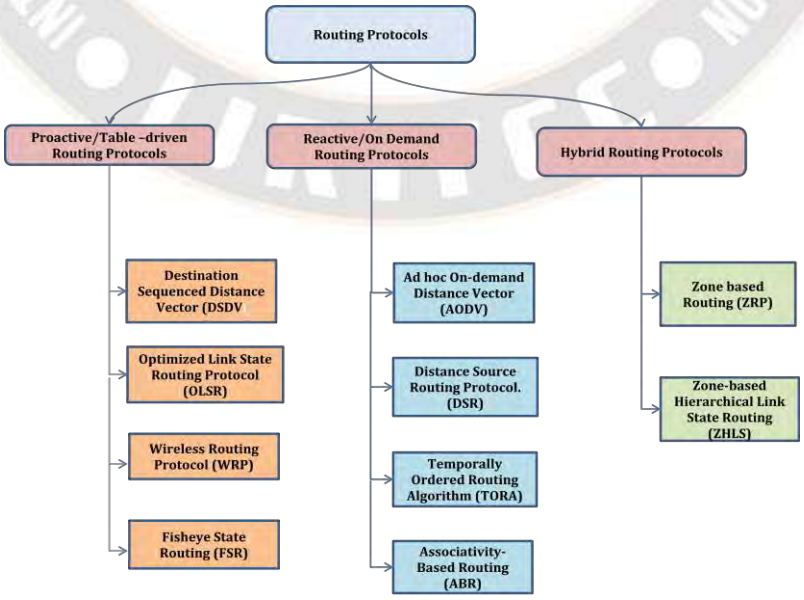


Figure.2: Types of Routing Protocols

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

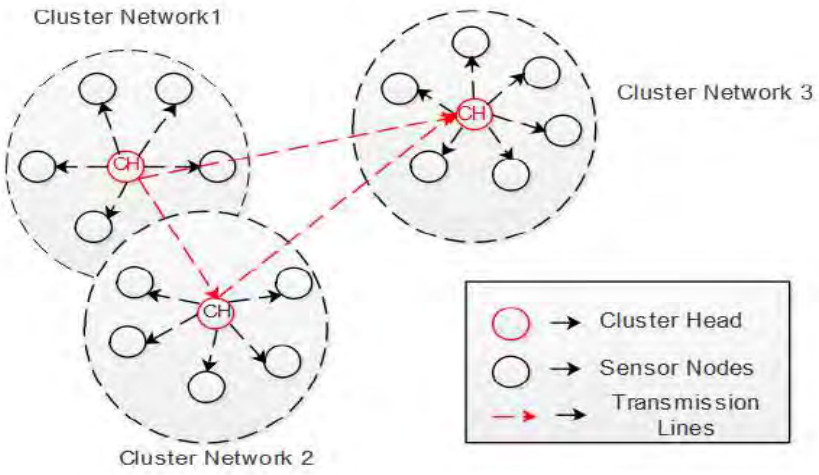


Figure.3: Basic Clustering in Mobile ad-hoc network

Table.1: Comparative Analysis of Proactive Routing Protocols

Parameter	DSDV	OLSR	WRP	FSR
Routing Algorithm	Distributed Bellman-Ford	Optimized Link State Routing	Improved distributed Bellman-Ford	Link State routing algorithm
Routing Structure	Flat	Flat	Flat	Hierarchical
Route Metric	Shortest Path	Shortest Path	Shortest Path	Shortest Path
Number of Tables	2	3	4	3 and a List
Frequency of Updates	Periodic and as required on occurrence	Periodic	Periodic	Periodic exchange of updates more frequent in nearby nodes. Less frequent for nodes far apart.
Control Overheads	High	Low	High	Low
Scalability	Small and medium Size	Large and Dense Network	Small and medium Size	Large and Dense Networks
Convergence of routes	Slow as compared to WRP	Fast	Fast	Fast
Loop-free	Sequence number prevents loops	Link State entry prevents loops	Records Predecessor and Successor nodes. Additional overheads are incurred for consistency checking.	Link State entry prevents loop.

Table.2. Comparative Analysis of Reactive Routing Protocols

Parameter	AODV	DSR	TORA	ABR
Routing Algorithm	Distance Vector	Source Routing	Link Reversal	Degree of Association Stability.
Multiple Routes	No	Yes	Yes	No
Route Metric	Shortest path	Shortest path	Height	Link Associativity
Control Traffic Volume	Low Overheads	Overheads increase with the size of the network	Large Overheads due to maintenance of height and status of connected link.	Large overhead. due to beacon updates.
Scalability	Small, medium and Large Dense Networks	Small and Medium Networks	Large Dense Networks	Small and Medium networks.
Link Failure Information	Flood, Erase Route and inform Source Node	Flood, Erase Route and inform Source Node	Link Reversal and Route Repair	Local Broadcast Query.
Loop Free	Yes Sequence number	Yes Checking root record field,	Yes Link State entry prevents loop	Yes Link State entry prevents loop.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
Article Received: 19 February 2023 Revised: 02 April 2023 Accepted: 20 April 2023

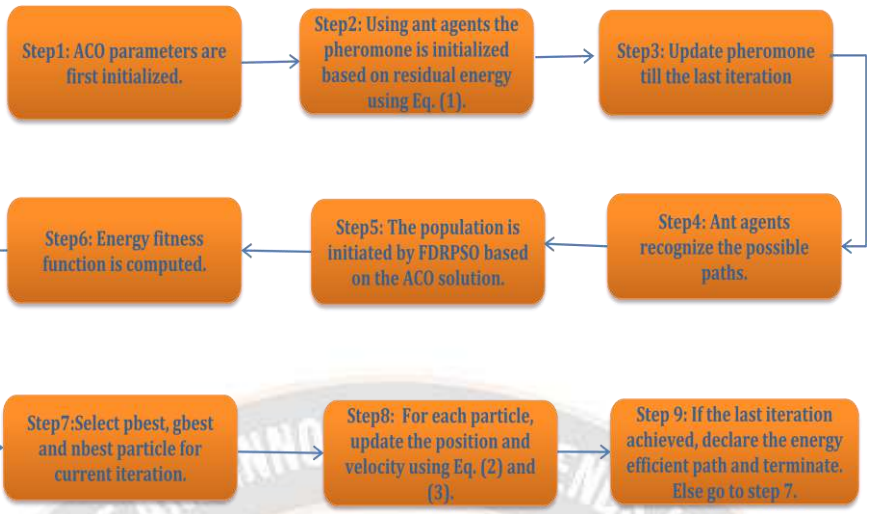


Figure.4: Steps of ACO-FDR PSO

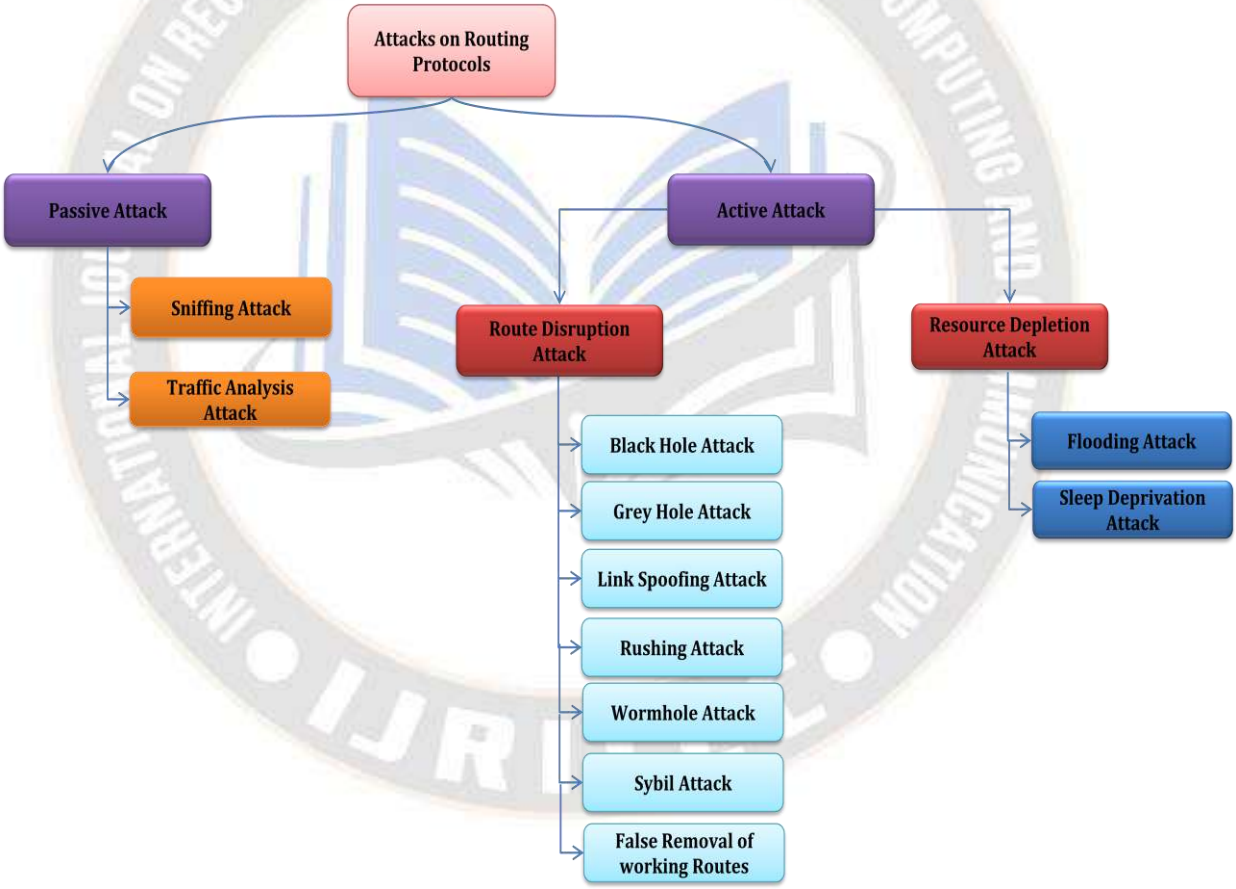


Figure.5: Classification of Attacks on Routing Protocols

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

VII. RESULTS AND DISCUSSION

Data transmission is accomplished by utilizing an efficient clustering and routing technique that uses the AODV routing protocol with the ACO optimization algorithm written in NS2. The I7 machine with 8 GB of RAM is used for all of the tasks. PSO algorithm and the DHKE method are both used to find the most efficient way for wireless mobile nodes to transmit data securely. AODV-ACO and DHKE algorithm findings are presented in this section in further depth. The nodes' communications are secure thanks to the AODV-ACO-PSO-DHKE method. Throughput, routing overhead, latency and energy usage are all taken into account when calculating the performance. A reduction in routing overhead results in an improvement in throughput and PDR.

$$Throughput = \frac{Total\ packets\ received\ at\ the\ destination}{Node\ total\ Simulation\ time}$$

$$Roughting\ Overhead = \frac{Total\ no.of\ routing\ packets}{Total\ no.of\ delivered\ data\ packets}$$

Energy consumption: The enormous quantity of incoming energy usage is mirrored by the enormous number of hops.

As shown in the simulation results, AODVACO-PSO-DHKE Methodology improves throughput by 10% in comparison to AODV-PSO Methodology [16]. AODV-PSO Methodology has a routing overhead of 7 percent less than AODVACO-PSO-DHKE. the DHKE method and the AODVACO-PSO method are compared. AODVACO-PSO-DHKE reduces delay by 8% compared to AODV-PSO Methodology. AODVACO-PSO-DHKE Methodology reduces energy consumption by 5% in comparison to AODV-PSO Methodology. Figure.6. shows the comparison of nodes versus throughput for AODV-ACO-PSO-DHKE and AODV-PSO.

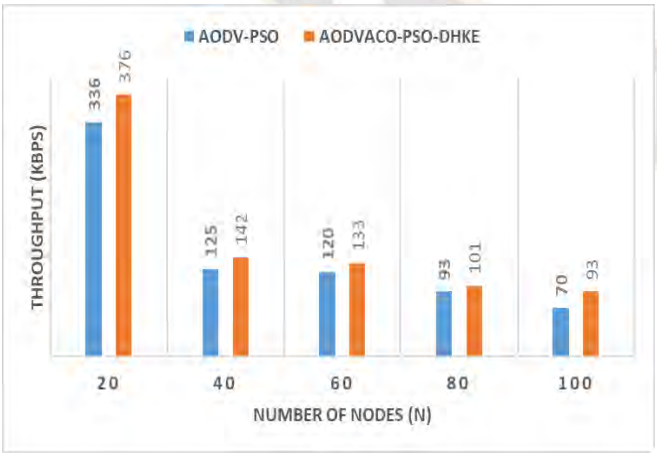


Figure.6. Node vs. Throughput

Figure.7. shows the comparison of AODV-ACO-PSO-DHKE vs. AODV-PSO in terms of nodes and routing overhead. When comparing the AODV-ACO-PSO-DHKE technique to the AODV-PSO method, the routing overhead is reduced by adjusting the number of nodes.

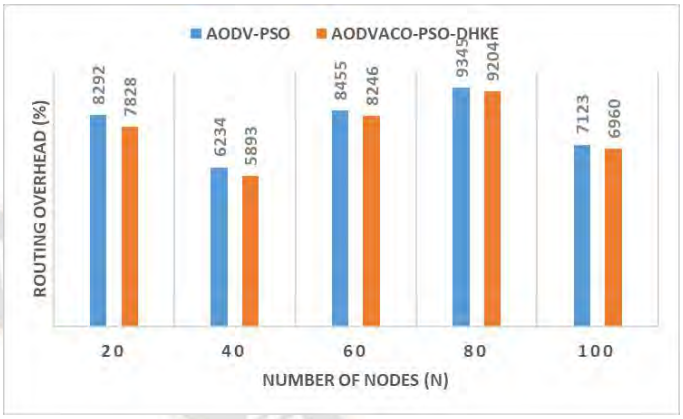


Figure.7. Node vs. routing overhead

PSO-DHKE and AODV-PSO are shown in Figure.8., which compares the number of nodes and the latency between the two protocols.

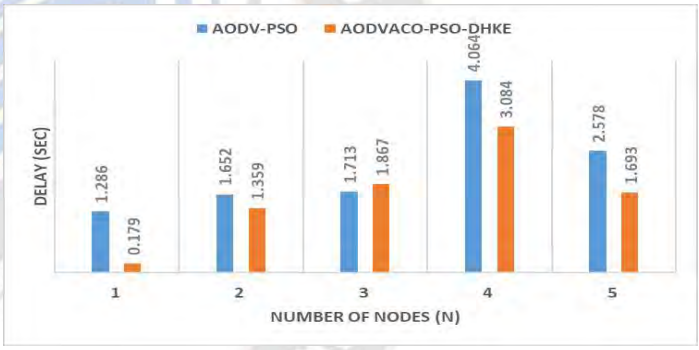


Figure.8. Node vs. delay

For AODV-PSO and AODV-ACO-DHKE, the node vs. energy comparison is shown in Figure.9. When AODV-ACO-PSO-DHKE technique is compared to AODV-PSO method, the energy consumption is reduced by adjusting number of nodes.

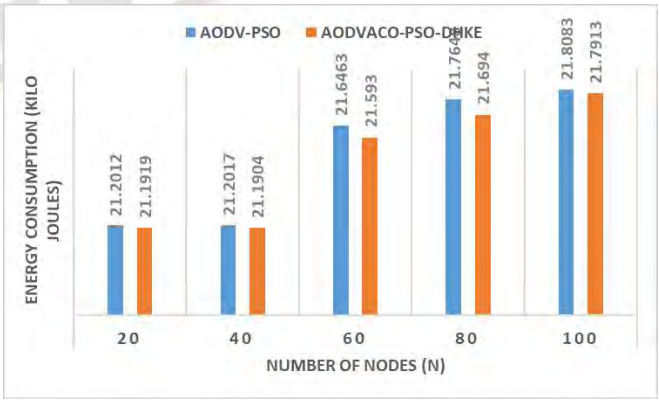


Figure.9. Node vs. Energy Consumption

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

By lowering the amount of overhead, time delay and energy used during the routing process, AODVACO-PSO-DHKE approaches have proven useful in sensor networks for safeguarding data packets from source to destination.

As a result of the M-TCFPA approach's reduced communication latency, the route discovery phase experiences shorter path identification and fewer control packet transfers. There is a significant difference in efficacy between M-TCFPA and TBSMR.

Table.3. Comparative analysis of M-TCFPA

Performances	Method	Number of Nodes				
		10	50	100	150	200
PDR (%)	TBSMR [20]	98	97	95	92	90
	M-TCFPA	99.77	99.37	98.50	99.04	98.85
PLR (%)	TBSMR [20]	2	3	5	8	10
	M-TCFPA	0.22	0.62	1.49	0.95	1.14
Throughput (kbps)	TBSMR [20]	510	500	470	440	410
	M-TCFPA	1090.36	1086.1	1077.85	1083.72	1083.27
AEED (sec)	TBSMR [20]	0.1	0.15	0.2	0.25	0.3
	M-TCFPA	0.024	0.027	0.038	0.060	0.070

Table.3. compares the TBSMR [20] and M-TCFPA [21] methods in terms of accuracy. Changing the number of sensor nodes used, i.e. 10, 50, 100, 150, and 200, allows comparisons to be made between different numbers of sensor nodes. Based on Table 2, it's clear that the M-TCFPA method [20] beats the TBSMR method. Only trust, traffic, and residual energy values were taken into account by the TBSMR [20] during data transmission. When transferring data packets, this TBSMR [20] does not take distance under account. There is a delay in transmitting the data packets because of this. The M-TCFPA approach makes use of trust and integrity to make sure the nodes are reliable and the data they exchange is accurate. As a result, the M-TCFPA method's data delivery is improved while transmission latency is reduced.

VIII.CONCLUSION

MANET's is an alternate wireless network architecture which is finding increasing use in defence, law enforcement, disaster management and commercial applications. This research work includes Enhancing Security Solutions, Energy Efficient Transmission, Throughput enhancement and Identifying the vulnerabilities. Ad-hoc networks may employ the "AODVACO-PSO-DHKE" approach for secure data transfer. As a result, the AODV routing protocol with ACO optimization is utilized for efficient data transfer. PSO Clustering is used to keep each node's energy consumption under control. DHKE approach is used to secure data transfer from source to destination. Regarding energy usage and throughput, the new approach is superior to the current method. Because of this, AODVACO-PSO-throughput DHKE's improves by 10% compared to AODV-PSO Methodology, while routing overhead and latency are down by 8%, and the energy consumption is down by 5% using AODVACO-PSO-DHKE. Performance analyses of routing protocols, the vulnerabilities of the routing protocols which lead to several attacks have been reviewed. Several proposed security enhancements for the routing protocols have been examined. Most of the existing routing protocols are developed to mitigate various network layer attacks such as black hole, wormhole and rushing attack used cryptography based technique and as well as trust based techniques. This may add additional overhead and processing time. It also decreases the throughput of the network. Our research is focused on developing secure routing protocol to mitigate various network layer attacks with less cryptographic techniques. The parameters considered are PDR, throughput, delay, packet loss and routing overhead. In order to construct a safe cluster-based routing system that can effectively transfer data, the M-TCFPA technique is proposed. Data transmission on the MANET can be improved by using the K-means clustering technique. Based on trust, integrity factor, residual energy, distance, and the M-TCFPA method's selection criteria, the best CH and routing are determined.

REFERENCES

[1]. S Sarkar and R Datta, A Secure and Energy-efficient Stochastic Multipath Routing for Self-organized Mobile Ad Hoc Networks, Ad Hoc Networks, 37 (2019), 209-227.

[2]. L Abusalah, A Khokhar, and Mohsen Guizani, A Survey of Secure Mobile Ad Hoc Routing Protocols, IEEE Communications Surveys & Tutorials, **10** (4) (2018), 78-93.

[3]. M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.

[4]. S. B. Sridhara, K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," Adv. Math. Sci. J., vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.

[5]. C. Toh, H.Cobb and D. Scott. "Performance Evaluation of Battery-Life-Aware Routing Schemes for Wireless Ad Hoc Networks", Proc. IEEE International Conference on Communication, Helsinki, Finland, pp.2824-2829, 2018.

[6]. Shivashankar, and S. Mehta, "MANET topology for disaster management using wireless sensor network," in International Conference on Communication and Signal Processing, ICCSP

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 5s
DOI: https://doi.org/10.17762/ijritcc.v11i5s.6643
***Article Received**: 19 February 2023 **Revised**: 02 April 2023 **Accepted**: 20 April 2023*

2016, 2016, pp. 0736–0740, doi: 10.1109/ICCSP.2016.7754242

[7]. Jayashri N, Veeresh Rampur, Durgaprasad Gangodkar, Abirami M, Balarengadurai C, Anil Kumar N, Improved block chain system for high secured IoT integrated supply chain, Measurement: Sensors, Volume 25, 2023, 100633.

[8]. Rajashanthi, M. and Valarmathi, K.: A Secure Trusted Multipath Routing and Optimal Fuzzy Logic for Enhancing QoS in MANETs. Wireless Personal Communications, 112, 75-90 (2020).

[9]. B.S., Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (eds) Recent Advances in Metrology . Lecture Notes in Electrical Engineering, vol 906. Springer, Singapore. https://doi.org/10.1007/978-981-19-2468-2_3.

[10]. S. B. M, P. Pavankumar, N. K. Darwante, "Performance Monitoring and Dynamic Scaling Algorithm for Queue Based Internet of Things," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2022, pp. 1-7, doi: 10.1109/ICSES55317.2022.9914108.

[11]. A. Singla, N. Sharma, "IoT Group Key Management using Incremental Gaussian Mixture Model," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), 2022, pp. 469-474, doi: 10.1109/ICESC54411.2022.9885644.

[12]. S. Reddy P, P. S. Patwal, "Data Analytics and Cloud-Based Platform for Internet of Things Applications in Smart Cities," 2022 International Conference on Industry 4.0 Technology (I4Tech), 2022, pp. 1-6, doi: 10.1109/I4Tech55392.2022.9952780.

[13]. A. Sharma, K. S and M. R. Arun, "Priority Queueing Model-Based IoT Middleware for Load Balancing," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), 2022, pp. 425-430, doi: 10.1109/ICICCS53718.2022.9788218.

[14]. M. Nagabushanam, H. G. Govardhana Reddy & K. Raghavendra (2022) Vector space modelling-based intelligent binary image encryption for secure communication, Journal of Discrete Mathematical Sciences and Cryptography, 25:4, 1157-1171, DOI: 10.1080/09720529.2022.2075090.

[15]. Pankaj Mudholkar, Megha Mudholkar, B S Puneeth Kumar and S. Srinivasulu Raju (2021), Smart Villages: IoT Technology Based Transformation, Journal of Physics: Conference Series, 2070(1), pp. 012128. https://doi.org/10.1088/1742-6596/2070/1/012128.

[16]. Gurung, S., Chauhan, S.: A dynamic threshold based algorithm for improving security and performance of AODV under black-hole attack in MANET. Wireless Networks, 25(4), 1685-1695 (2019).

[17]. Jacob, S.S., Hussain, N., Chennappan, R., Sakhare, D.T. (2023). Convergence of Communication Technologies with Internet of Things. In: Hemanth, J., Pelusi, D., Chen, J.IZ. (eds) Intelligent Cyber Physical Systems and Internet of Things. ICoICI 2022. Engineering Cyber-Physical Systems and Critical Infrastructures, vol 3. Springer, Cham. https://doi.org/10.1007/978-3-031-18497-0_48

[18]. Rajeswari, A.R., Kulothungan, K., Ganapathy, S. and Kannan, A., 2019. A trusted fuzzy based stable and secure routing algorithm for effective communication in mobile adhoc networks. Peer-to-Peer Networking and Applications, 12(5), pp.1076-1096.

[19]. Krishnan, R.S., Julie, E.G., Robinson, Y.H., Kumar, R., Son, L.H., Tuan, T.A. and Long, H.V., 2020. Modified zone based intrusion detection system for security enhancement in mobile ad hoc networks. Wireless Networks, 26(2), pp.1275-1289.

[20]. Bisen, D. and Sharma, S., 2018. An enhanced performance through agent-based secure approach for mobile ad hoc networks. International Journal of Electronics, 105(1), pp.116-136.

[21]. Vatambeti, R., Sanshi, S. and Krishna, D.P., 2021. An efficient clustering approach for optimized path selection and route maintenance in mobile ad hoc networks. Journal of Ambient Intelligence and Humanized Computing, pp.1-15.

[22]. Vatambeti, R., 2020. A novel wolf based trust accumulation approach for preventing the malicious activities in mobile ad hoc network. Wireless Personal Communications, 113(4), pp.2141-2166.

[23]. Xu, H., Si, H., Zhang, H., Zhang, L., Leng, Y., Wang, J. and Li, D., 2020. Trust-based probabilistic broadcast scheme for mobile ad hoc networks. IEEE Access, 8, pp.21380-21392.

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

Enhancing Microcomputer Edge Computing for Autonomous IoT Motion Control

S. Bangaru Kamatchi[1], Vaishali N. Agme[2], S.Premkumar[3], KDV Prasad[4], Dankan Gowda V[5], I Gugan[6]

[1]Assistant Professor, Department of Computer Science and Engineering, Sathyabama Institute of Science and Technology (Deemed to be University), Chennai, Tamilnadu, India.
e-mail: bangarukamatchi.cse@sathyabama.ac.in

[2]Assistant Professor, Department of Applied Science, Bharati Vidyapeeth College of Engineering, Navi Mumbai, Navi Mumbai, Maharashtra, India.
e-mail: vaishali.agme@bvcoenm.edu.in

[3]Assistant Professor, Department of Computer Science and Engineering, Galgotias University, Gautam Buddha Nagar, Greater Noida, Uttar Pradesh, India.
e-mail: s.premkumar@galgotiasuniversity.edu.in

[4]Assistant Professor (Research), Symbiosis Institute of Business Management, Hyderabad, Symbiosis International (Deemed University), Pune, India.
e-mail: kdv.prasad@sibmhyd.edu.in

[5]Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India.
e-mail: dankanies@gmail.com

[6]Assistant Professor, Department of CSE, Dr NGP Institute of Technology, Coimbatore, Tamilnadu, India.
e-mail: apgugan@gmail.com

Corresponding Author: Dankan Gowda V, e-mail: dankanies@gmail.com

Abstract— Devices microprocessors, microcontrollers, and Field Programmable Gate Arrays (FPGA) play the core rule at the IoT edge level and it should be right provisioned. For proper controller performance, control algorithms should be implemented near the actuator eliminating the delay effects. In the IoT domain, this means to implement the mentioned algorithm at the edge level and prior data transmitting. The efficient IoT-enabled motion control can be obtained by considering two main factors; the first factor is from the actuator design point of view and the second factor is from the controller performance point of view. Therefore, in this article, the two mentioned factors are treated concerning the microprocessor rule and importance as a core for proper prototype design and as the main platform to implement the control algorithms. A comparison of controller performance indices for both prototypes is done using previously distributed motion control schemes and newly developed schemes after tuning the respective schemes gains in an optimal manner. The scheme with better behavior of both prototypes are selected for the IoT integration process, this scheme ensures optimal edge computing for the distributed motion control, making the implementation of all control computation take place at the IoT-edge level. As a result, the dynamic pipeline stages (DPS) based prototype gives better controller performance indices for most strategies, less power consumption, and optimally utilized resources encouraging the use of the microprocessors with reconfigurable components at the IoT-edge level.

Keywords- Controller, Internet of Things, Edge Computing, Pipeline, Dynamic pipeline stages, Embedded system and Motor.

I. INTRODUCTION

This article highlights the importance of decentralized motion control as one of the main industrial uses of the rising state-of-the-art Internet of Everything. IoT basics and IoT definitions are introduced. A broad overview of the problem is presented, and the objectives of this paper are outlined. From cloud level the operator can specify individual channels for individual motors and individual channels for n-motors working together [1]. IoT enables the researchers to develop new systems which can be observed, analyzed and controlled remotely. With IoT enabled distributed motion control, both data visualization and data analysis for n-motors working in the industry can be carried out and abnormal cases can be detected early [2]. Microprocessors, microcontrollers, and field-programmable gate arrays (FPGAs) have evolved to the point where they constitute the backbone of Internet of Things (IoT) systems, beginning with nodes and progressing to gateways and servers [3]. Edge computing reduces delays, energy costs, and communication bottlenecks. Researchers in the Internet of Things sector have identified edge-level exploitation of reconfigurable microarchitecture as a potential research topic [4]. Taking into consideration the fact of using the microprocessors, microcontrollers, DSPs and FPGAs for control algorithms implementations, the emerging of IoT with the industry will require new modifications in the platforms architecture in order to be more suitable for the integration

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

process and to insure optimal behavior in real-time working conditions. Existing meanings of the term "Internet of Things" (IoT) vary. Many of these definitions have advanced as a result of the evolution of prior technologies and the effects of more recent ones [5]. The primary IoT techniques should be understood before working on the IoT architecture. In 1998, Ashton coined the phrase "Internet of Things" (IoT) to describe a network in which everyday things are linked by RFID and their data is utilised in online supply chains. Communication range, memory, and computing power are the constraints. Many requirements are reflected through those constraints and are needed to connect the internet to constrained devices by using TCP/IP protocols. However, a lot of meanings are missed through this definition. In 2009, CASAGRAS definition [6] is introduced. We were inspired to take on the task of transferring our focus from the home automation area, where most IoT research has been conducted, to the industrial control domain, where it has received much less attention.

Even IoT is considered as a new name of recent developed embedded systems, the device microprocessor which considered as brine for any embedded system has got less attention with respect to the research in the new technology IoT and more attention is given to the connectivity and protocol aspects [7]. The device microprocessor is intensively used at edge level with both sensors and actuators enabling them to be intelligent. However, the current fixed architecture of the device microprocessor is found to be unsuitable for utilization with IoT applications. Reconfigurable architecture makes the microprocessor more flexible with less power consumption and resource utilization as per application requirements. The focus of this work is to illustrate the suitability of using microarchitecture with reconfigurable (dynamic) pipeline stages against a one with fixed pipeline stages by using them as cores in developing two actuators for one of the important industrial control applications. That is a distributed motion control. The developed Distributed motion control scheme is to be enabled through the new technology IoT using the microprocessors in an optimal manner at edge level i.e. to have optimal performance indices through Edge Computing [7]. For this purpose, two different systems have been developed. The first system is a microprocessor (with fixed pipeline stages (FPS)) based PMDC Motor prototype, and the second system is a reconfigurable microarchitecture (with dynamic pipeline stages (DPS)) based PMDC Motor prototype. The aim of such developments is to show the possibility of overcoming the drawbacks with the current available fixed microprocessor architectures and obtaining the flexibility of avoiding the data loss in an optimal manner when high frequency measurement is required. The idea behind proposing the reconfigurable microarchitecture is due to expectation of having high speed machines in the industry and when high speed devices are connected with slow devices it is very important to see not to lose any data in transmission. If the input device speed is high then the pulses frequency coming from its sensor will be higher than the receiving element in processor, it is very difficult in synchronizing the data rate between them, so it is very important in considering some handshaking device between fast and slow devices [8]. For instance, a buffer or some delay system can be inserted between processor element and measurement system. But simple buffer and delay elements may increase the design complexity, slow down the process and shows impact on accuracy. The main reason of data loss caused when high-speed machines connected with low-speed devices is due to the frequency difference between them, and when the processor receiving element enters into its NRS status due to the processor “clear” signal effect, the coming pulses from the data source i.e. the high-speed machine will not be counted and therefore it will be lost [9]. The accuracy loss will occur only if a large number of delay elements are inserted without proper clocking protocol; however, the value of accuracy loss depends on the frequency value to be measured and the number of utilized delay elements and buffers. In the proposed micro architecture, dynamic pipeline is integrated to overcome these problems and its suitability for the industrial IoT utilization is confirmed.

II. LITERATURE REVIEW

The more current embedded systems are now referred to by their new term, Internet of Things (IoT). On the other hand, a fixed architectural design is seen as a disadvantage in the context of the internet of things, particularly in light of the growing number of things that are linked to the internet.

The device microprocessor acts as the brain of any embedded system. This section illustrate the previous research work related to the device microprocessor in the IoT domain [10]. In another hand and as per current work requirements, the distributed motion control application is selected as an important industrial topic for IoT integration process, so the research work related to both frequency measurement and distributed motion control is reviewed within current article, as proper frequency measurement results in proper motion control, along with illustrating the role of the computational hardware in PMDC motor control applications. Several application domains are offered computing potential by the IoT. The application domains include healthcare, logistics and transportation, smart environments, social and personal domains. Sensing includes data collection about case or event. The data can be temperature, motion or pressure, etc. The data acquired from the sensors should be manipulated to a more meaningful form. For example: in sensor fusion [11], the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

sensed data using different sensors are merged to produce a more robust and accurate quantitatively/qualitatively data. The algorithms that used for sensor fusion can occupy different levels of memory/compute intensity. Many nodes are connected together and the data travels through them. Wi-Fi and Bluetooth are examples of the communication technologies. Software defined ratio (SDR) is a communication system where filters, modems and other physical layer functions that are presented in hardware are implemented in software [12]. The most important characteristics of SDR are the flexibility in which the system enhancements do not require hardware updates. With SDR, both instruction memory and data footprints are small. However, its applications are compute intensive. Signal processing is the major function in image processing. The parameters of an input video stream or image are to be extracted or converted to a more meaningful form [13]. Face recognition, traffic sign recognition and automatic license plate recognition are IoT applications which use image processing. More computation capabilities are required with image processing as it includes matrix multiplications. In addition to that, large amount of data is to be stored either as input, intermediate or output, so this raise the memory size requirements. Compression ensures that the data is transmitted fastly, analyzed or/and retrieved by reducing the communication requirements. If the storage on the edge level is required then compression reduces the amount of required storage thus reducing memory size, and this will meet with fact of resource-constrained of the IoT devices [14]. There are two types of compression and those are lossy and lossless compression. With first type, i.e. lossy compression, the perceptibility of data in question is exploited, but the unnecessary data is removed. With second type, i.e. lossless compression, the redundant data are removed statistically for the aim to concisely represent the data. Second type is more memory intensive and more compute [15]. In order to keep the integrity of both the data and the device, security applications are required; those applications prevent unauthorized access to sensitive data. Malicious attacks affect the IoT devices which functioning in open environments. Data encryption ensures data confidentiality [16]. With data encryption, the encryption algorithm is utilized to create encrypted data which require decryption in order to be read / utilized. Encryption speed is related to the memory access latency for storage and data retrieval, so this is again forms both memory and compute intensively. When the IoT devices are utilized in unattended environments, the system should be capable to function properly if a failure of some their component accrues [17]. This is referred to as fault tolerance and it is important for quality of service (QoS) insurance. There are two types of fault tolerance and they are hardware-based and software based. First type i.e., hardware-based fault tolerance are a storage devices and use redundancy. Second type includes algorithms and applications that make operations like error detection and correction, cyclic-redundancy checks and memory scrubbing. However, in our work, we focus on configurable pipeline stages due to their effect on area and power consumption and thereby on performance. With configurable pipeline stages, both not ready sequence (NOT) and data propagation delay are avoided. We observe how the controller performance indexes values are enhanced when using configurable (dynamic) pipeline stages for speed measurements by comparing with their values in the case of non-configurable (fixed) pipeline stages. Even after linking the controller with IoT technology, the published data shows the enhanced performance of the controller behavior with configurable pipeline stages. With heterogeneous architectures, the same instruction set can be executed with different cores, but the cores have different performance and capabilities. The software of the concern application will determine the core that best ensures the optimization behavior based on evaluating the resources requirements [18]. There will be several cores with the heterogeneous architectures and those cores may be GPUs, CPUs, DSPs, etc Heterogeneous architectures provide a much reduced design space, but this requires more design time and efforts. However, if the number of applications is increased, then heterogeneous architectures may provide less optimization potential than adaptable architectures. The number and choice of the cores is a major challenge and scheduling of applications to the best core is another major challenge in designing the heterogeneous architectures for the IoT. The distributed heterogeneous architectures in the form of a network are an alternative to heterogeneous cores on one device. In this case, many research challenges are to be addressed in order to use distributed heterogeneous architectures. IoT devices are under more number of attacks due to their pervasiveness, mobility and connectedness [19]. IoT devices resources constraints make security a challenging as they will not have hardware support for enhance security features. In addition to resources constraints, IoT devices may generate sensitive information. Most of hardware security techniques that proposed for embedded systems can be used in IoT devices with runtime configurable security policies which adaptable to the requirements. Several control and instrumentation applications require accurate measurement of frequency in digital form. The device microprocessor plays the main rule in frequency digital measurement and there are several methods in which the microprocessor can be programmed to perform the task of frequency measurement [20]. The latency requirements at computational stages should be given significant importance to insure proper and trusted control algorithm implementation via

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

a specified hardware target. However, those requirements are greatly depended on the processor working frequency while performing the control algorithm program cycling and the operator should ensure a complete program scale on each cycle before every controller necessary action to satisfy the desired output. This goal can be achieved by suitably selecting the sampling time period in such a manner to allow the smooth and accurate controller algorithm implementation using specified hardware boards. This section presents the relevant work related to the device microprocessor in the IoT context and its optimization for the utilization with IoT especially at edge level. Reconfigurable microarchitecture is recommended for the components that can be configured within the run-time[21]. A review of the velocity measurement methods based on the device microprocessor is introduced within this article considering the fact of obtaining proper motion control through proper velocity measurement. The DMA method is found to be the best for frequency measurement but it has the problems of NRS and propagation delay which can be solved via pipeline system, but with fixed processor architecture the pipeline stages number is fixed which a drawback, and therefore this can be solved via dynamic (reconfigurable) microarchitecture.

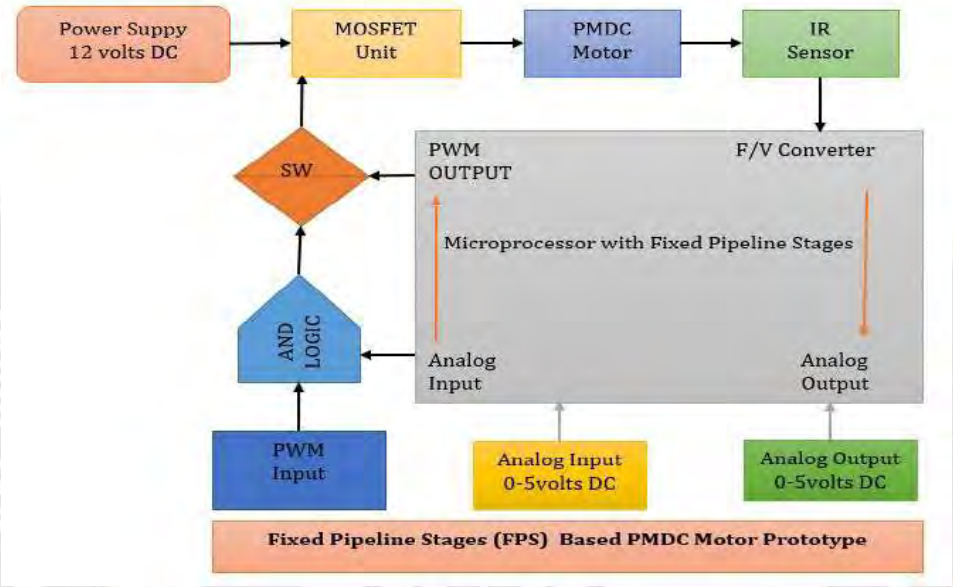


Figure1: Block schematic for the proposed PMDC motor prototype

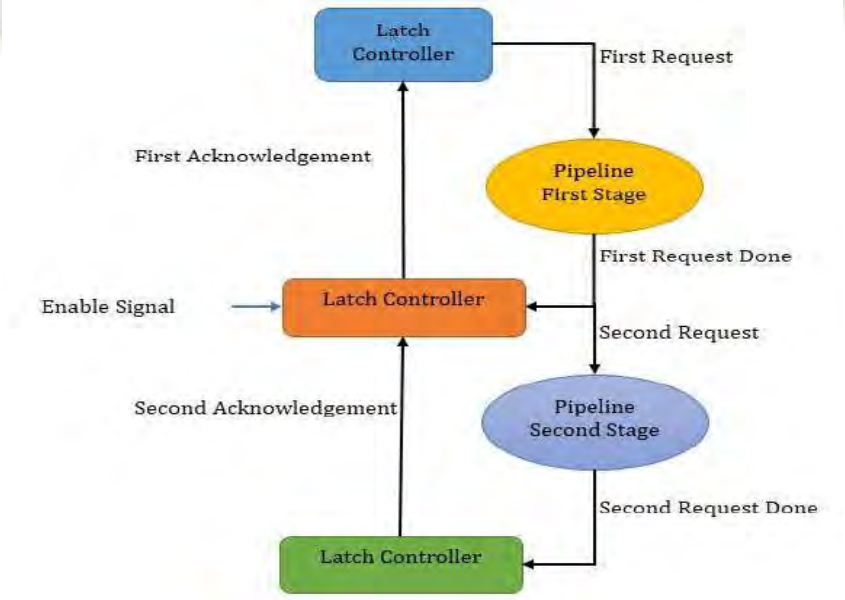


Figure.2: Two pipeline stages

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

III. PROPOSED SYSTEM

Digital electronic devices become much faster, smaller and cheaper due to two main reasons, first reason is the massive progress in semiconductor technology paced by planer processing and the second one is the Very-Large-Scale Integration (VLSI) process [22]. The applications of the microprocessors are extended to the industry due to the spectacular development semiconductor memory such as ROMs, PROMs and EPROMs [23]. In many industrial applications, microprocessors are presented [24]. One important industrial application is motor speed control with the ability of changing the speed within wide ranges.

The block diagram of the prototype PMDC motor design is shown in Figure.1. The front panel is the interface for the PMDC motor, via which the user may send commands and get feedback. A switch on the actual front panel lets the user choose between an analogue input with a range of 0-4.5 or a pulse width modulation (PWM) input. If the user chooses analogue input, a potentiometer on the physical front panel may be used to adjust the input voltage within the aforementioned range. Selecting an analogue input working scenario provides the microcontroller with the input voltage it needs to transform the analogue input into a pulse width modulation signal.

Dynamic pipeline stages are used to create a PMDC prototype based on reconfigurable microarchitecture. The Reconfigurable Microarchitecture is designed to be used for edge computing in the IoT. This microarchitecture is built utilising an FPGA board and a reconfigurable computer system. The PMDC motor's feedback circuit is utilised with reconfigurable microarchitecture to create a speed control prototype that can be successfully integrated with IoT technology [25]. The goal of designing a reconfigurable microarchitecture with the aid of a reprogrammable computing system based PMDC motor speed control prototype was to address the shortcomings of the fixed architecture (fixed pipeline stages as per our aim) microprocessor that are encountered in the previous designed PMDC motor speed control prototype. Both Not Ready Sequence (NRS) and data propagation latency of the DMA controller are eliminated with this produced prototype. To determine the best speed measuring technique at the IoT edge level, many approaches are explored utilising various pipeline architectures.

IV. EDGE LEVEL TECHNOLOGY FOR IOT

Some of the costs of using cloud computing include higher energy use, longer reaction times, and more demands on available communication bandwidth. Edge computing (carrying out the calculations near the data source) is favoured for the purpose of reducing these overheads, decreasing latency, preventing bottlenecks, and boosting performance. Because of the heterogeneity of IoT edge node platforms, it is challenging to create a common application based on the edge computing paradigm. Fog computing is presented as a virtual platform that offers networking services, computation, and memory between edge nodes and cloud-based data centres and is also known as mobile edge computing (MEC). The high throughput is a result of the pipeline layout's design. Like a factory line, a pipeline allows for the processing of fresh data even before the previous data has completed. In today's high-tech computers, the pipelining method is indispensable. It's put to use in a variety of contexts, including but not limited to parallel processing, cutting-edge computer design, high-velocity ALUs, and instantaneous I/O data retrieval. When the bulk of data processing is sequential, pipelines may increase system throughput. Power consumption rises proportionally with pipeline depth and falls with pipeline diameter. Transparent or opaque latches/registers allow pipeline stages to be merged or divided at runtime, resulting in dynamic pipeline stages. Think about the pipeline's first two phases, as seen in Figure 2.

In order to achieve precise signals, it is very important to get non-overlapping timing signals. The input signals are activated either at falling or rising edge of the clock signal. In this case, the clock edges imply that the main reason of inaccuracy is the timing jitter. For the aim of further suppressing the charge injection errors, the rising edge of pre-phase would be placed slightly before than that of past phases. It occurs when a pulse is readily available at Direct Memory Access Request (DRQ) of the Direct Memory Access (DMA) controller in Not Ready Sequence (NRS). NRS occurs in two cases: first case when the DMA is executing the past pulse and second case when the DMA is in inactive state. In this situation, second pulse should wait for DMA permission to enter into DRQ. Maximum efficient operation of the pipeline system cannot be obtained because of the internal propagation delay. Wave and mesochronous pipeline systems are used to manage the propagation delay [16]. However, it is difficult to estimate the required delay element. The clock pulses, in most digital systems, are used to control the execution of sequential functions. Non-overlapping clock pulses are difficult to be avoided, especially in two phase systems. Higher performance digital gates are achieved by minimizing the clock period. This population-based strategy finds the optimal answer by generating a swarm of particles that then flies about the study area. Each particle in the research space uses its own experience (Pbest) and the experiences of other particles in the swarm (Gbest) to update its flying direction and location using its fitness value and velocity. This process is carried out at the start of each iteration and continues until the termination requirements are fulfilled. Nonetheless, the PSO algorithm flowchart is shown in figure.3. The idea of developing the

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

new prototype comes to make the designed prototype more suitable for using with Internet of Things (IoT) technology as it is based on a processor with reconfigurable pipeline stages and this is recommended for the new generation of IoT microprocessors. A procedure of both System Identification and Optimal Speed Controller Design is performed for the dynamic pipeline stages (DPS) based PMDC motor. The merits of the new proposed PMDC motor, i.e. DPS_based one are its dynamic power consumption and better performance indexes, but the dis-merit is: it has less data matching percentage in comparison with FPS based PMDC in system identification procedure.

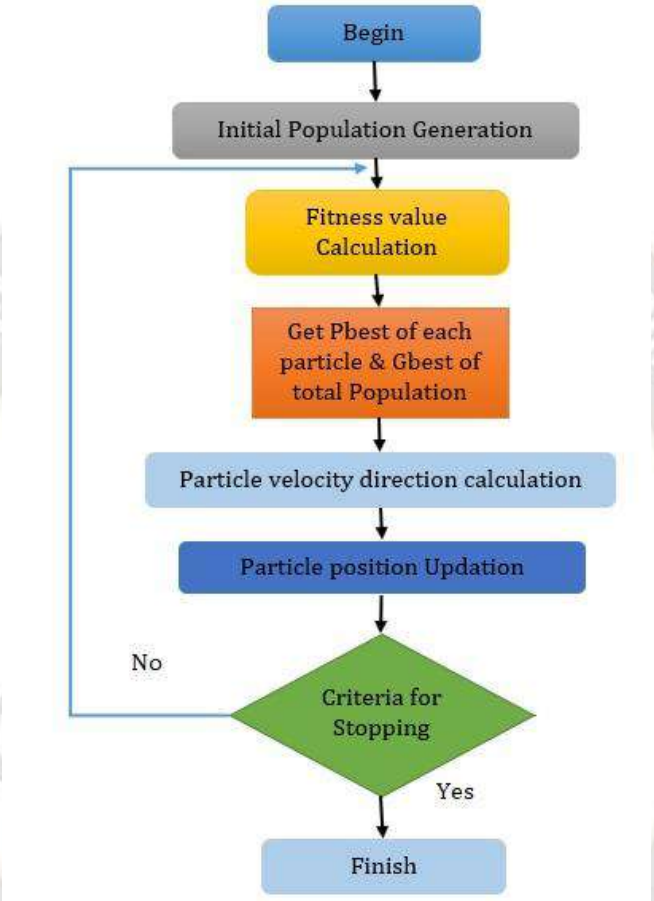


Figure.3: Schematic representation of the PSO algorithm

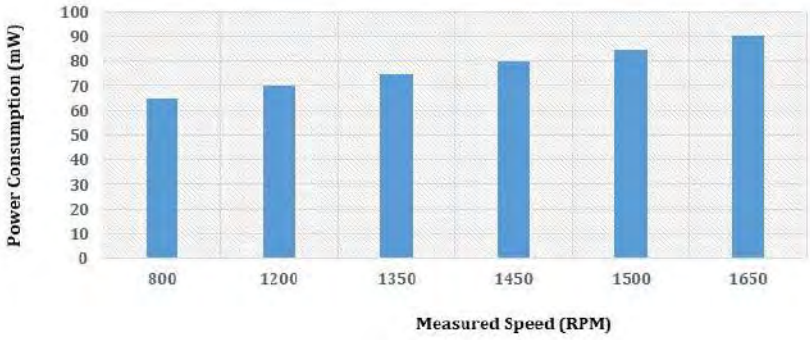


Figure.4: Examination of the prototype power consumption based on fixed microarchitecture

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

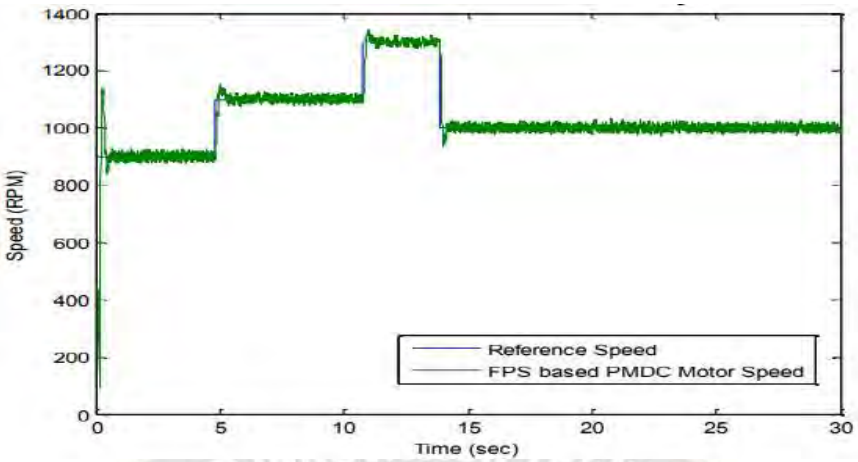


Figure.5: Set point tracking work Real Time test

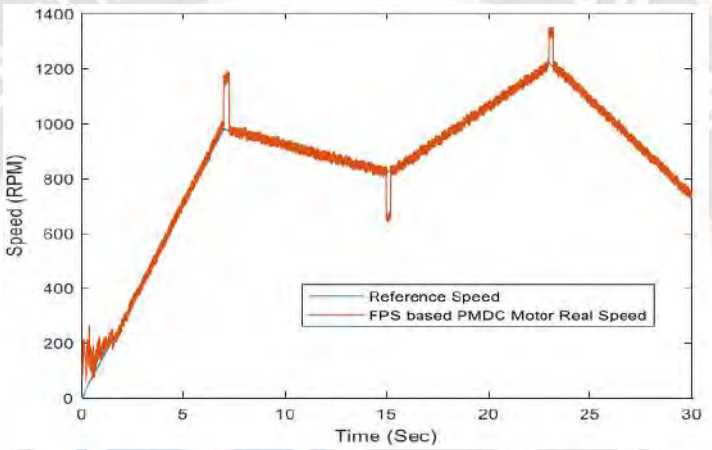


Figure.6: Prototype real time response for variable ramp input sequence

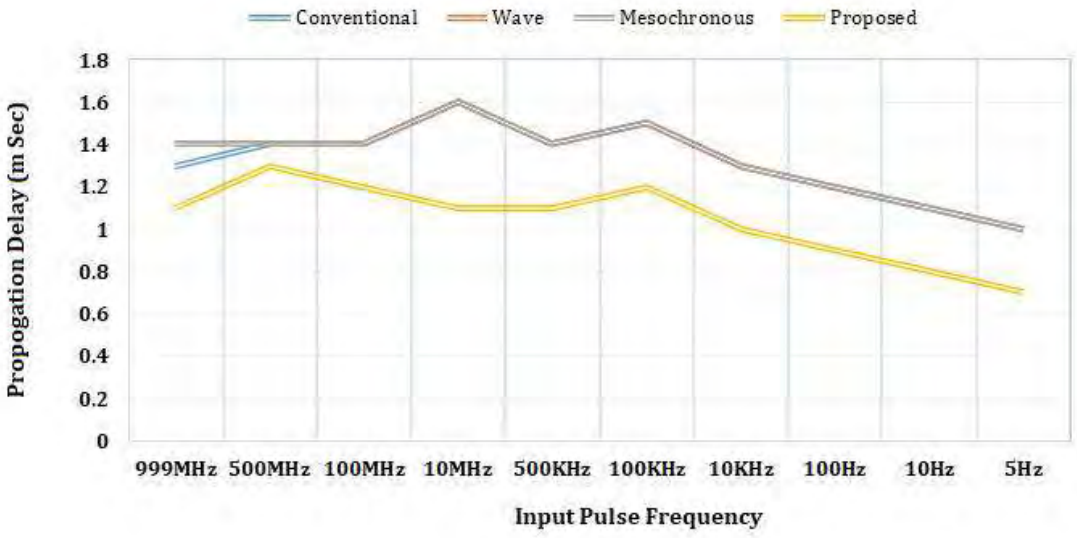


Figure.7: Three- Pipeline Stages

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

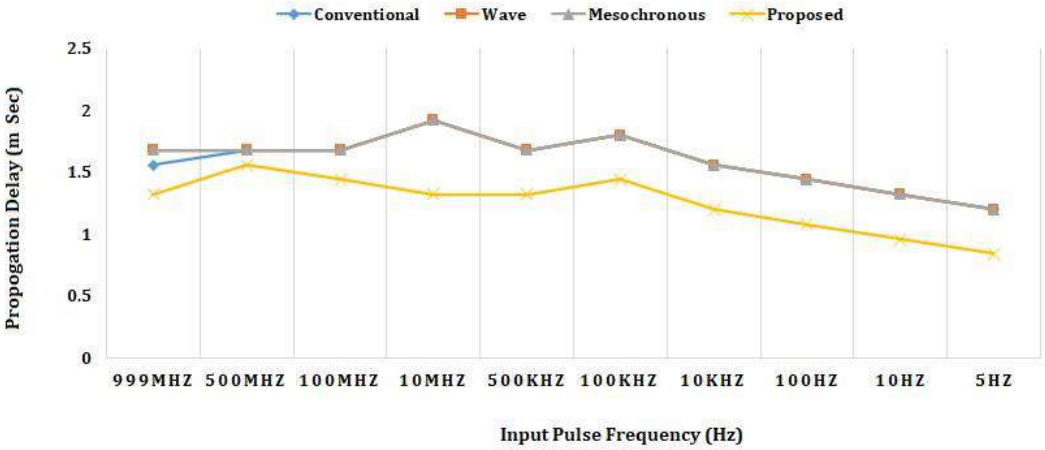


Figure.8: Four Pipeline Stages

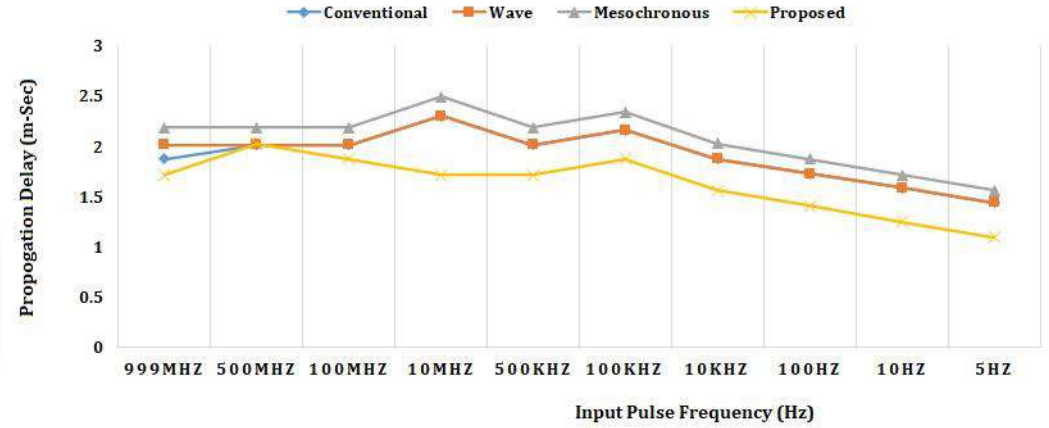


Figure.9: Five Pipeline Stages

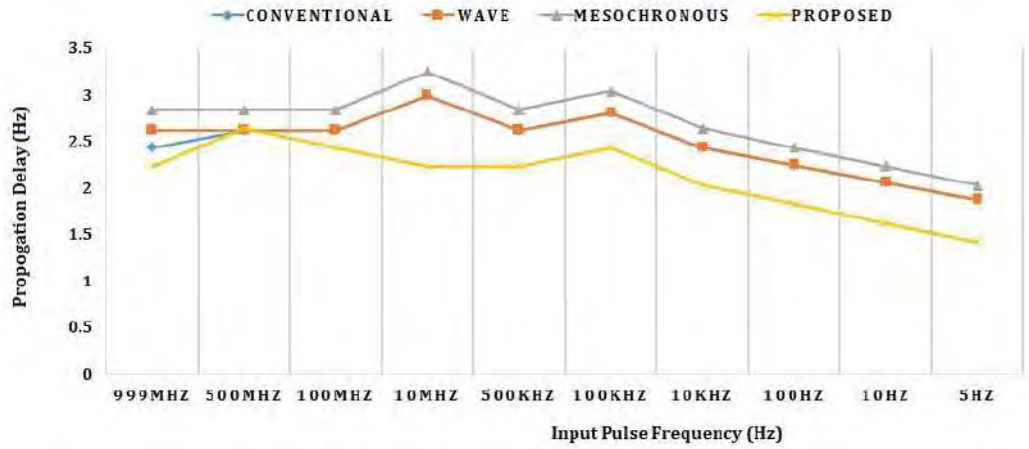


Figure.10: Six Pipeline Stages

International Journal on Recent and Innovation Trends in Computing and Communication
ISSN: 2321-8169 Volume: 11 Issue: 3
DOI: https://doi.org/10.17762/ijritcc.v11i3.6202
***Article Received**: 21 December 2022 **Revised**: 28 January 2023 **Accepted**: 15 February 2023*

V. RESULTS AND DISCUSSION

As it is mentioned and tested, current development makes the procedures of both system identification and controller design straightforward and easy due to its single input single output (SISO) configuration, but the fixed microarchitecture based microprocessor as a core for such development is considered as a drawback for its utilization with the-state-of-the-art IoT technology. When the user is going to use the prototype at the IoT edge-level as an actuator, the same amount of power consumption is required regardless the application speed requirements. However, the INA 219 current sensor is linked in contact with the feedback circuit of this fixed microarchitecture based developed prototype, and its output is collected by means of an Arduino mega board in order to examine the developed's power consumption. Power consumption is calculated by multiplying the computed current by the voltage difference between the power supply and the load, and a consistent value of 54 mW is recorded over a range of observed speeds (figure 4). In addition to that, only two pins i.e. the PWM input pin and the Analog output pin have been utilized from whole the microcontroller pins and the others are left idle which is considered as resources waste utilization. The fixed operating frequency of such fixed microarchitecture design limits its utilization when high frequency measurements are required due to the fixed number of its pipeline stages (FPS) and, in such case, the problems of the not ready sequence (NRS) due to the clear signal and the data propagation delay will take place. The reconfigurable microarchitecture-based prototype will ensure dynamic power consumption, proper resources utilization and overcome the NRS and data propagation delay problems by means of its dynamic pipeline stages (DPS) structure. In order to carry out the PMDC motor set point tracking job, the simulation test is being carried out. Figure 5 shows the response. In this particular illustration, the different paces is established at 900 revolutions per minute (RPM), and the flow rate conforms to this trajectory. Following that, at a real time interval of five seconds, a fresh trajectory consisting of 1100 RPM is implemented, and the controller makes it such that the motor speed follows this trajectory. After that, after ten seconds of real time, a fresh trajectory of 1,300 RPM is given, and the controller makes it such that the motor speed follows this trajectory. In the last step, which occurs after 15 seconds of actual time has passed, a new trajectory consisting of 1000 RPM is provided. In addition to that, the real-time reaction to a varied ramp input sequence was evaluated, and the results are shown in the figure.6. by carrying out a series of steps that are similar to those described in the section on simulation tests. The following statistics (Figure.7 to Figure.10) provide a graphical comparison of the propagation delays between the existing pipeline systems and the proposed pipeline system. A logical analysis is performed. When the logic is added, the suggested pipeline system has the lowest latency in propagation among the systems.

VI. CONCLUSION

The fundamental benefits of integrating the-state-of-the-art IoT with distributed motion control as an important industrial application along with IoT definitions, problem statements and article objectives are discussed in the first article. Literature review for enabling right-provisioned microprocessors and their optimizations for IoT along with microprocessor-based speed measurement and distributed motion control are discussed in article two. From the review, it is observed that: the microprocessors utilizations with IoT domain and especially for edge computing task should be designed in an optimal manner which can be obtained through several aspects and the reconfigurable architecture is one of them, in other side, the DMA method is found to be the best for speed measurement but it has the problems of NRS and data propagation delay which can be solved via pipeline system, so the fixed processor architecture is a drawback and therefore the selection of dynamic (reconfigurable) microarchitecture took place. The idea of developing the new prototype comes to make the designed prototype more suitable for using with Internet of Things (IoT) technology as it is based on a processor with reconfigurable pipeline stages and this is recommended for the new generation of IoT microprocessors. The merits of the new proposed PMDC motor, i.e. DPS based one are its dynamic power consumption and better performance indexes, but the demerit is: it has less data matching percentage in comparison with FPS based PMDC in system identification procedure.

REFERENCES

[1]. Lee, H., Paik, S. and Shin, Y. ―Pulse width allocation and clock skew scheduling: Optimizing sequential circuits based on pulsed latches‖. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, 29(3). March 2020, pp. 355-366.

[2]. Han, K., Park, J., Lee, J.W., Chung, J., Byun, E., Woo, C.J., Oh, S. and Abraham, J.A. ―Off-Chip Skew Measurement and Compensation Module (SMCM) Design for Built-Off Test Chip‖. Journal of Electronic Testing, 27(4). 2019, pp. 429-439.

[3]. Seetharaman, G., Venkataramani, B. and Lakshminarayanan, G. ―Automation techniques for implementation of hybrid wave-pipelined 2D DWT‖. Journal of Real-Time Image Processing, 3(3). 2018, pp.217-229.

[4]. insertion‖. Journal of VLSI signal processing systems for signal, image and video technology, 42(2). 2019, pp. 107-116.

[5]. Dailin, Z., Youping, C., Tom, K.C., Wu, A. and Zude, Z. ―Compensation scheme of position angle errors of permanent-magnet linear motors‖. IEEE transactions on magnetics, 43(10). 2017, pp.3868- 3871.

[6]. Anderson, R.G., Meyer, A.J., Valenzuela, M.A. and Lorenz, R.D. ―Web machine coordinated motion control via electronic line-shafting‖. IEEE Transactions on Industry Applications, 37(1). 2020, pp. 247-254.

[7]. Wang, S.D., Luo, X., Xu, S.J., Luo, Q.S., Han, B.L., Liang, G.H. and Jia, Y. ―A Planning Method for Multi-Axis Point-to-Point Synchronization Based on Time Constraints‖. IEEE Access, (8). 2020, pp. 85575-85604.

[8]. Jacob, S.S., Hussain, N., Chennappan, R., Sakhare, D.T. (2023). Convergence of Communication Technologies with Internet of Things. In: Hemanth, J., Pelusi, D., Chen, J.IZ. (eds) Intelligent Cyber Physical Systems and Internet of Things. ICoICI 2022. Engineering Cyber-Physical Systems and Critical Infrastructures, vol 3. Springer, Cham. https://doi.org/10.1007/978-3-031-18497-0_48.

[9]. Zhang, C.F., Xiao, Y.Y., He, J. and Yan, M. ―Improvement of electronic line-shafting control in multi-axis systems‖. International Journal of Automation and Computing, 15(4). 2018, pp. 474-481.

[10]. P, A. Sharma, S. Reddy P, P. S. Patwal, "Data Analytics and Cloud-Based Platform for Internet of Things Applications in Smart Cities," 2022 International Conference on Industry 4.0 Technology (I4Tech), 2022, pp. 1-6, doi: 10.1109/I4Tech55392.2022.9952780.

[11]. Fang, Y., Hu, J., Liu, W., Shao, Q., Qi, J. and Peng, Y. ―Smooth and time-optimal S-curve trajectory planning for automated robots and machines‖. Mechanism and Machine Theory, (137). 2019, pp. 127-153.

[12]. Swaminathan, K., Mukundrajan, R., Soundararajan, N. and Narayanan, V. ―Towards resilient microarchitectures: Datapath reliability enhancement using STT-MRAM‖. In 2019 IEEE Computer Society Annual Symposium on VLSI. IEEE. July 2019, pp. 236-241.

[13]. S. B. M, P. Pavankumar, N. K. Darwante, "Performance Monitoring and Dynamic Scaling Algorithm for Queue Based Internet of Things," *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2022, pp. 1-7, doi: 10.1109/ICSES55317.2022.9914108

[14]. Kanuparthi, A., Karri, R. and Addepalli, S. ―Hardware and embedded security in the context of internet of things‖. In Proceedings of the 2018 ACM workshop on Security, privacy & dependability for cyber vehicles. November 2018, pp. 61-64.

[15]. Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (eds) Recent Advances in Metrology . Lecture Notes in Electrical Engineering, vol 906. Springer, Singapore. https://doi.org/10.1007/978-981-19-2468-2_3.

[16]. Sharma, A. Sharma, K. S and M. R. Arun, "Priority Queueing Model-Based IoT Middleware for Load Balancing," *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2022, pp. 425-430, doi: 10.1109/ICICCS53718.2022.9788218.

[17]. Cloud Storage, Cloud Computing Security: From Single to Multi-Clouds, Journal of Physics: Conference Series, 1921 (1), pp. 012072. https://doi.org/10.1088/1742-6596/1921/1/012072.

[18]. G. Naveena Pai, M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.

[19]. K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," *Adv. Math. Sci. J.*, vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.

[20]. Guermaz, M.B., Bouzerara, L., Slimane, A., Belaroussi, M.T., Lehouidj, B. and Zirmi, R. ―High speed low power CMOS comparator for pipeline ADCs‖. In 2016 25th International Conference on Microelectronics. IEEE. May 2016, pp. 428-43.

[21]. Gaura, E.I., Brusey, J., Allen, M., Wilkins, R., Goldsmith, D. and Rednic, R. ―Edge mining the internet of things‖. IEEE Sensors Journal, 13(10). October 2018, pp.3816-3825.

[22]. Singh, S.P. and Maini, R. ―Comparison of data encryption algorithms‖. International Journal of Computer Science and Communication, 2(1). January 2018, pp.125-127.

[23]. Xiao, Y., Pang, Y., Ge, X. and Sun, J. ―Synchronous control for high-accuracy biaxial motion systems‖. Journal of Control Theory and Applications, 11(2). 2017, pp. 294-298.

[24]. Suryanarayana and José G. ―A Mesychronous pipelining scheme for high performance digital systems‖, Circuits and Systems I: Regular Papers, IEEE Transactions, 53(5). May 2016, pp. 1078 – 1088.

[25]. Anderson, R.G., Meyer, A.J., Valenzuela, M.A. and Lorenz, R.D. ―Web machine coordinated motion control via electronic line-shafting‖. IEEE Transactions on Industry Applications, 37(1). 2015, pp. 247-254.

SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348-8379/ https://doi.org/10.14445/23488379/IJEEE-V10I1P106

Volume 10 Issue 1, 63-73, January 2023

Original Article

Multimodal Biometric Identification system using Random Selection of Biometrics

Sampada Abhijit Dhole[1], Jayamala Kumar Patil[2], S. M. Jagdale[3], H. G. Govardhana Reddy[4], V. Dankan Gowda[5]

[1,3]*Department of Electronics and Telecommunication, Bharati Vidyapeeth's College of Engineering for Women, Pune, Maharashtra, India.*
[2]*Department of Electronics and Telecommunication Engineering, Bharati Vidyapeeth's College of Engineering Kolhapur, Maharashtra, India.*
[4]*Department of Mathematics, Alliance School of Applied Mathematics, Alliance University, Bangalore, Karnataka, India.*
[5]*Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India.*

[5]*Corresponding Author : dankanies@gmail.com*

Received: 25 November 2022 Revised: 04 January 2023 Accepted: 14 January 2023 Published: 29 January 2023

Abstract *- Biometric systems employ their biometric features to identify people. Identification systems that solely employ one biometric modality would not be able to meet the demands of demanding biometric applications in terms of performance, acceptance, and uniqueness. The majority of unimodal biometrics systems have problems with concentrated data noise, variances within and across classes, non-universality, etc. Multimodal biometric systems, which may establish identity from many sources of information, can bypass some of these restrictions. Identifying a person using multimodal biometric technology is more accurate and dependable. Early integration tactics are anticipated to perform better than late integration strategies. In this paper, feature-level fusion using the random selection of biometrics is presented. Block variance features and contourlet transform features are used to carry out the feature-level fusion. LDA is used to reduce the feature vector's dimensions. When compared to alternative integration approaches and their unimodal cousin, integrating the contourlet transform features of two independently determined biometric qualities delivers a consistent gain in performance accuracy. In this work, we use a random selection of biometric traits to guarantee the presence of a real human being at the time of data collection. Only fingerprints, palm prints, and faces will be included in the random selection.*

Keywords *- Hand geometry, Contourlet transform, Multimodal, Feature level fusion, Biometric.*

1. Introduction

Information transparency is a fundamental concern in the age of information technology. The information must be protected from unwanted access because its confidentiality and integrity are crucial. Security is the act of preventing unauthorized individuals from accessing sensitive information or priceless assets. Personal identification in various contexts (ATM cards, driver's licences, passports, citizen cards, mobile phones, voter ID cards, etc.) must be reliable and automated [1]. Furthermore, securing yourself and your belongings is essential. The traditional identification methods, including PINs and passwords, are unreliable because fraud is possible. Utilizing biometric identification provides a solution to this issue [2].

The term "biometric" refers to a person's physiological (such as fingerprints, the face, or the iris) and behavioral (such as speech) traits completely unique to him. Unlike traditional authorisation systems like smartcards, biometrics identifies a person based on who they are rather than what they have on them. You can never misplace, guess, or fake a biometric id. Regarding biometrics, the vast majority of real-world applications are single-modal [4]. When just one biometric identifier is employed, the process is said to be "unimodal." However, these systems have limitations that may be overcome by using additional sensors in a multimodal setup [5]. These limitations include noise in sensed data, intra-class variations, inter-class similarities, nonuniversality, and spoof assaults.

An attempt is made in this study to present a multimodal system that is flawless. Section 2 describes the materials and methods used to design the multimodal framework. Section 5 includes the results of the framed system, and finally, the conclusion highlights conclusive remarks of the resultant system.

2. Literature Survey

2.1. Multimodal Biometric System

Most existing biometric systems rely on what is known as a unimodal system, in which just a single biometric attribute is utilised to identify a person. Considering several different biometric characteristics, Multimodal biometric systems can circumvent some restrictions imposed by single-trait biometric methods. In theory, many biometric features make multimodal biometric systems more robust [7]. The lack of universality is solved by multimodal biometric systems, which include numerous characteristics that provide enough features for identification. As it is difficult for a fraudster to fake many biometric features at once, multimodal authentication systems effectively reduce the likelihood of identity theft.[3] In addition, a multimodal system may verify the presence of a 'live' user at the time of data collection by having the subject show a random subset of biometric features [8].

The following are five examples of integration possibilities for multimodal biometric systems: 1) a network of sensors whose results are averaged. There are four types of biometric identifiers: 2) multiple instances, 3) multiple samples, and 4) various biometric traits incorporate a number of biometric features (e.g., face and iris). Fifthly, several methods for recognising the same biometric—for instance, a fingerprint matcher that uses both texture and fine-grained features—are merged [6,9,10,18].

Four fusion stages are present in multimodal systems (sensor, feature, matching, and decision) [11]. It is generally accepted that the device and feature levels are where pre-mapping fusion takes place, while post-mapping fusion is assumed to take place at the corresponding score and determination levels. In premapping fusion, the biometric information is combined before classification; however, in postmapping fusion, the biometric information is patterned separately before being structured into a corresponding score/decision space and fused. Premapping fusion is contrasted with postmapping fusion, described in the following sentence. After each stage of a biometric system's processing, the quantity of data accessible for fusion decreases [13]. In this context, "fusion" at the classification stage refers to the process of merging feature sets that correlate to numerous modalities. When contrasted to the match score or the option, the feature set connects to more data about only the basic biometric information. This finding suggests that integrating at this stage might result in recognition results that are more accurate [14].

2.2. Feature Level Fusion

Features are derived from all biometric characteristics in feature-level fusion. By combining the retrieved features, a final output vector of increased dimension may be generated. Integration at the classification stage produces better identification results than score level or decision level fusion because the feature set contains more detail about the input information [16]. As shown in Fig.1, feature-level fusion involves pre-processing a feature set from two different sensors (such as a face and a palm print) before extracting features separately from each sensor to construct a trait vector [17]. Characteristics are then combined into a single new vector via composition. In this paper, to select biometric traits logistic regression method is used. So weights to the four biometric traits have been assigned depending upon the accuracy of biometric traits [19]. Weights assigned to fingerprint, palmprint, face and Hand geometry are 0.3, 0.25, 0.25 and 0.2, respectively. A higher value of weight assigns to the more accurate biometric trait.

Let's assume F_p and F_s are two feature vectors obtained by applying the contourlet transform to any two multimodal biometrics at random. Each trait's dimension will be denoted by f_p and f_s. Two-trait feature vectors are denoted by, $F_p = \{P_1, P_2, \ldots . . f_p\} and\ F_s = \{S_1, S_2, \ldots . . f_s\}$

In order to create a new different feature $F_c = [F_p \ldots . F_s]$, we join together two existing feature vectors. According to Fig.2, Fc has a size proportional to (fp + fs). Dimension 128 is achieved by concatenating the estimated correlation Characteristics of dimension 64 taken from each characteristic [20]. Fc is added to the database as a new template for finding a good match.

3. Feature Level Fusion using Block Variance Feature

The extracted features of 128 features are generated by concatenating the features of 64 different modalities into a single vector. The first modality contributes 64 features, and the second modality contributes 64 features [22]. The size of the picture being input is 256 by 256. The image is then cut into windows that are 32 pixels by 32 pixels, which results in eight blocks across and eight blocks down, as illustrated in Figure 3. Because of this, 64 features are taken from the first set of possibilities and 64 features are taken from the second set of paradigms [23]. The overall size of the feature vector is 128 when the characteristics of both modalities are concatenated together.

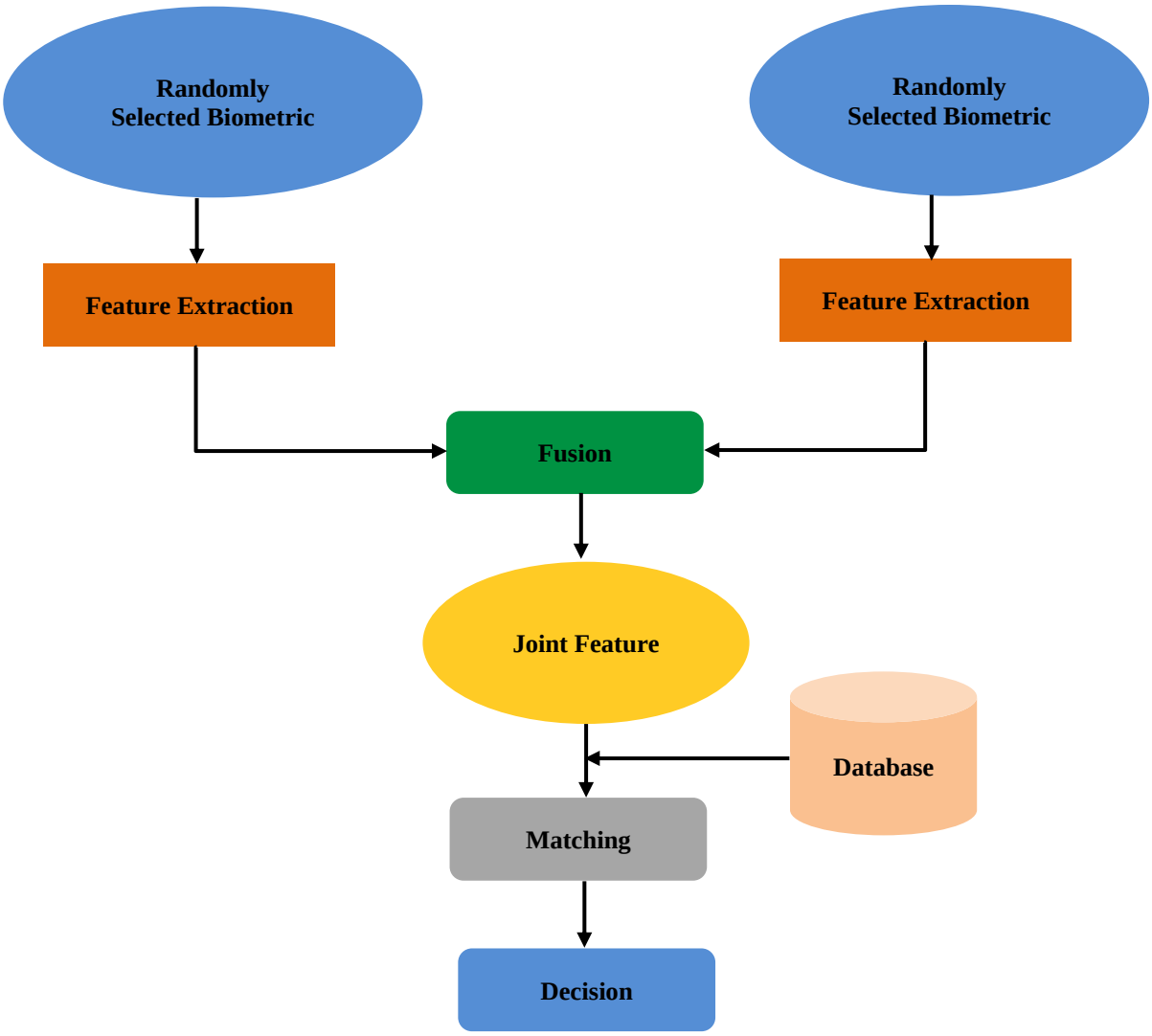


Fig. 1 Block Diagram of Feature Level fusion

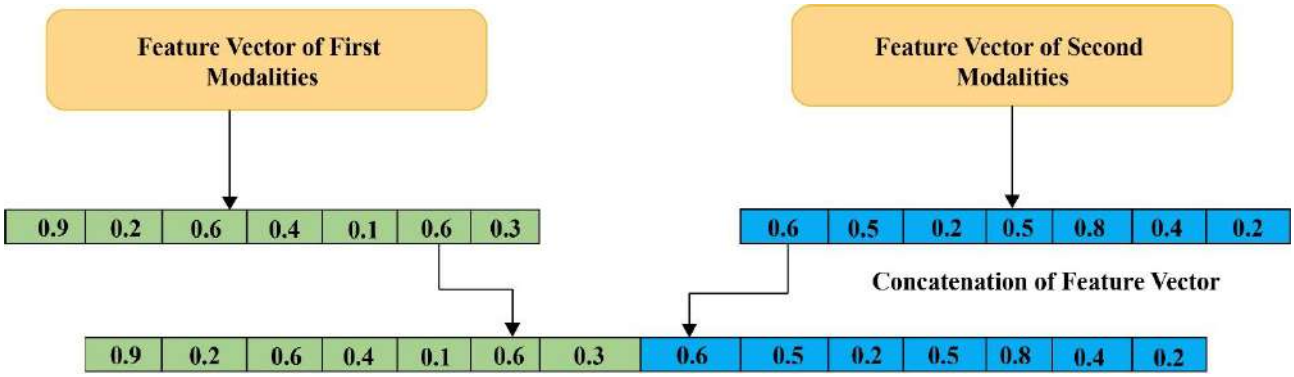


Fig. 2 Combining of two distinct characteristic vectors

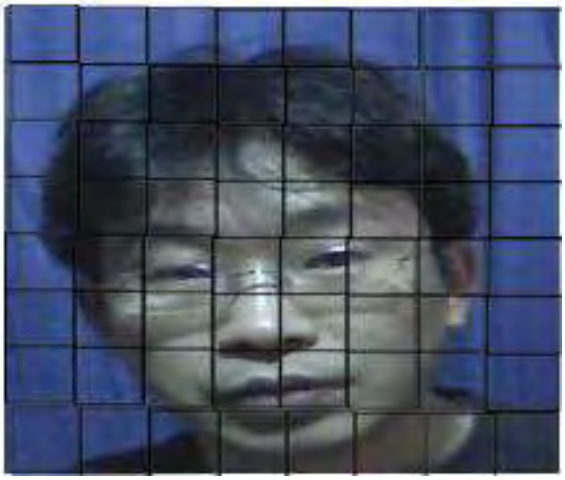


Fig. 3 Dividing Input Image into 64 blocks

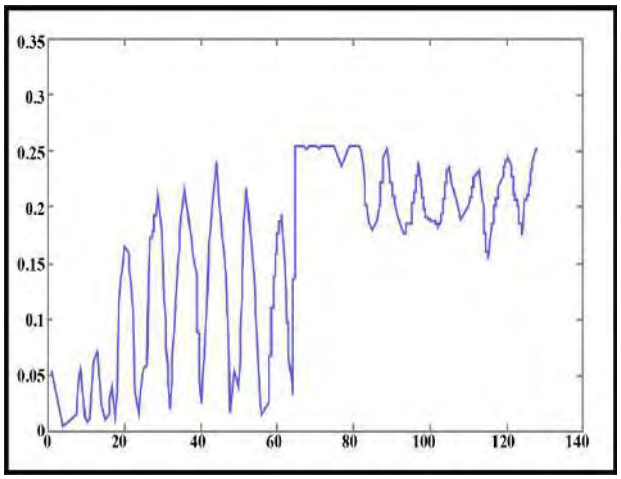


Fig.4. The combined face and fingerprint feature vector has 128 components.

Fig. 5 Dividing Input Image into 16 blocks

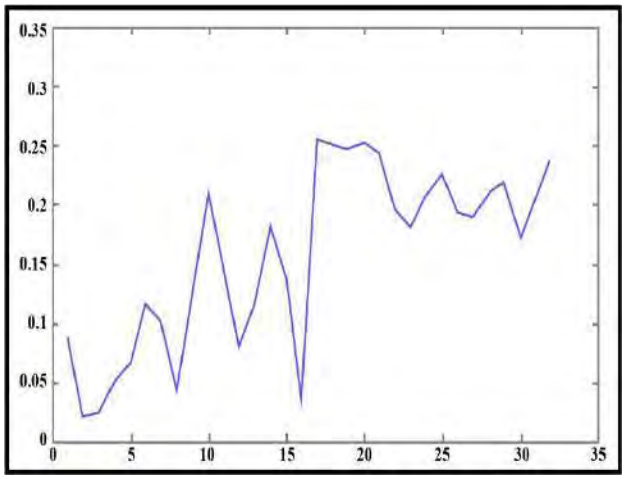


Fig. 6 Face and fingerprint features concatenated into a 32-feature vector

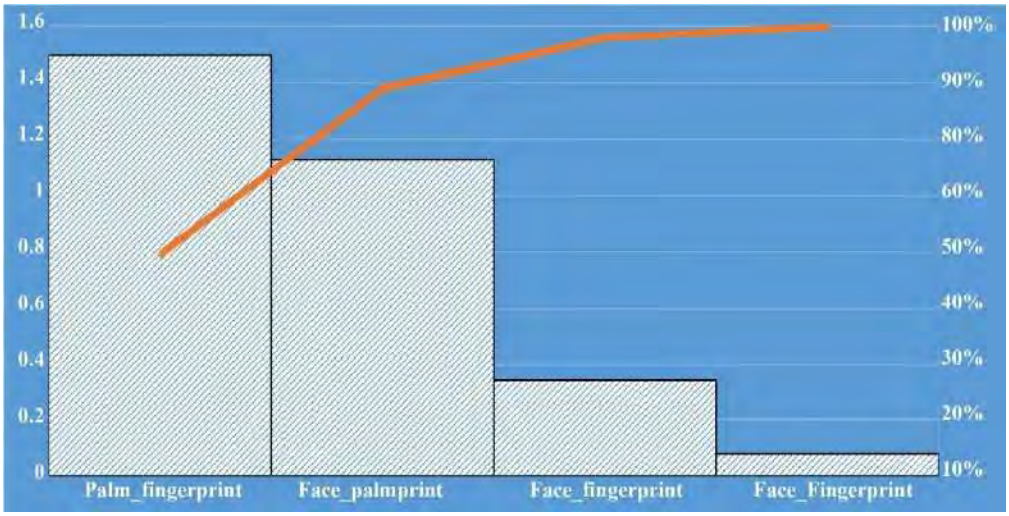


Fig. 7 Comparison Level Fusion with different concatenated biometric features modality for Casia database

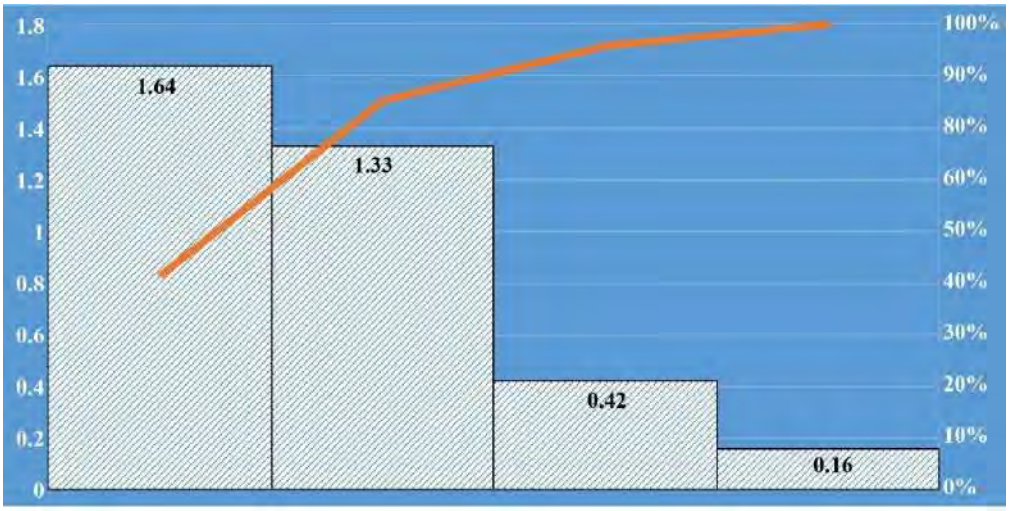


Fig. 8 Comparison level fusion of different with concatenated biometric features Modality for Local Database

Figure 4 depicts the integrated face and fingerprint feature vector. This figure exhibits a total of 128 features, consisting of the first 64 characteristics of the face biometric and the following 64 features of the fingerprint biometric, making the total number of features 128—a fusion of 32 feature vectors at the feature level [25].

Together, the 16 features from the first modality and the 16 features from the second modality make up the 32 features that make up the feature vector [26]. A 256x256 input picture is expected. As shown in Fig.5, the image is partitioned into 64x64-pixel windows, creating four blocks across and four down. It allows the extraction of 16 features from the first modality and 16 features from the second. Combining the feature vectors from the two modalities into one yield 32 features in total.

The first 16 characteristics of the face biometric and the following 16 features of the fingerprint biometric are shown concatenated in Fig.6.

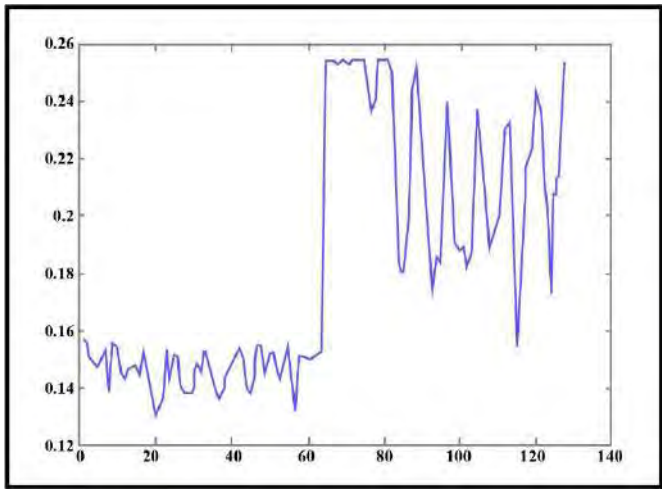


Fig. 9 Concatenated Feature Vector of Feature Fingerprint and Palmprint

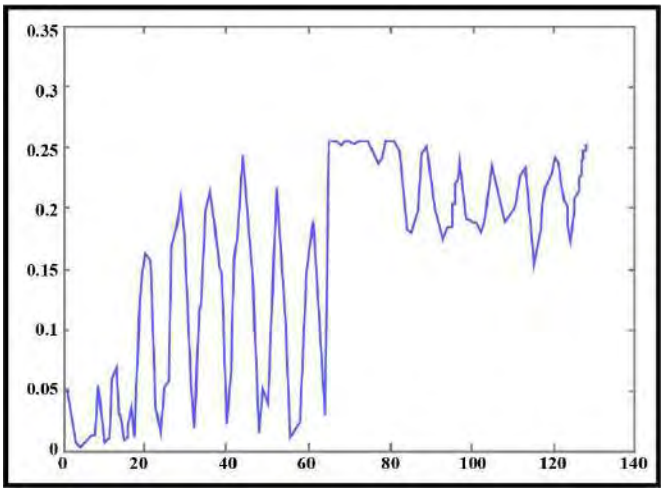


Fig. 10 Concatenated Feature Vector of Level Fusion of Feature Level Fusion of Face and Fingerprint

Therefore, the combined size of all the features in the feature vector is 32.

Figures 7 and 8 compare the results of fusing facial, palm, and fingerprint traits using the Casia and local databases. There was a 128-point correlation between palmprints and fingerprints, proving that the two could be fused into a single Accuracy is improved when combining many characteristics.

The integrated feature vector of a fingerprint and palmprint is shown in Fig. 9; it has 128 features in total, the first 64 of which come from the palmprint and the remaining 64 from the fingerprint. The magnitude of features is represented by the X-axis, while the size of the feature vector is shown by the Y-axis. Similarly, the joined face and fingerprint feature vector are shown in Fig. 10.

4. Feature Level Fusion Using Contourlet Transform Features

DWT does not provide directions other than horizontal, vertical, and diagonal. Curves with abrupt transitions are difficult for the wavelet to process. As the size of the feature vector grows, the computing time needed for feature extraction also increases, which slows down the retrieval speed [28]. The curvelet transform is useful for displaying discontinuities in curves but inevitably leads to the continuous domain. There is a transition from the discrete to the continuous domain in the contourlet transform. The contour more accurately represents the image's lines, edges, contours, and curves let transform than the wavelet or curve let transforms, thanks to the transform's directionality and anisotropy. Combining multiscale decomposition with directed decomposition results in the contourlet transform. The multiscale decomposition provided by the Laplacian pyramid allows the picture to be transformed into a coarse level and a collection of laplacian sub-bands. Critical downsampling at a directional stage allows for easy and adaptable sub-band partitioning of the overall frequency spectrum [29]. The basis function used in the contourlet expansion may have a variety of scales and orientations, and its aspect ratio can be varied as needed[31]. With such a wide variety of basic functions, the contourlet transform can accurately capture the smooth contours that are often the most prominent aspect of a picture. In Fig.11., we see the contourlet transform at level four, with orientations 0, 2, 3, and 4 representing the coarse, medium, and fine scales, respectively. The input picture is divided into two bands of equal width using a filter of order 'n' for each 'k' resolution [32]. If you want to see your picture as it was originally captured, start with the highest resolution setting (level 1). This setting has an input size of 256x256. The 128x128 pixel size is the next step up in resolution. Subsampling at levels 3 and 4 further reduces the input picture size to 64x64 and 32x32, respectively. Here, the contourlet transform performs to extract features, and then sum, max, and min are applied to those features to create a unified feature vector during the fusion stage [33]. Linear analytical thinking is used to reduce the feature vector's dimensionality.

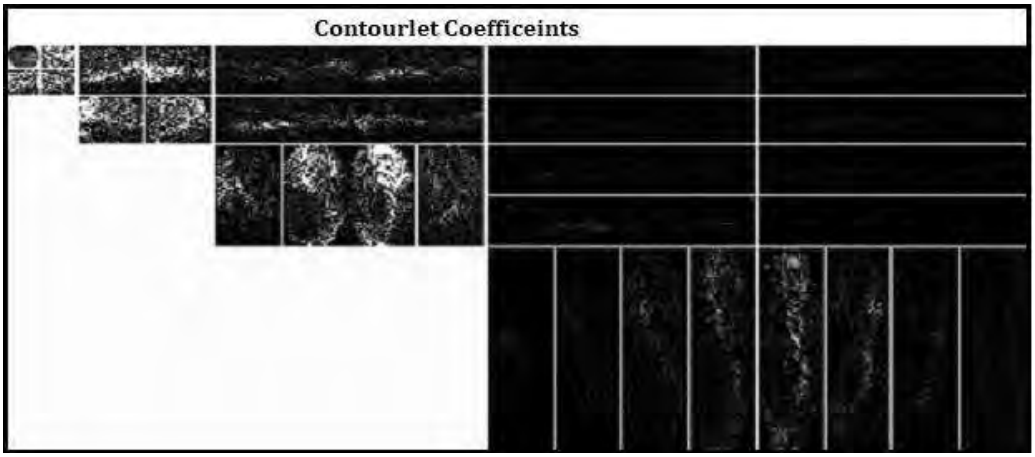


Fig. 11 Contourlet Transform decomposition for fingerprint image

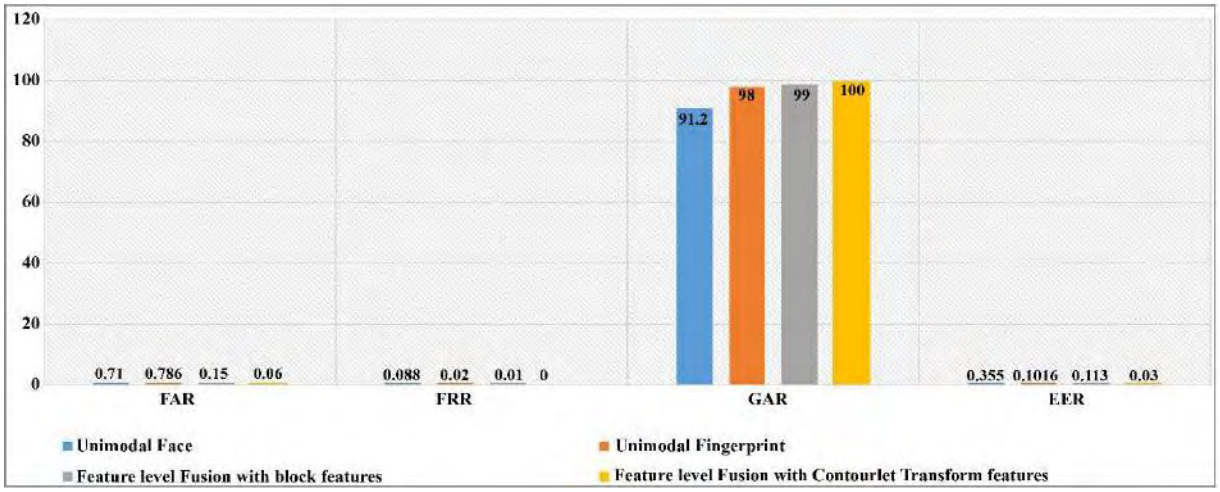


Fig. 12 Comparison of fingerprint and face multimodal identification system

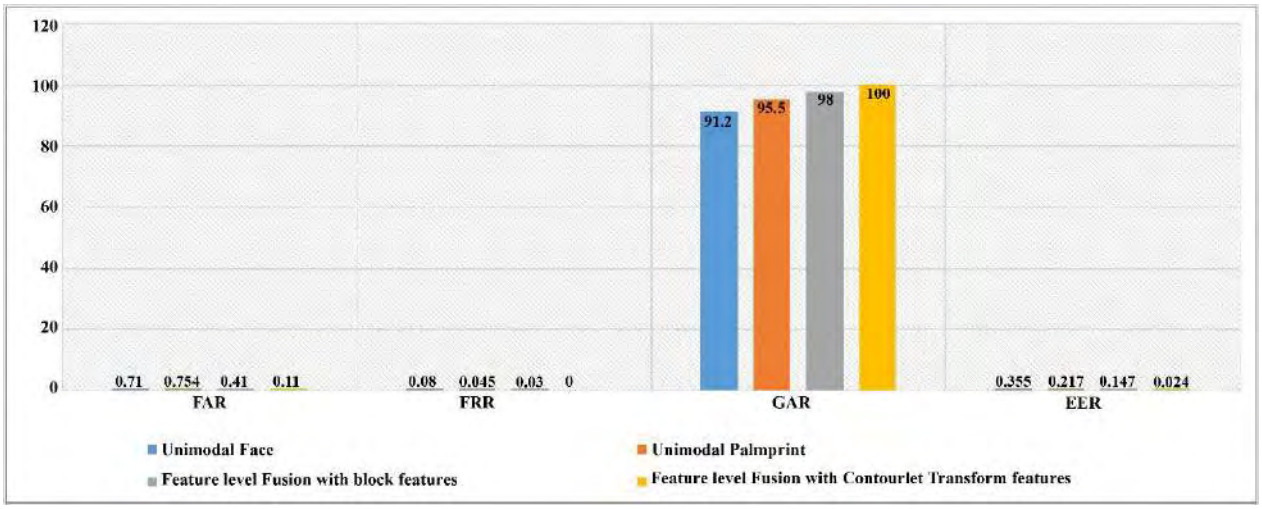


Fig. 13 Comparison of palmprint and face identification System

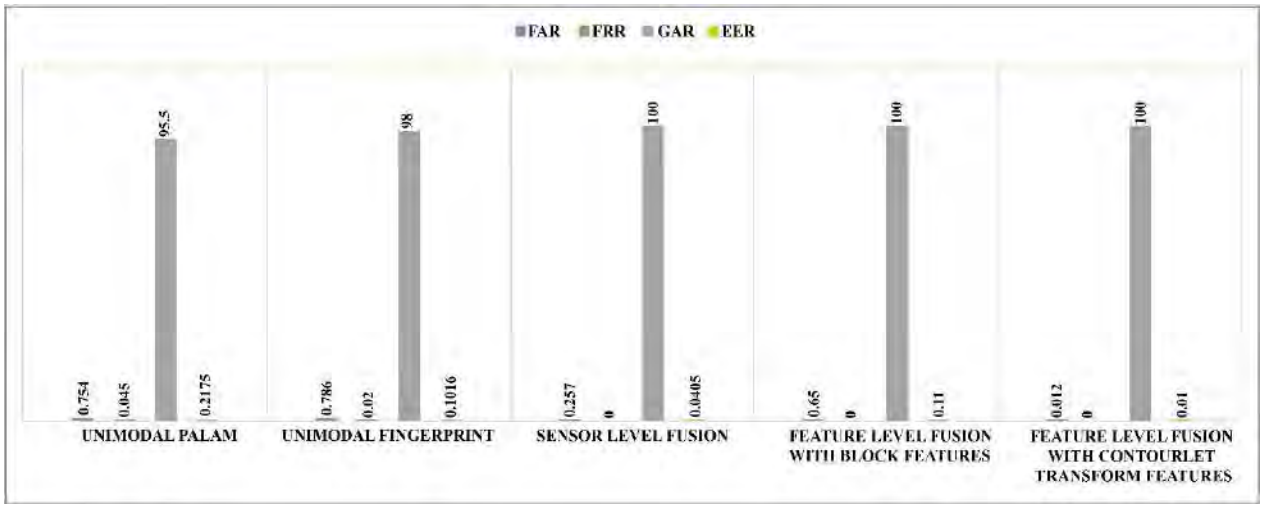


Fig. 14 Comparison of palmprint and fingerprint multimodal identification system

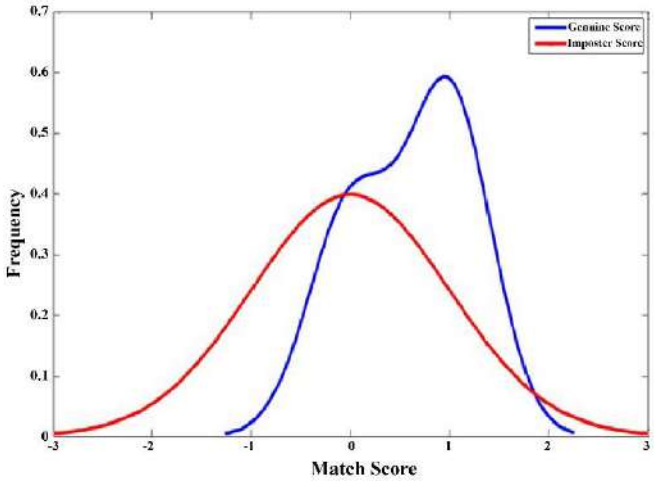


Fig. 15 Genuine Vs imposter score for the unimodal face identification system

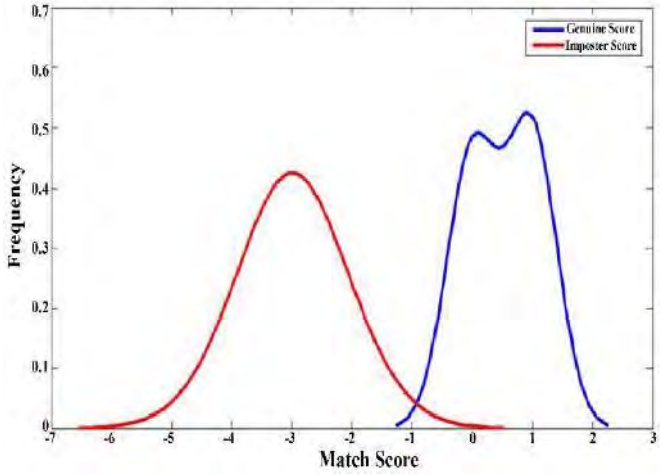


Fig. 16 Genuine Vs imposter score for the unimodal finger identification system

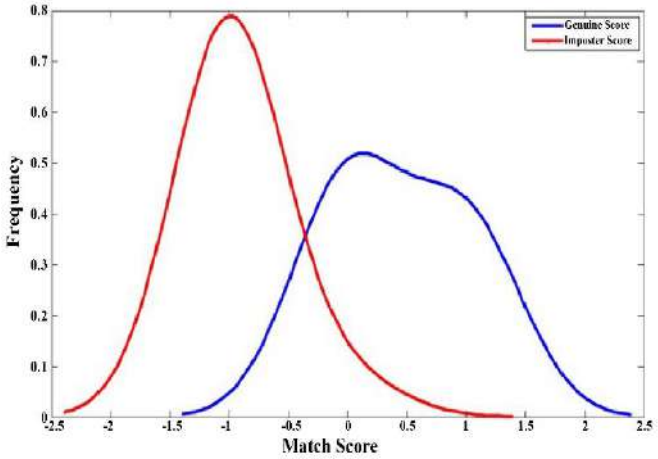


Fig. 17 Genuine Vs imposter score for unimodal Palm print identification system

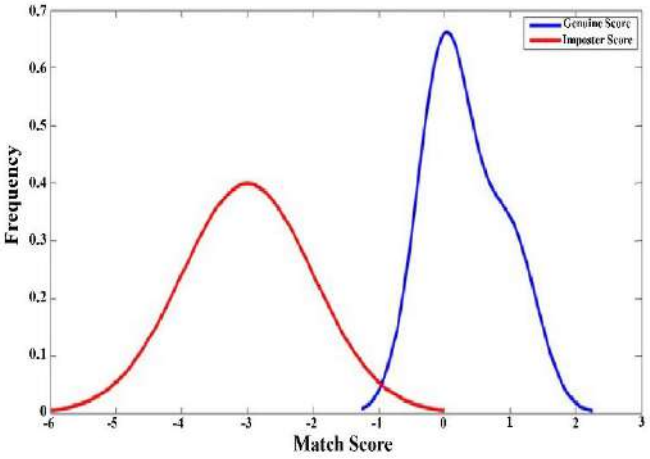


Fig. 18 Genuine Vs imposter score for Multimodal face and palm print identification system

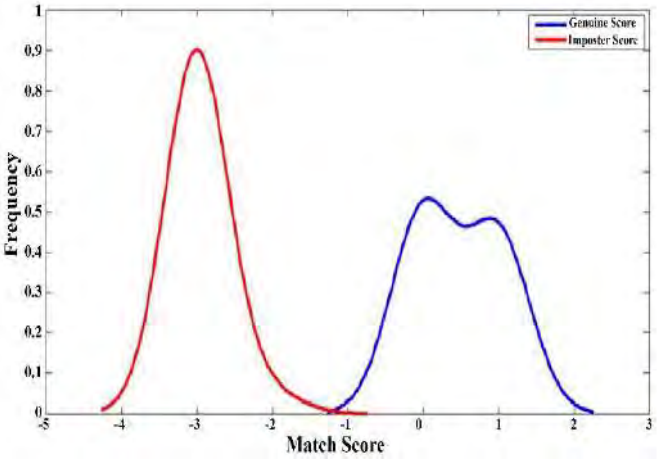


Fig. 19 Genuine Vs imposter score for Multimodal print and fingerprint Identification system

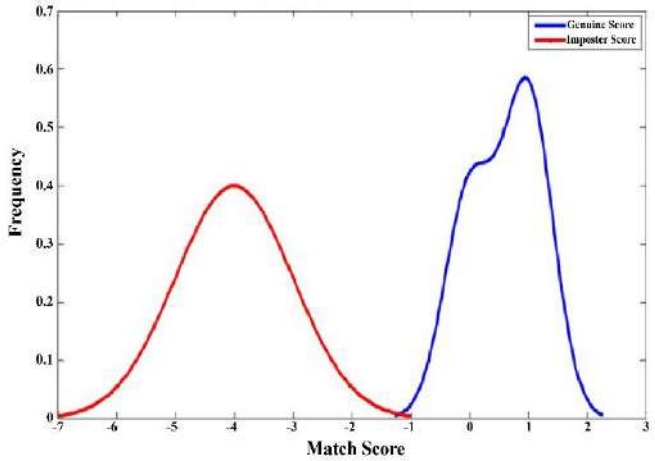


Fig. 20 Genuine Vs imposter score for Multimodal palm face and Fingerprint Identification system

Table 1. Feature vector size before and after applying the LDA algorithm

Biometric Modality	Feature vector size Before applying LDA	Feature vector size after applying LDA
Fingerprint_Palmprint	62	20
Fingerprint_Face	62	20
Palmprint_Face	62	20

4.1. Linear Discriminate Analysis (LDA)

LDA linearly combines unrelated information to generate the highest possible differences between the means between the target classes. In LDA, the goal is to use a scatter matrix to discover a transform where the characteristic clusters are most distinguishable following the linear transformation [34]. To achieve its goals, LDA attempts to maximise the scatter matrix's measure across classes while minimising the scatter matrix's measure within classes.

Table.1 shows how the feature vector size of feature level fusion is reduced after applying the LDA algorithm [12, 35]. A total of 31 features are present for fingerprint, palmprint and face after applying the contourlet transform. So after feature concatenation for feature level fusion total feature vector becomes 62(31+31). After applying the LDA feature vector size becomes 20.

5. Results and Discussion

The presented method's experimental evaluations are carried out on the standard database (CASIA database) and locally captured. The database used in this investigation consists of 1,000 photos across all accessible modalities, gathered from 200 unique participants and 5 representative samples of each participant. Light levels, age, and the natural bending of lines all have a role in the variability seen in palm print and fingerprint databases. Alterations in lighting, ageing, expression, stance, and camera angle were made to the face photos. The methods discussed in section 2 are applied, and results are analyzed with respect to FAR, FRR, GAR and EER for all modalities. Figure.12., Figre.13. and

Fgure.14. provide a comparison of face, fingerprint, and facepalm print modalities in unimodal and multimodal biometric technology, respectively. Compared to feature-level block characteristics and their unimodal counterpart, the performance of a feature-based fusion with a contourlet transform is much higher.

Fig. 15, 16, and 17 show the genuine imposter graphs for the unimodal biometric identification system. The overlap region of genuine and imposter scores for unimodal fingerprint identification is less than the unimodal face palm print identification system. So the accuracy of fingerprint identification system is more than Face, Palm print Recognition System

Figures 18., 19 and 20 show Genuine Vs imposter scores for multimodal biometric identification systems using feature-level fusion, which shows more accuracy than their unimodal counterparts. Overlap region of Genuine and imposter graphs for multimodal biometric identification using fingerprint and palm print is less compared to other multimodal biometric recognition systems. So the accuracy of multimodal biometric identification using fingerprint and palm print is high compared to other multimodal biometric identification systems.

This article demonstrates feature-level fusion using block variance features and contourlet transform features. Because the contourlet transform decomposes the picture into low and high-frequency curvelet coefficients at varying scales and angles, it is superior to block features for feature-level fusion. Contourlet transform can extract unique textural patterns from palmprint, face and fingerprint images. The presented feature level fusion using palmprint and fingerprint achieved the best recognition rates of 99.98%, and the ROC curve shows that the system achieved 100% GAR.

6. Conclusion

Multimodal biometric identification systems are utilised for person recognition since unimodal systems have problems with things like noise in sensed data, intra-class differences, inter-class similarities, non-universality, spoof assaults, etc. Multimodal systems, which **consider** both fingerprints and hand geometries, solve the problem of non-universality by providing better population coverage than single-modal systems. Because subjects deemed unfit for one biometric modality may utilise another, the reported findings show that it is exceedingly difficult for an attacker to fake numerous biometric features concurrently with a validly registered person. Compared to previous integration approaches and the unimodal analogue, integrating two separate, uncorrelated biometric features at the feature level delivers a consistent gain in performance accuracy. The accuracy of the Multimodal biometric system is shown to increase when biometric qualities are selected at random, as shown by the experiments. In addition, the system would know it communicates with a real person since biometric features would be picked randomly. Commercial, governmental, and even forensic uses may all benefit from this effort.

References

[1] Anil K. Jain, Arun ross, and Salil prabhakar, “An Introduction to Biometric Recognition,” *IEEE Transactions on Circuits and Systems for Video Technology*, vol. 14, no. 1, pp. 4 – 20, 2004. Crossref, https://doi.org/10.1109/TCSVT.2003.818349

[2] Ashish Mishra, “Multimodal Biometrics It Is: Need for Future Systems,” *International Journal of Computer Applications*, vol. 3, no. 4, pp. 28-33

[3] Ms.P.Jennifer, and Dr. A. Muthu Kumaravel, "An Iris Based Authentication System by Eye Localization,” *International Journal of Biotech Trends and Technology (IJBTT)*, vol. 3, no. 4, pp. 9-12, 2013.

[4] Karki Maya, and Sethuselvis, “Multimodal Biometrics at Feature Level Fusion Using Texture Features,” *International Journal of Biometrics and Bioinformatics (IJBB)*, vol. 7, no. 1, pp. 58-73, 2013.

[5] M. Arunkumar, and S. Valarmathy, “Palmprint and Face Based Multimodal Recognition Using PCO Dependent Feature Level Fusion,” *Journal of Theoretical and Applied Information Technology*, vol. 57, no. 3, pp. 337-346, 2013.

[6] Pooja G Nair, and Sneha R, "A Review: Facial Recognition Using Machine Learning," *International Journal of Recent Engineering Science* vol. 7, no. 3, pp. 85-89, 2020. Crossref, https://doi.org/10.14445/23497157/IJRES-V7I3P115

[7] Sneha A. Taksande, and Pushpanjali M. Chouragade, “Fusion Based Multimodal Biometrics Using Face and Palm Print at Various Levels,” *International Journal of Scientific & Engineering Research*, vol. 7, no. 2, pp. 106-109, 2016.

[8] Arun Jain, Surenderjangra, and Mehzabeenkaur, “Multimodal Biometric Identification Using Face and Palmprint,” *International Journal of Advanced Research in Computer Science and Software Engineering*, vol. 7, no. 5, pp. 11401144, 2017.

[9] Kalpana Chauhan, and Mrs. Mamta Yadav, "Automated Multi Face Identification,” *International Journal of Computer & Organization Trends*, vol. 7, no. 5, pp. 5-7, 2017.

[10] Praveen Kumarnayak, and Devesh narayan, “Multimodal Biometric Face and Fingerprint Recognition Using Adaptive Principal Component Analysis and Multilayer Perception,” *International Journal of Research in Computer and Communication Technology*, vol. 2, no. 6, pp. 313-321, 2013.

[11] Shiraz Anwar, and Surinder, “Multimodal Biometrics Identification Using Face and Palm-Print,” *International Journal of Advanced Research in Engineering Technology & Science*, vol. 4, no. 5, pp. 45-50, 2017.

[12] Raman Kumar, and Satnam Singh, "Face Recognition Using Principle Component Analysis for Biometric Security System,” *International Journal of Computer & Organization Trends*, vol. 3, no. 4, pp. 38-40, 2013.

[13] V Dankan Gowda, “Signal Analysis and Filtering Using One Dimensional Hilbert Transform,” *Journal of Physics: Conference Series*, vol. 1706, no. 1, 2020. Crossref, https://doi.org/10.1088/1742-6596/1706/1/012107

[14] Avinash Sharma et al., "Extraction of Fetal ECG Using ANFIS and the Undecimated-Wavelet Transform," *2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT)*, 2022, pp. 1-5, Crossref, https://doi.org/10.1109/GCAT55367.2022.9972078

[15] Utkarsh Chouhan, and H N Verma, "Blood Vessel Segmentation for IRIS in Unconstrained Environments Using Moment Method," *SSRG International Journal of Computer Science and Engineering*, vol. 5, no. 8, pp. 8-14, 2018. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V5I8P103

[16] Navdeep, and Surinder, “A Novel Multi-Model Biometric Fusion Approach Using Palm-Print & Face Biometric,” *International Journal of Latest Trends in Engineering and Technology*, vol. 8, no. 3, pp. 240-247, 2017.

[17] Varsha H. Patil et al., “An Efficient Secure Multimodal Biometric Fusion Using Palm Print and Face Image,” *International Journal of Applied Engineering Research*, vol. 11, no. 10, pp. 7147-7150, 2016.

[18] Deepak Singh, and Mr. Mohan Rao Mamdikar, "Identify a Person From Iris Pattern Using GLCM Features and Machine Learning Techniques," *SSRG International Journal of Computer Science and Engineering*, vol. 7, no. 9, pp. 25-29, 2020. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V7I9P105

[19] Ranjeet Suryawanshi et al., “Enhanced Diagnostic Methods for Identifying Anomalies in Imaging of Skin Lesions,” *International Journal of Electrical and Electronics Research (IJEER)*, vol. 10, no. 4, pp. 1077-1085, 2022. *Crossref*, https://doi.org/10.37391/IJEER.100452

[20] Avinash Sharma et al., “Vector Space Modelling-Based Intelligent Binary Image Encryption for Secure Communication,” *Journal of Discrete Mathematical Sciences and Cryptography*, vol. 25, no. 4, pp. 1157-1171, 2022. *Crossref*, https://doi.org/10.1080/09720529.2022.2075090

[21] Gatheejathul Kubra.J, and Rajesh.P, "Iris Recognition and Its Protection Overtone Using Cryptographic Hash Function," *SSRG International Journal of Computer Science and Engineering*, vol. 3, no. 5, pp. 1-9, 2016. *Crossref*, https://doi.org/10.14445/23488387/IJCSE-V3I5P101

[22] K. Tripathi, “A Comparative Study of Biometric Technologies With Reference to Human Interface,” *International Journal of Computer Applications*, vol. 14, no. 5, pp. 10–15, 2011. Crossref, https://doi.org/10.5120/1842-2493

[23] Aarti Hemant Tirmare et al., “A Morphological Change in Leaves-Based Image Processing Approach for Detecting Plant Diseases,” *International Journal of Electrical and Electronics Research (IJEER)*, vol. 10, no. 4, pp. 1013-1020, 2022. *Crossref*, https://doi.org/10.37391/IJEER.100443.

[24] Opeyemi Oyelesi, and Akingbade Kayode Francis, "Face Recognition for Access Control Using PCA Algorithm," *SSRG International Journal of VLSI & Signal Processing*, vol. 4, no. 2, pp. 22-27, 2017. Crossref, https://doi.org/10.14445/23942584/IJVSP-V4I3P105

[25] R. Gagan, and S. Lalitha, “Elliptical Sector Based DCT Feature Extraction for Iris Recognition,” *2015 IEEE International Conference on Electrical, Computer and Communication Technologies (ICECCT)*, pp. 1–5, 2020. *Crossref*, https://doi.org/10.1109/ICECCT.2015.7226026

[26] Ajay. P et al., "Intelligent Breast Abnormality Framework for Detection and Evaluation of Breast Abnormal Parameters," *2022 International Conference on Edge Computing and Applications (ICECAA)*, 2022, pp. 1503-1508, *Crossref*, https://doi.org/10.1109/ICECAA55415.2022.9936206.

[27] Anamika Baradiya, and Vinay Jain, "Speech and Speaker Recognition Technology Using MFCC and SVM," *SSRG International Journal of Electronics and Communication Engineering*, vol. 2, no. 5, pp. 6-9, 2015. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V2I5P105

[28] R. S. Gejji et al., “Understanding the Subject-Specific Effects of Pupil Dilation on Iris Recognition in the NIR Spectrum,” *2015 IEEE International Symposium on Technologies for Homeland Security (HST)*, pp. 1–6, 2015. *Crossref*, https://doi.org/10.1109/THS.2015.7225317

[29] DankanGowda V et al., “A Novel Method of Data Compression Using ROI for Biomedical 2D Images,” *Measurement: Sensors*, vol. 24, 2022, *Crossref*, https://doi.org/10.1016/j.measen.2022.100439

[30] Anjali Soni, and Prashant Jain, "Iris Recognition Using Four Level HAAR Wavelet Transform: A Literature Review," *SSRG International Journal of Electronics and Communication Engineering*, vol. 3, no. 6, pp. 14-18, 2016. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V3I6P106

[31] Madhavi Gudavalli et al., “Multimodal Biometrics-Sources, Architecture and Fusion Techniques: An Overview,” *2012 International Symposium on Biometrics and Security Technologies*, pp. 27–34, 2012. *Crossref*, https://doi.org/10.1109/ISBAST.2012.24

[32] Fabio Calefato et al., "Mobile Speech Translation for Multilingual Requirements Meetings: A Preliminary Study," *2014 IEEE 9th International Conference on Global Software Engineering*, pp. 145–152, 2014. *Crossref*, https://doi.org/10.1109/ICGSE.2014.10

[33] Zeng Wei et al., "A New Inertial Sensorbased Gait Recognition Method via Deterministic Learning," *2015 34th Chinese Control Conference (CCC)*, pp. 3908–3913, 2019. *Crossref*, https://doi.org/10.1109/ChiCC.2015.7260243

[34] Mohamad El-Abed et al., "A Study of Users' Acceptance Satisfaction Biometric Systems," *44th Annual 2010 IEEE International Carnahan Conference on Security Technology*, 2010. *Crossref*, https://doi.org/10.1109/CCST.2010.5678678

[35] Saiyed Umer, Bibhas Chandra Dhara, and Bhabatosh Chanda, "Face Recognition Using Fusion of Feature Learning Techniques," *Measurement*, vol. 146, pp. 43–54, 2019. *Crossref*, https://doi.org/10.1016/j.measurement.2019.06.008

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

Performance Analysis of Energy Efficiency and Security Solutions of Internet of Things Protocols

Manjunath Itagi[1], Dankan Gowda V[2*], KDV Prasad[3], Pullela SVVSR Kumar[4], Shekhar R[5] and B. Ashreetha[6]

[1]*Associate Professor, Department of Civil Engineering, Nagarjuna college of Engineering and Technology, Devanahalli, Bangalore, Karnataka, India, drmanjunath.itagi@ncetmail.com*
[2]*Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India, dankan.v@bmsit.in*
[3]*Assistant Professor (Research), Symbiosis Institute of Business Management, Hyderabad, Symbiosis International (Deemed University), Pune, India, Kdv.prasad@sibmhyd.edu.in*
[4]*Aditya College of Engineering, Surampalem, Andhra Pradesh, India, pullelark@yahoo.com*
[5]*Professor, Department of Computer science and Engineering, Alliance University, Bangalore, Karnataka, India, shekhar.r@alliance.edu.in*
[6]*Assistant Professor, Department of Electronics & Communication Engineering College, School of Engineering & Technology, Mohan Babu University (Erst while Sree Vidyanikethan Engineering College), Tirupati, Andhra Pradesh, India ashreetha.b@vidyanikethan.edu*

*__Correspondence:__ Dankan Gowda V; dankan.v@bmsit.in

ABSTRACT- The scientific and business communities are showing considerable interest in wireless sensor networks (WSN). The availability of low-cost, small-scale components like CPUs, radios, and sensors, which are often combined into a single chip, is crucial. Parallel to the evolution of WSNs, the concepts of the IoT have been evolving in recent years. Wireless communication technologies may play a significant role in the implementation of IoT, despite the fact that IoT does not need or require any particular technology for communication. WSN assisted IoT networks can drive several applications in many industries. The proposed research explores the possibility of enhancing energy efficiency in WSN-assisted IoTN by balancing various challenging sensor network performance metrics. The base station's current placement inside the sensing field is predetermined by the preexisting routing algorithms. Our study examines the impact of base station placement outside and within the prescribed sensing domains on energy consumption and network longevity. In addition, methods for transferring data from the distributed source sensor to the base station while minimizing energy consumption are investigated. In this preliminary study, we focus on developing an algorithm for WSN-Assisted IoTN that can balance network factors such as hop count, communication distance, and residual energy. To further optimize the routing route between local cluster heads and the base station, a novel network architecture is built based on the Ant-optimization model, which uses centroid routing to balance energy consumption among local clusters. An open-source Network Simulator (NS-3) is used to model the behaviour of the proposed routing protocols and compare them to comparable existing network protocols. All of the suggested protocols have the same fundamentals for creating networks, however they vary in terms of routing, optimization, and performance depending on the development effort under consideration.

Keywords: Internet of Things, Security, Protocols, Efficiency, Simulator and Optimization.

ARTICLE INFORMATION

Author(s): Manjunath Itagi, Dankan Gowda V, KDV Prasad, Pullela SVVSR Kumar, Shekhar R and B. Ashreetha;

Received: 25/02/2023; **Accepted**: 31/05/2023; **Published**: 26/06/2023;
e-ISSN: 2347-470X;
Paper Id: IJEER 2502-11;
Citation: 10.37391/IJEER.110226
Webpage-link:
https://ijeer.forexjournal.co.in/archive/volume-11/ijeer-110226.html

1. INTRODUCTION

Companies in a wide range of sectors are increasingly using IoT, a network of linked devices, for a variety of uses. In view of digital transformation of businesses, the IoT adoption has become essential for human and societal activities. To automate and collect data, many IoT devices are being deployed across the world. Many breakthrough innovations are achieved and supported by IoT in modern world. IoT technologies are being used to mitigate global warming, saving water and increasing yield in smart farming.

IoT is always a booming concept as its information processing capability became more reasonable due to its high computation power and cheap storage price [1]. As the available bandwidth of the network increased rapidly, sensors evolved to be smaller, affordable and more accurate. However, unlocking full potential of IoT generates key issues such as lowering the resolution complexity, solving security concerns and fighting the communication flaws in diverse environments.

Perception, the network, and applications are the three tiers that make up the Internet of Things. The perception layer consists of

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

a collection of devices that can see, sense, collect data, and share it with one another through the internet. This layer contains sensors [2]. The information obtained from the network layer is then processed by the application layer. To facilitate the creation of intelligent and context-aware software, the Internet of Things relies on WSNs, which function as "cells" to gather and disperse data. These sensor network devices are true enablers in IoT with regards to metrics like longevity, energy efficiency, reduced cost, and interface to resources since they use a variety of power sources and keep it running for a long time.

The "base station" in most Internet of Things designs is a collection of sensors that receives data from the many "things." The data will be sent to remote servers through the base station [3]. While there are Internet of Things (IoT) designs that use smartphones and other pervasive computing devices as "base stations" or "sensors," the emphasis of this study will be on using such devices as "sensors" in a simulated environment [4]. For enterprises and consumers alike to gain an edge via greater productivity and fewer overhead expenses, the data amassed by a sensor network must be made readily available to them in the form of actionable insight.

Setting up a 100-node network simulation in NS-3 (Network Simulator 3) involves creating a network topology with 100 nodes and configuring their communication and behavior. NS-3 is a discrete-event network simulator that provides a wide range of tools and modules to simulate network protocols, applications, and scenarios. Although NS-3 does not have specific tools dedicated to setting up a 100-node network, it provides a rich set of functionalities to facilitate the process.

In order to provide operational efficiency at lower costs, sensors conducting IoT operations are essential. The base station has superior processing capability, more memory, and an endless supply of electricity so that it can operate continuously [5]. The base station can collect, store, visualise, and evaluate the data sensed by the sensor nodes. The detected data may be sent to a distant server via the Internet, and the base station offers a Graphical User Interface (GUI) for communicating with user agents [6]. This information will be sent from the base station to the approved users through the internet. Access to sensing field data through web pages anywhere in the globe may improve data analysis beyond what is possible with a single base station [7]. *Figure 1* depicts a basic IoT setup that makes use of WSNs to some degree.

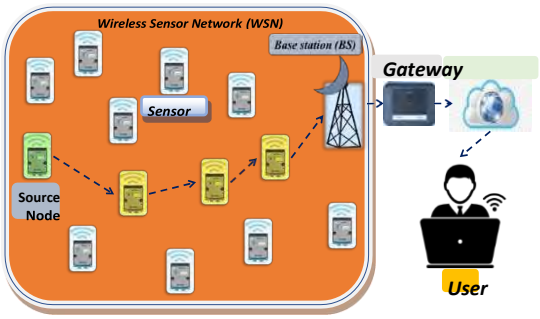


Figure 1: A basic WSN-assisted Internet-of-Things setting

The sensor network's "heartbeats" may be regularly sent from the base station to the cloud servers to update the latter on the network's operational state. The server responds to orders from the base station by acknowledging them and sending the base station any new config or software updates it has available. Water water features, parking lot surveillance systems, and vending machines may all benefit from having sensors included into their designs so that their status can be reported in real time [8]. Using the Internet of Things, cities may find problems like leaking pipes beneath. By putting RFID tags with the transported items and monitoring them using wireless sensors installed in nearby buildings and shipping containers, warehouse and logistic organisations like Amazon's delivery services may minimise misdirected shipments and prevent mistakenly lost merchandise [9]. Sensor data gathered in advance may alert Internet of Things gadgets to problems before they have an effect on the end user. Without the need for human labour and also data integration from diverse sources, previously stored IoT information may be utilised to make future planning choices via forecasting.

2. LITERATURE SURVEY

In a WSNs, energy efficiency is a very important performance metric. Several energy efficient routing schemes are introduced by researchers in the very recent years. To overcome multiple problems like energy consumption, load balancing and transmission cost many routing algorithms are published. The literature review in this paper is mostly based on Hierarchical routing in which multiple cluster-based routing techniques and route optimization algorithms are explored [10]. One of the biggest problems with WSNs is that the act of routing consumes a lot of power. Routing in this context means determining the most efficient and least energy-intensive route from a sensor's point of origin to the network's hub [11]. Therefore, when building a routing protocol for WSN, an intensive algorithm with reduced complexity is necessary. *Figure 2* depicts a classification of WSN routing methods.

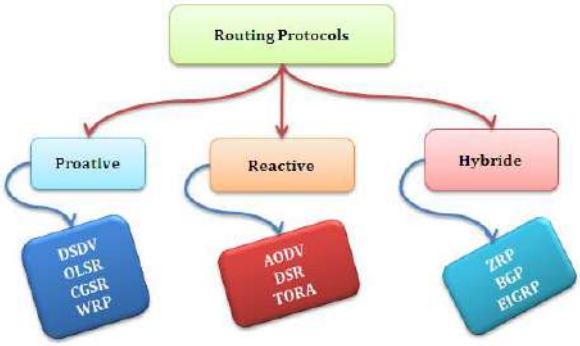


Figure 2: Schematic on a classification of WSN routing techniques

Figure 3 depicts a hierarchy of protocols, in which the lowest-energy sensor nodes are activated to collect data, while the highest-energy for processing and relaying the data to the final destination [12]. So, these protocols are utilised to accomplish routing while minimising energy use. Location-Based: These protocols don't assume any prior knowledge of the sensor nodes' coordinates.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

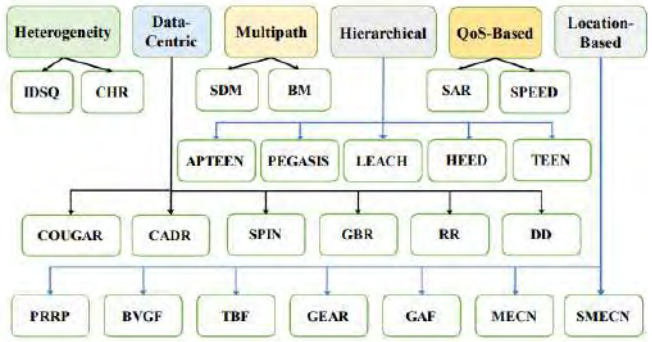


Figure 3: Existing energy efficient sensor network routing protocols

Using Global Positioning System (GPS) coordinates, sensor node locations may be determined; this data will then be utilised in conjunction with floods to determine the best possible route [13]. Data Centric Protocols: These protocols are dependent on the sequence number of the desired data and known to be query-based protocols. Base station will send a query to a region inside the sensor network to get certain information and waits for the acknowledgement from the sensor nodes in that region [14]. Procedures for Clustering: Here, the area covered by the sensor network is divided into smaller sections called "clusters," and leaders are chosen to oversee each section.

A reactive sensor network makes use of a protocol that reacts to rapid changes in variables like temperature. It has the ability to regulate a mission-critical data transmission and is designed to adapt quickly to changes in time-sensitive applications [15]. The lack of threshold-specific data at the base station may prevent it from accurately detecting the number of live and dead nodes [16]. A more realistic implementation may have prevented cluster collisions. It may be used to detect intrusions and explosions. Its inefficiency and uneven energy distribution are major drawbacks. Its usage of heterogeneous nodes and knowledge of residual energy for cluster head selection are both advantages when dealing with high energy levels [17]. The chosen cluster head has the disadvantage of sending data to the sink in a single hop [18]. The main drawback is that it has a large information processing overhead. Behera's I-SEP proposes a more reliable election process that evenly distributes power amongst cluster nodes and their members. Depending on their available starting energy, sensor nodes are divided into three categories: normal, intermediate, and advanced. In order to reduce power consumption, this protocol swaps energy states between CH and member nodes [19]. Using an opportunistic routing method, in which a group of candidate nodes is identified and employed to improve the dependability of long-distance node data transmission, the success rate in data transmission is raised. Power-saving and Safety-assured [20]. To solve the cluster formation issue with little costs and maximum coverage, a suggested protocol for an unequal-sized clustering technique is presented.

By calculating the ideal cluster size interval, ECUC is able to circumvent certain energy consumption constraints. A well-functioning design that meets the needs of the study requires optimisation. Sensor network lifespan may be increased by optimising the network to use as little energy as possible. Non-deterministic polynomial time hard (NPHard) issues are what optimisation is best at solving[21]. To solve a given issue, optimisation methods seek to maximise or minimise an objective function. These procedures might have one or several goals in mind. It may be classified as either a single- or multi-objective function, depending on the number of performance parameters chosen for optimisation. For most optimisation problems in science, the multi-objective function is the first choice [22].

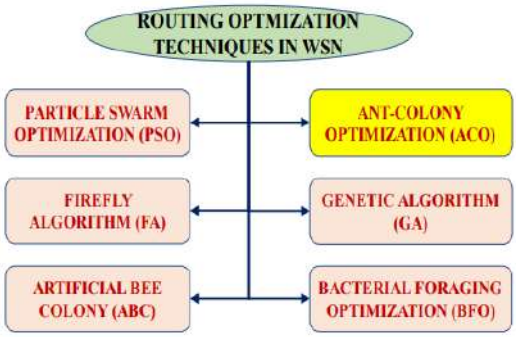


Figure 4: Efficient optimization techniques in WSNs

The majority of this study concerns a routing issue, whereas a subset concerns both routing and clustering. The primary goal of the optimisation approaches used to WSNs in *Figure 4* is to lessen the amount of power required to run the network and lengthen its useful lifespan [23]. Since the opted wireless sensor network is usually deployed in intricate areas or environment, only a few research approaches are found to be well suited. Various research sources provide better energy efficient schemes while simultaneously increasing network lifetime and to balance the network load among sensor nodes [24]. We reviewed an ample of load balancing and energy efficient protocols, where a quite few are found to be work excellent for energy consumption in sensor networks. Our research on several route optimisation methods has led us to the conclusion that Antcolony Optimisation (ACO) is the most trustworthy method for multi-hop data transmission in WSNs[25]. In view of real time computations, ACO shows superior performance when compared to other existing route optimization schemes. Although, ACO has a weak point of consuming more energy compared to PSO, it can be combined with centroid routing approach to build a hybrid protocol.

3. ENERGY SAVING SCHEME

The efficiency of data collection from the sensing field is improved by sink nodes in sensor networks. The sink node's job is to relay data from the sensor network's sensing field to the various user agents. The sink node collects all the information from the sensor nodes. The base station (BS), which acts as a central administrator by collecting and processing data, is the best possible sink. Therefore, a base station is seen as a unique node that constantly transmits data and has access to an infinite supply of power [26]. The ES3 method has two groups of nodes located inside the sensing area. First, the nodes in the network. Nodes outside of the network, number two. Packet transmission

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

occurs between network nodes and the base station, where the collected data is processed. In this paper, we present ES3, an Efficient Sink Selection Scheme for tree-like networks [27]. As was previously indicated, this design divides the connected devices in the network into three distinct categories. One: the primary hub or primary sink. There are also gateway nodes. Third, sensors that double as sinks. Each sensor node inside the sensing field is assigned an energy limit. If a sensor node's battery life drops below a certain percentage, it is removed from the network and its data isn't used in any further routing. This routing protocol may be used by mobile nodes in ad hoc networks [28]. The AODV is meant to reduce overhead and regulate traffic. The AODV routing system addresses both Route Discovery and Route Maintenance. Two separate functions are responsible for discovering new routes and fixing current ones, they are: Route Discovery and Maintenance. The reactive protocol maintains no permanent route table. AODV can analyze network topology changes fast. AODV routing protocols' information flow is seen in *figure 5*.

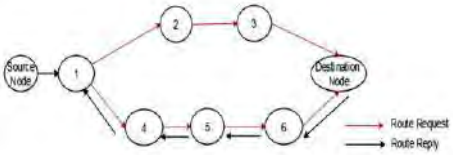


Figure 5: AODV Routing Protocols-Data Transfer

The ACO method for sensor network routing is inspired by the natural properties of ants and the related field of ACO. In ants, the ability to use Pheromone to find the quickest route between their colony and a food supply is the primary source of incentive (Impulsive Chemical Substance) [29]. Pheromones are left behind by ants when they go from one place to another. To find food, ants go out in high concentrations of pheromones. Faster completion of minimal pathways results in more significant amounts of pheromone. The colony can find the quickest route via the positive strengthening process. In generation u, the chance that an ant l at node k would move to node l is given by the equation 1.

$$Q_{J,L}{}^{K}(u) = \frac{tj,k(u)d_{j,k}{}^{-\beta}}{\sum_{u\varepsilon}\Gamma^{kt}j,v_{j,v}{}^{d-\beta}}\ ,k\epsilon\Gamma^{k}{}_{j} \qquad (1)$$

The pheromone intensity at the edge is proportional to node I is proportional to the distance between these nodes. Pheromone-updating rules in equations 2, 3 and 4 are used to update the pheromone on all edges once all the ants have completed their tours.

$$t_{j,k}\ (t+1) = (1-\rho)_{j,k}(t) + \Delta t_{j,k}(t) \qquad (2)$$

Where $\Delta t_{j,k}(t) = \sum_{K=1}^{NP}\Delta t^{k}{}_{j,k}(t)$ (3)

$$\Delta t_{j,k}(t) = \begin{cases} \frac{Q}{LK}\ ,if\ \ (j,k) \in tour\ done\ byantk \\ 0, \qquad\qquad otherwise \end{cases} \qquad (4)$$

Pheromone decay parameter, which measures the trial evaporation after the ant has decided where it will relocate is the pheromone decay parameter The ant k's tour duration is Lk, and the ant k's tour number is m. The administrative burden of centralized critical management systems may be reduced using this approach [30]. Additionally, the storage cost on each node is decreased since each node no longer has to store all public keys. After giving the cluster head the keys, it is essential to encrypt the information in the nodes not to be accessed by attackers. Encryption is explained in detail as follows.

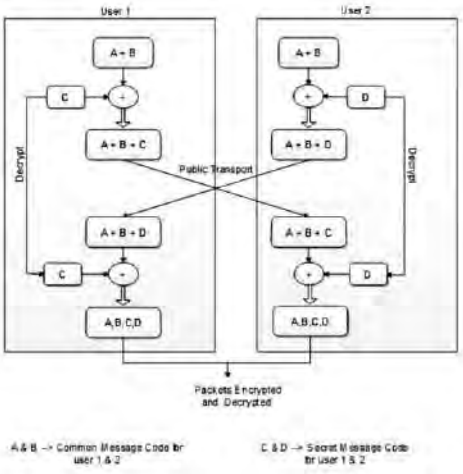


Figure 6: DHKE Method

Encryption: Data or information is encoded into a code to protect it from unwanted access. When it comes to data encryption, the DHKE technique is a good choice. The following is a detailed breakdown of how this algorithm works. Diffie-Hellman Key Exchange Algorithm: Using DHKE, cryptographic keys may be exchanged precisely. It's one of the first real-world applications of cryptography's key exchange mechanism. Any subsequent communications may be encrypted with the freshly generated private key using an asymmetric key cypher. *Figure 6* shows a block diagram for the DHKE Method.

The DHKE Cryptography is a widely used cryptography technology for ensuring network security. The DHKE's security perimeter should be raised. Duplicating P and Q subtracts the modulus n. In both the general population as well as amongst operators, the number is exploited. Using the secret key, a single user may transmit plain text to the encrypted public key.

4. PROPOSED METHOD

The ES3 network starts out with a single base station and a hop value of 0. The base station will then send out broadcast packets and look for its gateway node and its child nodes as the nodes are installed within the sensing field. Once they receive broadcast packets, that topology information is obtained by the neighboring non-network nodes. Then these non-networking nodes selfevaluate different metrics such as current energy level, hop count, communication distance to available sink and child node count of the available sink to match nearest sink weight and join the network and become one of the networking nodes. In a local tree, the sink for a given iteration of data is always the node with the highest weight. Usually, the selected sink in a local tree will have higher current energy level than the other sensor nodes which forms a local tree. The modeled

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

approach conserves power, builds a tree-based network quickly, and extends its lifespan by optimizing internal performance measures.

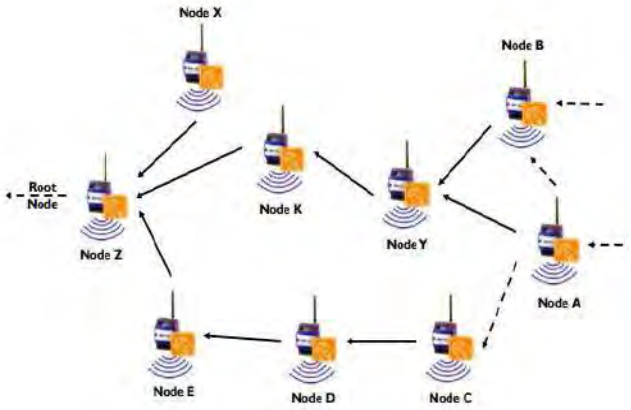


Figure 7: Local tree topology of sensor network in ES3

Figure 7 depicts the local tree topology in the hypothetical architecture. In this setup, Node A may share its data directly with Nodes B, Y, and C. In contrast. If Node A selects Node Y as the sink node, the packets will be routed to the node that is geographically closest to the sink. To account for this, each deployed sensor node's hop count and distance between sink nodes are recorded and considered. The number of a sensor node's offspring must be recorded as part of the decision-making process for which sink node to use. Node A's sink node in *figure 7*. Node B may serve as a sink node for Node A for a certain simulation period t sec, hence boosting Node A's lifespan. After a certain amount of time, Node A may reselect Node C as the sink node in order to equalise the power use of Nodes B and C. When Node A detects that Node B's energy has fallen below that of Node C, an event is triggered. While the simulated tree-based network organises itself. Therefore, in WSNs-assisted IoT contexts, a dependable network model and appropriate routing technique are needed to prolong the useful life of the networks.

Figure 8: Block diagram of the M-TCFPA approach

5. ENERGY EFFICIENT OPTIMAL CLUSTERING

Selecting an ideal CH and routing path with this M-TCFPA algorithm ensures reliable data transfer across networks. For the clustering and CH selection phases of this M-TCFPA approach there are three stages.

Typically, network clustering is utilised to reduce energy consumption while aggregating data. Because it factors in trust and integrity into its fitness function, M-TCFPA can thwart attacks that are meant to do harm. *Figure 8* depicts a block schematic of the M-TCFPA method. First, nodes in the targeted network region are distributed randomly, and then clusters are formed using K-means. Since that time, K-means clustering has primarily utilised the calculation of euclidian distances. This section explains how to use the M-TCFPA to choose the best CHs in each cluster after the network has been grouped. Nodes' energy usage is reduced by selecting the most secure optimal CHs from the clusters in the second step of selection. The Flower Pollination Algorithm (FPA) was developed by Xin-She Yang, a population-based optimization method. This FPA mimics the pollination process of flowers. Pollination, in the broadest sense, refers to the process by which pollinators move pollen from one plant to another during the course of a mating cycle (e.g., insects). The flower/pollen of the M-TCFPA is used as a metaphor for the possible solution. An acceptable candidate node is selected from the population of sensor nodes. As the population grows, so does the demand for CHs in the network. One to N sensor nodes are assigned to each population and each population has a random node ID assigned between 1 and N. The equation for u-initialization TCFPA's can be found here (5).

$$x_i = \left(x_{i,1}, x_{i,2}, .., x_{i,NCH}\right) \quad (5)$$

Self-pollination and cross-pollination are the two methods used in the M-TCFPA for pollination. Self-pollination refers to the pollination of a single flower by the pollen of another flower. Cross-pollination occurs when a plant receives pollen grain from another plant and uses it to reproduce. Furthermore, there are a variety of techniques in which flowers attempt to disperse pollen. Pollen is dispersed by the wind in a biotic pollination, one of the several pollination techniques. The term "biotic pollination" refers to the process of pollination carried out by animals such as birds, bats, insects, and more. Insect pollinators, such as honeybees, are also checked for floral constancy. These pollinators prefer a specific type of bloom and shun the rest, which are employed to boost the reproduction of the same type of flower.

The mobile nodes in the MANET have built up enough mutual confidence to exchange data with one another. In this situation, the trust value is generated through direct discussions, whereas in the prior case, it was linked to packet forwarding behaviour.

$$\text{Trust} = \frac{\text{Amount of broadcasted packets}_{a,b}}{\text{Amount of received packets}_{a,b}} \quad (6)$$

The chosen CH must complete a challenging challenge in order to carry out be the best option for data transmission. Equation expresses the remaining energy (7).

$$RE = \sum_{i=1}^{NCH} E_{CH_i} \quad (7)$$

Where RE stands for residual energy and E_{CH_i} stands for the residual energy of the i^{th} CH.

Keeping in mind the nodes' energy consumption, Euclidean distances between them, as well as between CH and their final destination, should be minimised. To make a single goal

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

function that can be represented mathematically, the weighted sum technique is employed (8).

$$\text{Fitness} = \beta_1 \times \text{Trust} + \beta_2 \times \text{IF} + \beta_3 \times \text{RE} + \beta_4 \times \text{Distance} \quad (8)$$

Where, $\beta_1, \beta_2, \beta_3$ and β_4 indicates the relative importance of the various fitness values (0.3, 0.3, 0.2, and 0.2). Using the previously described fitness function, the cluster's optimal CHs may be derived from the data. M-robustness TCFPA's against malicious attacks is boosted by the trust value it uses, while the integrity factor assesses whether or not data packets are correctly formatted and sent in a timely manner as specified by the protocol. As a result, the PDR benefits from the component of trust and integrity.

6. RESULTS AND DISCUSSION

Experimental set up: Install NS-3: Begin by installing NS-3 on your machine. You can download the latest version of NS-3 from the official NS-3 website. Create a network topology: NS-3 allows you to define network topologies using various classes and modules. You can create a network topology with 100 nodes by using the existing Node and Point-to-point Net Device classes provided by NS-3. Configure node behavior: Once the network topology is created, you can configure the behavior of each node in the simulation. This includes setting up network addresses, protocols, traffic patterns, and application-level behavior. Choose appropriate routing protocols: Depending on your simulation requirements, you may need to select and configure routing protocols for the network. NS-3 supports a wide range of routing protocols such as OSPF, RIP, and AODV. You can choose the appropriate routing protocol based on your simulation scenario. It's important to note that NS-3 is a highly flexible and customizable network simulator, and the specific tools required for a 100-node setup may vary depending on your simulation objectives. The above steps provide a general guideline for setting up a 100-node network simulation in NS-3, but you may need to further explore NS-3's documentation, examples, and modules to tailor the simulation to your specific requirements.

The focus of this section is on the NS-3 simulation findings, which includes data analysis. Using computer simulations, we can see how the network nodes respond to a Black Hole attack (see *figures 9-12*). Metrics including latency, throughput, and network load are used to approximate the potential behaviour of an intrusion-based black hole assault.

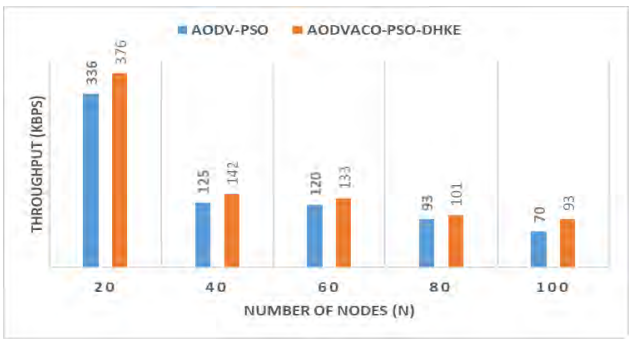


Figure 9: Node vs. Throughput

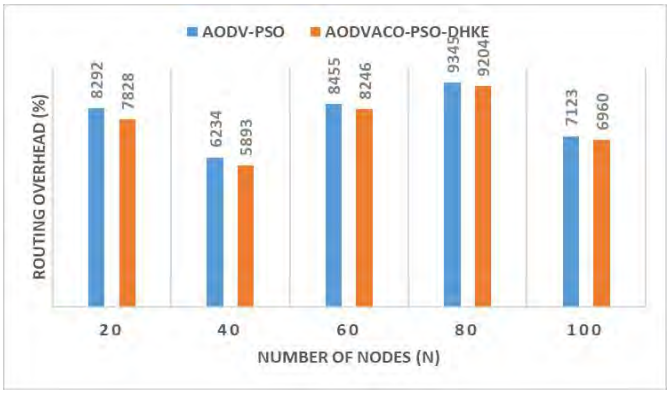


Figure 10: Node vs. routing overhead

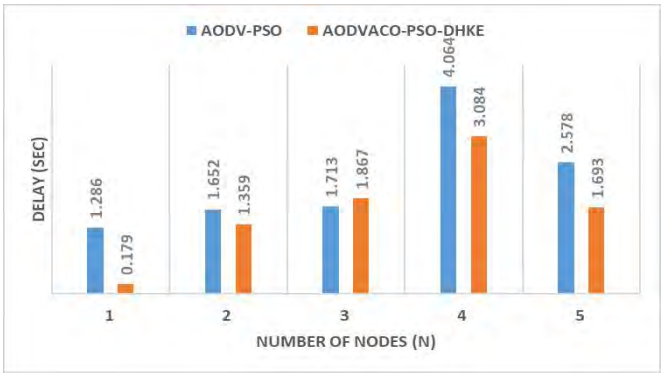


Figure 11: Node vs. delay

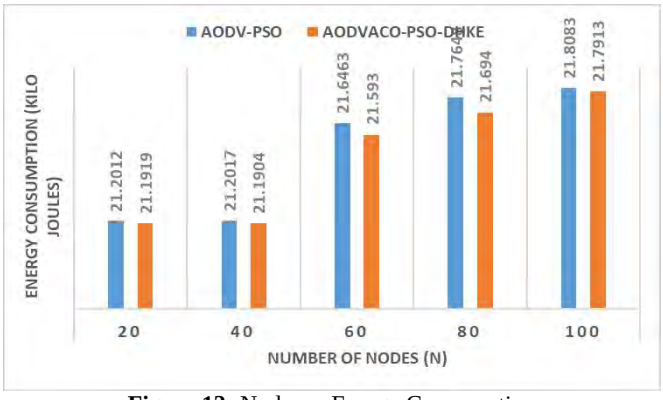


Figure 12: Node vs. Energy Consumption

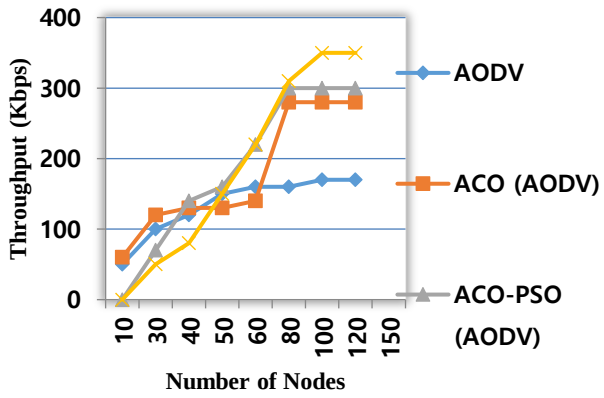


Figure 13: No. of Nodes versus Throughput

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

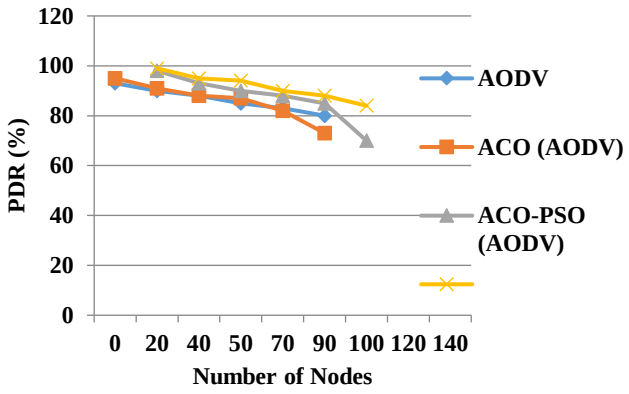


Figure 14: No. of Nodes versus PDR

A packet's end-to-end latency was tested using two simulations. As the number of nodes engaged grows, so does the likelihood of an assault being successful. When there are 20 nodes, as shown in *figure 9*, the AODV and OLSR delays are shown in red. To be able to observe the graph shown below shows how the black hole attack affects the network as a whole, relative to the normal operating protocol. The graph demonstrates a larger latency when there is no rogue node in the network. Similarly, a simulation would be performed with 40 nodes and the presence of a rogue node in the network. There is a notable lag when employing 40 nodes, as seen in *figure 13*. Both protocols' latency decreased dramatically when a rogue node was present. At 20 nodes, average latency rises by around 5%. This was shown to be the case for both protocols, which were tested to determine which was more affected by the attack.

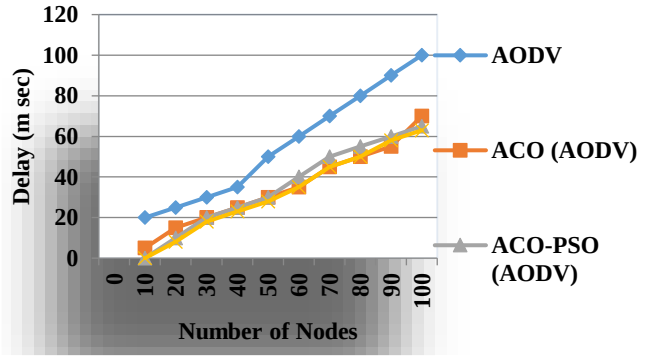


Figure 15: No. of Nodes versus Delay

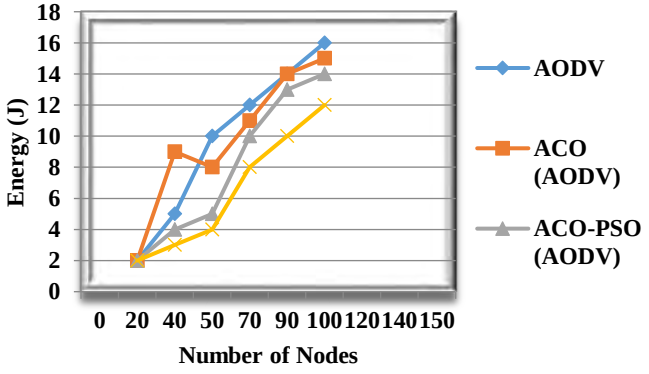


Figure 16: No. of Nodes versus Residual Energy

In comparison to the AODV, the OLSR has a little longer delay, as seen in *figure 14*. This is also true when there are fewer nodes. Latency increases as the number of nodes increases. A node network with 40 distinct integers may demonstrate this. *figure 15*. depicts that AODV has a substantial delay when compared to OLSR for a network of 40 nodes

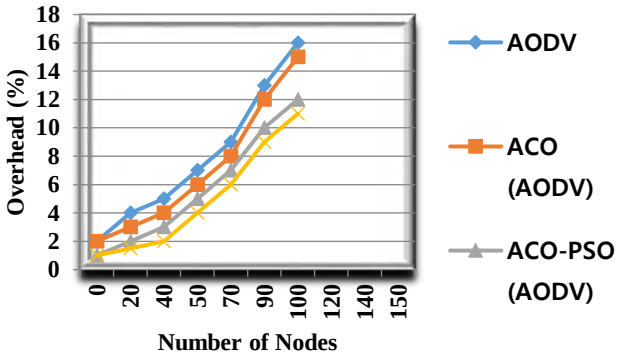


Figure 17: No. of Nodes versus overhead

There will be a higher throughput for OLSR if there is no attack (no malicious node present) (see *Figure 16* for 20 nodes). There is less routing forwarding and traffic going through the system as a consequence of this reduction in traffic. The rogue node just discards the data instead of delivering it, causing a reduction in total throughput. Also, when the malicious node is not destroying packets, AODV has superior throughput than while under assault. OLSR outperforms AODV in terms of throughput when compared side by side.

Figure 17 displays the same trends with 40 nodes, showing that although throughput rises with more nodes, the attack and non-attack throughput patterns are similar to those seen with 20 nodes. *Figures 14* and *15* show that there is a rogue node in the system. and ten more. OLSR and AODV's throughput is fairly high in a normal working protocol when no assaults are applied. These two protocols both drastically degrade throughput when attacked.

Table 1: Comparative analysis of M-TCFPA

Performances	Method	Number of Nodes				
		10	50	100	150	200
PDR (%)	TBSMR [20]	98	97	95	92	90
	M-TCFPA	99.77	99.37	98.50	99.04	98.85
PLR (%)	TBSMR [20]	2	3	5	8	10
	M-TCFPA	0.22	0.62	1.49	0.95	1.14
Throughput (kbps)	TBSMR [20]	510	500	470	440	410
	M-TCFPA	1090.36	1086.1	1077.85	1083.72	1083.27
AEED (sec)	TBSMR [20]	0.1	0.15	0.2	0.25	0.3
	M-TCFPA	0.024	0.027	0.038	0.060	0.070

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

In *table 1*, a comprehensive comparison is presented between the TBSMR [20] and M-TCFPA [21] methods in terms of accuracy. The comparison encompasses various configurations of sensor nodes, ranging from 10 to 200, providing insights into their performance across different scenarios. The results clearly demonstrate that the M-TCFPA method [21] outperforms the TBSMR method. The TBSMR approach [20] solely focuses on trust, traffic, and residual energy values during data transmission, disregarding the consideration of distance. As a consequence, there are delays in transmitting data packets. Conversely, the M-TCFPA method takes into account both trust and integrity, ensuring the reliability of nodes and accurate data exchange. Consequently, the M-TCFPA approach significantly improves data delivery while reducing transmission latency. When comparing the vulnerability of both protocols to a black hole attack, it was demonstrated that OLSR had a far higher throughput than AODV. Similarly, it is because OLSR is an active routing protocol that the protocol initially checks to verify whether there is a routing route before directing traffic to it. More sources have lesser influence on throughput than fewer sources, according to our research. AODV's low total throughput may be attributed to route reply. As soon as the malicious node delivers its route reply, the malicious node receives the data and may act quickly and discard it. Thus, network throughput suffers greatly as a result of this.

Twenty nodes were used to measure the load on the OLSR and AODV networks for the first time. During the experiment, 20 nodes and two scenarios were tested: one faulty node in a network and no bad nodes in a network (typically functional protocol). With and without a rogue node, the network load is shown in *figures 12 and 13*.

Figure 14 shows that OLSR and AODV have reduced network demands while the network is under attack. The rogue node doesn't send any data out into the network; instead, it just drops all of the packets it receives. For example, when there is no assault, it shows that OLSR is successfully routing its packets to their eventual destination, as seen in the accompanying picture. While the system is under attack, packet rejection minimises network traffic. AODV follows a similar path in the graph. When there are 40 nodes, there is a little difference between OLSR with and without assault. Routing traffic has grown as a result of the huge number of nodes. AODV doesn't alter even if there are 20 or 40 nodes.

There is a greater network cost on OLSR than AODV when malicious nodes are present, (as seen in *figures 15 and 16*) when the two protocols are pitted against one another. The OLSR network, whether it has 20 or 40 nodes, has high network load because the routing protocols can adjust to its variations between node restart and node halting. There is a delay in switching to new routes at higher speeds since the routing operations require more time for change. AODV is quicker to respond. The complexity of the OLSR increases as the number of nodes in the network increases, such that the pressure on the network is significantly reduced. As the node becomes more reliable, the strain on the network grows with each stop and restart.

7. CONCLUSION

This paper extensively discusses the weaknesses of routing protocols, the corresponding attacks exploiting these vulnerabilities, and various preventive measures. It becomes evident that no singular approach can ensure comprehensive security against all types of attacks. Therefore, it is necessary to address the vulnerabilities posed by external and internal attackers separately. Several secure variations of routing protocols have been proposed and developed in existing literature. However, none of these protocols have successfully provided protection against all forms of attacks. Additionally, the secure versions of routing protocols heavily rely on cryptographic techniques, which impose significant computational burdens on nodes with limited resources. After conducting a performance study, it has been determined that the M-TCFPA strategy surpasses the TBSMR approach. Addressing vulnerabilities from external and internal attackers necessitates separate approaches. Various secure versions of routing protocols have been proposed and developed in existing literature, but none have achieved comprehensive protection against all types of attacks. Furthermore, these secure routing protocols heavily rely on cryptographic techniques, resulting in excessive computational burdens for nodes with limited resources. Consequently, there is a need to develop a lightweight cryptographic technique that fulfills security objectives. Our investigation indicates that a Black Hole attack poses a greater threat to the AODV protocol compared to the OLSR protocol.

REFERENCES

[1] Hiremath, P. S., Anuradha, T. and Pattan, P., “Adaptive fuzzy inference system for detection and prevention of cooperative black hole attack in MANETs,” International Conference on Information Science (ICIS), pp. 245-251, 2019.

[2] S. Yang, K. Kang, Y. Cho and S.Y. Chae. “PAMP: Power-Aware Multi-Path Routing Protocol for a Wireless Ad Hoc Network”, Proc. IEEE International Conference on Wireless Communications and Networking, Las Vegas, USA, pp.2247-2252. 2018.

[3] A.K. N, S. A. Farooq, D. G. V, S. P. M, A. M. Reddy and R. Tanguturi, "Improved Secure Communication Mechanism for IoT Platform Devices," 2022 International Conference on Smart Generation Computing, Communication and Networking (SMART GENCON), Bangalore, India, 2022, pp. 1-6, doi: 10.1109/SMARTGENCON56628.2022.10083714.

[4] Y. Kim, E. Lee and H. Park. “Ant Colony Optimization Based Energy Saving Routing for Energy-Efficient Networks”, IEEE Communications Letters, Vol. 15(7), pp. 779-781, 2019.

[5] S. R. Kawale, K. Prasad, N. Anil Kumar, N. S. Reddy and B. Ashreetha, "Technologies for Comprehensive Information Security in the IoT," 2023 International Conference for Advancement in Technology (ICONAT), Goa, India, 2023, pp. 1-5, doi: 10.1109/ICONAT57137.2023.10080332.

[6] R.R Rout and S.K. Ghosh. “Enhancement of Lifetime using Duty Cycle and Network Coding in Wireless Sensor Networks”, IEEE Transactions on Wireless Communications, Vol. 12(2), pp. 656-667, 2018.

[7] V. S. Chinamuttevi, K. Prasad, A. Y. Begum, R. C. Tanguturi and B. Ahmed, "Implementation of an Early Warning System using Internet of Things and Rainfall Threshold," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 491-496, doi: 10.1109/ICECA55336.2022.10009211.

[8] Merlin, R.T.: Ravi, R.: Novel Trust Based Energy Aware Routing Mechanism for Mitigation of Black Hole Attacks in MANET. Wireless Personal Communications, 104(4), 1599-1636. (2019).

International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 11, Issue 2 | Pages 442-450 | e-ISSN: 2347-470X

[9] G. Gondhalekar, B. Ashreeth, G. R. Thellaputta, D. Venkataramireddy, M. Sumithra and N. Karyemsetty, "A Safety Assessment Model for Automotive Embedded Systems Networks," 2022 IEEE 2nd Mysore Sub Section International Conference (MysuruCon), Mysuru, India, 2022, pp. 1-5, doi: 10.1109/MysuruCon55714.2022.9972628.

[10] Rajashanthi, M. and Valarmathi, K.: A Secure Trusted Multipath Routing and Optimal Fuzzy Logic for Enhancing QoS in MANETs. Wireless Personal Communications, 112, 75-90 (2020).

[11] K. N, S. Vijayalakshmi, D. Baswaraj, P. Selvarajan, S. Chandramohan and M. Tiwari, "Towards Internet of Things: Integration of Wireless Sensor Network to Cloud Services for Data Collection and Sharing," 2022 International Conference on Automation, Computing and Renewable Systems (ICACRS), Pudukkottai, India, 2022, pp. 363-369, doi: 10.1109/ICACRS55517.2022.10029163.

[12] Vanitha, K., Rahaman, A.Z.: Preventing malicious packet dropping nodes in MANET using IFHM based SAODV routing protocol. Cluster Computing, 22(6), pp.13453-13461 (2019).

[13] Avinash Sharma, K.D.V. Prasad, Sadashiva V. Chakrasali, Chanakya Kumar, Abhay Chaturvedi, A. Azhagu Jaisudhan Pazhani, Computer vision based healthcare system for identification of diabetes & its types using AI, Measurement: Sensors, Volume 27, 2023, 10075. https://doi.org/10.1016/j.measen.2023.100751.

[14] Ranjeet Suryawanshi, Revanna C R, B. Kameswara Rao and Parismita Sarma (2022), Enhanced Diagnostic Methods for Identifying Anomalies in Imaging of Skin Lesions. IJEER 10(4), 1077-1085. DOI: 10.37391/IJEER.100452.

[15] Parismita Sarma, Takrim UL Islam Laskar, and Ramesha M (2022), Human Emotion Recognition using Deep Learning with Special Emphasis on Infant's Face. IJEER 10(4), 1176-1183. DOI: 10.37391/IJEER.100466.

[16] Aarti Hemant Tirmare, Revanna C R, Ramesha M and N. K. Darwante (2022), A Morphological Change in Leaves-Based Image Processing Approach for Detecting Plant Diseases. IJEER 10(4), 1013-1020. DOI: 10.37391/IJEER.100443.

[17] El-Semary, A.M., Diab, H.: BP-AODV: Blackhole Protected AODV Routing Protocol for MANETs Based on Chaotic Map. IEEE Access, 7, 95185-95199 (2019).

[18] P, A. Sharma, S. Reddy P, P. S. Patwal, "Data Analytics and Cloud-Based Platform for Internet of Things Applications in Smart Cities," 2022 International Conference on Industry 4.0 Technology (I4Tech), 2022, pp. 1-6, doi: 10.1109/I4Tech55392.2022.9952780.

[19] S. B. M, P. Pavankumar, N. K. Darwante, "Performance Monitoring and Dynamic Scaling Algorithm for Queue Based Internet of Things," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2022, pp. 1-7, doi: 10.1109/ICSES55317.2022.9914108.

[20] Desai, A.M. and Jhaveri, R.H., 2019. Secure routing in mobile Ad hoc networks: a predictive approach. International Journal of Information Technology, 11(2), pp.345-356.

[21] Puneeth Kumar, B.S., Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (Eds) Recent Advances in Metrology. Lecture Notes in Electrical Engineering, vol 906. Springer, Singapore. https://doi.org/10.1007/978-981-19-2468-2_3.

[22] B. Kameswara Rao, Ravi Shankar, Parismita Sarma, Abhay Chaturvedi, Naziya Hussain, Industrial quality healthcare services using Internet of Things and fog computing approach, Measurement: Sensors, Volume 24, 2022, 100517, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100517 .

[23] P, A. Sharma, A. Singla, N. Sharma, "IoT Group Key Management using Incremental Gaussian Mixture Model," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), 2022, pp. 469-474, doi: 10.1109/ICESC54411.2022.9885644.

[24] Sharma, K. S and M. R. Arun, "Priority Queueing Model-Based IoT Middleware for Load Balancing," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), 2022, pp. 425-430, doi: 10.1109/ICICCS53718.2022.9788218.

[25] K. R. Swetha, Namitha A R, Manu Y M, Rashmi G R and Veera Sivakumar Chinamuttevi (2022), IOT Based Smart Health Care System to Monitor Covid-19 Patients. IJEER, 10(1), 36-40. DOI: 10.37391/IJEER.100105.

[26] Raja, R. and Ganeshkumar, P., 2018. QoSTRP: A trusted clustering based routing protocol for mobile ad-hoc networks. Programming and Computer Software, 44(6), pp.407-416.

[27] M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.

[28] K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," Adv. Math. Sci. J., vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.

[29] Usha, M.S. and Ravishankar, K.C., 2021. Implementation of Trust-Based Novel Approach for Security Enhancements in MANETs. SN Computer Science, 2(4), pp.1-7.

[30] Shivashankar, and S. Mehta, "MANET topology for disaster management using wireless sensor network," in International Conference on Communication and Signal Processing, ICCSP 2016, 2016, pp. 0736–0740, doi: 10.1109/ICCSP.2016.7754242.

Measurement: Sensors 24 (2022) 100439

Contents lists available at ScienceDirect

Measurement: Sensors

journal homepage: www.sciencedirect.com/journal/measurement-sensors

A novel method of data compression using ROI for biomedical 2D images

Dankan Gowda V [a,*], Avinash Sharma [b], Rajesh L [c], Mirzanur Rahman [d], Ghazaala Yasmin [e], Parismita Sarma [d], A. Azhagu Jaisudhan Pazhani [f]

[a] Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India
[b] Computer Science & Engineering Department, M. M. Engineering College, Maharishi Markandeshwar (Deemed To Be) University, Mullana, Ambala, India
[c] Department of Electronics and Communication Engineering, East Point College of Engineering and Technology, Bangalore, Karnataka, India
[d] Department of Information Technology, Gauhati University, Guwahati, Assam, India
[e] School of Computer Science Engineering and Technology, Bennett University, India
[f] Department of Electronics and Communication Engineering, Ramco Institute of Technology, Rajapalayam, Tamilnadu, India

ARTICLE INFO

Keywords:
Medical image
Compression rate
Region of interest
Vector quantization
Hierarchical trees
Lossless image

ABSTRACT

In the era of modern medical imaging communal, huge volume of medical image data are being acquired and need to be transmitted and archived which necessitate the development of efficient image compression techniques on both 2D and 3D images. The compression of medical images is an essential process to support remote healthcare services. The medical diagnostics through these services require more accurate information from an image. As the property of inverse proportionality between the compression rate and quality of the image takes place in any kind of compression method, there is a need to sacrifice any one of those credentials (Quality or Compression Rate). With this context, Region of Interest (ROI) codecs are emerging and reduces this proportionality that yields more compression rate without compromising the quality. In this paper, presents an ROI based near lossless image compression method that incorporates the Set Partitioning in Hierarchical Trees (SPIHT) and Vector Quantization coding for medical images.

1. Introduction

The present decade tenants at visually enchanting world, which explores extensive attributes such as colors, textures, motions and so on. The human visual system can easily bear any kind of these attributes and has the potential to discriminate and interpret such things. Imparting such notion to a machinery point of view also evolving through the image processing techniques. Thus, image processing applications becomes on integral part of our life in which the image can convey huge amount of information in a compact form. The substantial growth in utilization of images in variety of fields such as Satellite Imaging, Bio-imaging, Multimedia services and different types of web based applications is impactly elevated. In addition, the medical images also take an essential role in our daily life. Since we live in a Digital Age, image processing plays a significant role in our day-today life by any means. Image compression is a powerful tool to compactly represent an image. It reduces the actual number of bits needed to store the images which effected with low transmission costs. The reversible compression is termed as lossless due to no account of data loss after decoding. In contrast, irreversible compression endured with data loss and hence it is known as lossy compression. In the medical civilization, large number of medical-related data and images are handled. The advancements in medical imaging techniques also increases the practice of both 2D and 3D medical images in the diagnostics process. The quality of medical image is very crucial factor that related to diagnostics accuracy and feasibleness. The lossless compression methods give high reconstruction quality with low compression performance whereas near lossless compression gives appreciable reconstruction quality with better compression performance. With these concern, several image compression methods were proposed for 2D as well as for 3D medical images. Image compression is compulsory for various applications such as multimedia, documents and medical imaging which require huge data storage, transmission and retrieval. Significant storage capacity and transmission bandwidth are needed for uncompressed images. The intention of image compression is to represent image as a data in compact form and this can be done by reducing the redundancy of image data.

* Corresponding author.
E-mail addresses: dankan.v@bmsit.in (D. Gowda V), asharma@mmumullana.org (A. Sharma), rajeshlakshman.ece@eastpoint.ac.in (R. L), mr@gauhati.ac.in (M. Rahman), ghazaala.yasmin@gmail.com (G. Yasmin), parismita.sarma@gmail.com (P. Sarma), alagujaisudhan@gmail.com (A.A.J. Pazhani).

https://doi.org/10.1016/j.measen.2022.100439
Received 18 August 2022; Accepted 26 August 2022
Available online 31 August 2022

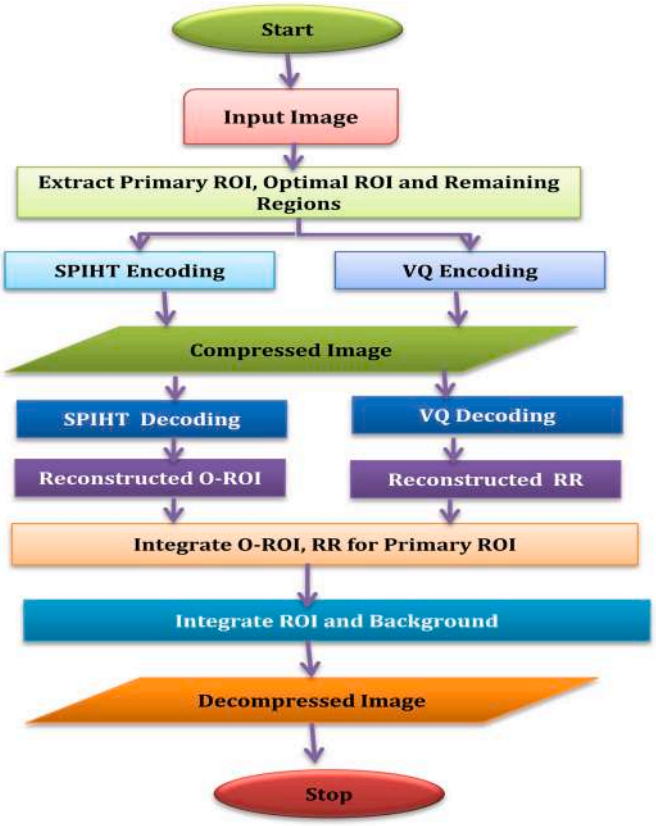


Fig. 1. Flowchart of the proposed method of data compression.

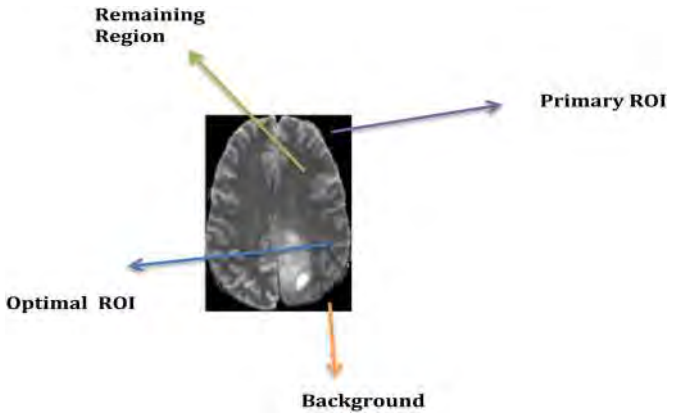


Fig. 2. Illustration of regions in the medical image.

2. Literature survey

In most of image processing operation, image encryption is done before compression. Major issues in such scheme are the design of pair of encryption and compression algorithm. Highly efficient encryption followed by compression scheme is developed using prediction error and random permutation (Jiantao Zhou et al., 2014 [1]). In this approach, highest level of security is provided by means of prediction error. The encrypted images are then compressed efficiently using arithmetic coding strategy. The result indicate that the compression efficiency is not satisfactory compared with state of art techniques.

An enhanced scheme of compression to compress the sequence of video based on adaptive prediction is presented (Ying Li & Sayood K 2007 [2]) for natural images. This compression technique consists of four stages such as pre-processing, symbol prediction, selection of

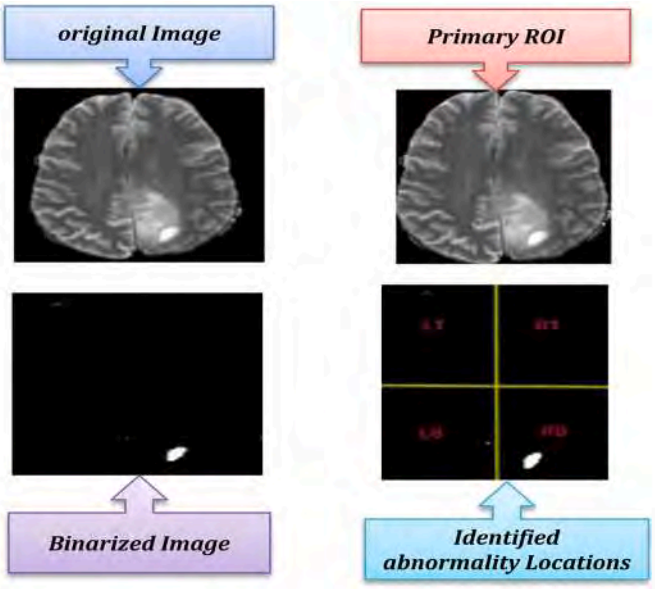


Fig. 3. Process of abnormality location setting (ALS).

prediction mode and lossless coding of prediction output. In the pre-processing stage, the video sequence processing are determined on a particular domain. The operational domain may be spatial or wavelet domain. In the prediction stage, three predictors are used to minimize redundancies like spatial and temporal. The predictors selected to reduce the redundancies are temporal predictor, spatial predictor and direct predictor. The prediction information is encoded using context based lossless arithmetic coding. The results signify that increase in compression efficiency is obtained compared with state of art methods.

Optimization technique based image compression scheme can improve the efficiency of prediction. A scheme of image compression based on particle swarm optimization concept is presented for compression of natural images (Nishat Kanvel et al., 2010 [3]). In this work, the prediction function for the lifting scheme is adaptively changed to obtain more accurate prediction of pixels. By means of adaptively changing the prediction function, compression ratio can be greatly increased.

During the image compression process, compression should be done to retain all the important information. The image details that need to be preserved during compression are decided using optimization techniques. Transform based compression techniques are generally used to compress the image, but these techniques produce blocking artifacts. Compression techniques integrated with particle swarm optimization (PSO) and genetic algorithm (GA) are proposed (Prabhjeet Kaur & Parminder Singh 2015 [4]) to eliminate blocking artifacts. In wavelet based image compression, image is transformed into wavelet coefficient then wavelet coefficients are selected using soft or hard thresholding. The threshold selection is done automatically using particle swarm optimization (Kranthi Kumar & Naidu 2014 [5]). Better PSNR and CR is obtained with automated coefficient selection using PSO.

Video image compression in lossless fashion using temporal information is investigated (Dania Brunello et al., 2003 [6]). This technique consists of motion compensation, 3D linear prediction and Golomb Rice coding. The performance of the compression is evaluated for different dimension of block and search area of motion vector. Best universal compression algorithm for finite N block is presented (Jacob Ziv 2009 [7]). In this, context tree algorithm and adaptive Huffman is also applied to N block successively. Best error free compression is obtained for sequences when N tends to infinity.

Real time image processing using Run Length coding for controlling Humanoid robot is implemented (Christopher Messom 2007 [8]). Using

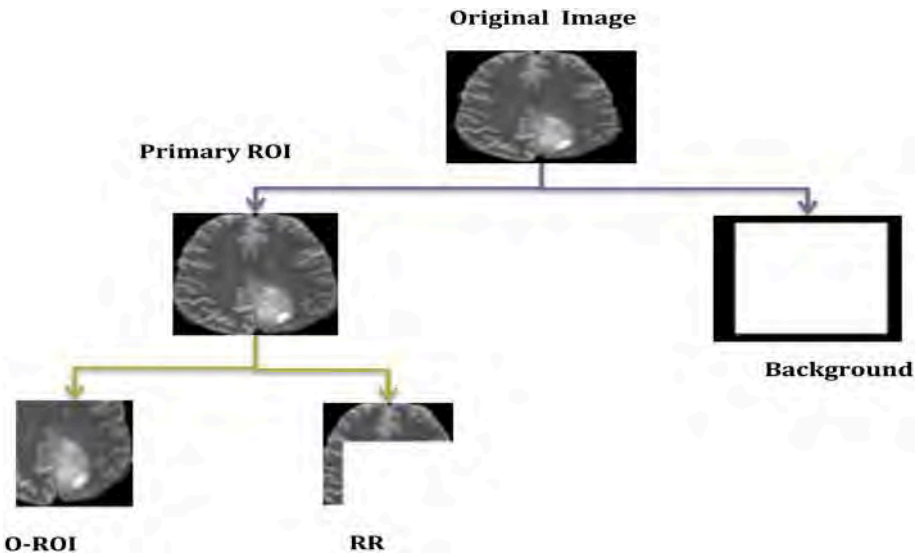


Fig. 4. Separation of regions.

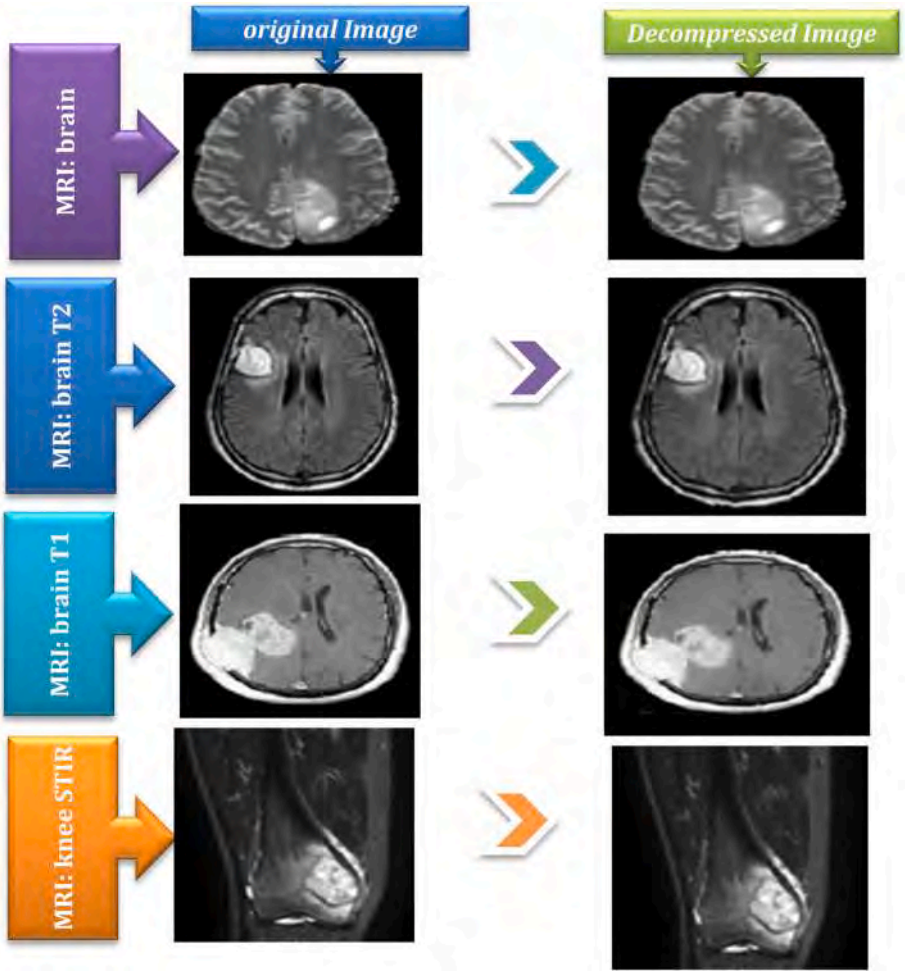


Fig. 5. Original and decompressed images by the proposed method.

RLE algorithm, any object present in the image can be identified also it provides size and position of the object. RLE based Hough transformation is utilized to recognize the land marks. This system is tested by simulation of dynamic robotic system and image processing sub system. Storage and transmission of three dimensional contents are very critical. Hence 3D compression is challenging area in image processing. Compression model using the 3 dimensional time varying facial expression is presented (Junhui Hou et al., 2014 [9]). In this work, isometric nature of human face expression is used to characterize the 3 dimensional faces into 2 dimensional canonical domain expressions.

Table 1
Performance comparison.

Test Images	PSNR (dB)	Proposed Method			
		CR	BPP (in bits)	Computation Time (in Sec)	
				Enc.	Dec.
MRI: brain	38.13	49.1	0.18	0.221	0.0912
MRI: brain T2	37.5	18.34	0.48	0.2132	0.0987
MRI: brain T1	37.88	26.35	0.36	0.2215	0.0944
MRI: knee STIR	38.45	25.13	0.38	0.2256	0.1102

These 2D expressions will generate 2D geometry videos. Using the inherent property of low rank sparse matrix decomposition is applied to 3D components and then compressed using video codec based on high precision rate and distortion model.

Generally video coding techniques consider video as a sequence of frames and are focussed on treating temporal and spatial redundancy separately. A new scheme of compression method is presented to form a non XY plane (Anmin Liu et al., 2011 [10]) in which optimal compression plane is determined in the pre processing stage to exploit the video redundancy. Then standard video coding is applied for further compression. This method can be suitable for lossy as well as lossless coding and also it can be applied for compression with or without prediction.

Commonly embedded compression codec is designed to reduce transmission bandwidth of the HD video (Yin-Tsung Hwang et al., 2015 [11]). This codec utilizes intra frame coding method and can be applicable for lossless as well as rate controlled near lossless scheme. Adaptive Golomb-Rice coding in combination with context modelling technique is used to minimize the complexity in computation. Piplined architecture and parallel computing kernel can be used to increase the throughput.

Hybrid coding algorithm consisting of two stage global estimation and hybrid coding is proposed (Yun-Chin Li & Yung-Chang chen 1998 [12]). Global estimation is determined in the first stage using feature based method. Results of the first stage are further refined using gradient method. Estimation algorithm is utilized to eliminate local facial expression. Then, by using clip and paste method facial areas are coded. Segmentation of different moving areas can be modelled using affine method. Model based coding method produces better results in terms of PSNR and CR.

The medical data transcriptions are increasing rapidly that handles MRI, CT and digital acquisition of medical images for the telemedicine applications. The medical images with higher capacity are difficult to transfer over the various wired and wireless networks (Penedo 2003, Pearlman 2004, Deepak 2014, Nikhil 2015, Gibson 2004 [13]). This problem is reduced using various encoding techniques like Huffman, Arithmetic coding, Embedded Zero tree Wavelet (EZW), LembelZiv (LZW). The medical images are to be stocked up in storage media for a period more than 30 years for diagnostic purpose. This increases the necessity of efficient compression algorithm.

In medical image compression, the diagnostic features are to be preserved without any loss which allows the radiologist to make accurate diagnosis. Even a little loss in quality will lead to a wrong prediction and create serious issues. Therefore, researchers focused on lossless compression techniques (Peter Schelkens et al., 1999, Jecob strom et al., 1997 [14]). Near lossless compression with scalability in resolution is proposed for CT and MRI images (Jonathan Taquetet al. 2012 [15]). This method composed of DPCM scheme with inbuilt adaptability and new hierarchical oriented prediction. This approach provides better scalability with improved compression performance.

A novel method for compressing the chromosome images based on the important features is investigated (Zhongmin Liu et al., 2002 [16]). This method comprises of a combination of lossless compression in the requisite part of chromosome and lossy coding in the rest of the image. In this, differential operation is performed to decorrelate the pixels and then finally wavelet transform is employed on ROI regions. SPIHT coding scheme is modified to produce the bit stream for ROI and other regions. Two types of chromosome samples such as spread and karyo types are utilized to test the performance of the compression. Obtained results are compared with JPEG 2000 standard. Lossless compression using the idea of adaptive prediction and multiplicative auto regressive technique is proposed for medical images (Zuo-Dian Chen et al., 1999 [17]). This method is focused to enhance the prediction accuracy. First, the image is divided into block and then every block is adaptively predicted using any prediction scheme used in JPEG mode. After the prediction, the residual values are processed using MAR model with entropy coding. Results are analysed with conventional coding methods.

Wavelet based compression methods are generally an efficient compression scheme for medical images. The implementation of compression schemes assumes that data sizes are powers of two. Hence, optimal selection of wavelet is the major issues in wavelet based compression. When the dimension of the data increases and data are not the multiple power of two then it will increase the computational risk. Four dimensional wavelet compressions using Zero tree concept was developed for arbitrary size ECG data (Li Zeng et al., 2002 [18]). Border artifacts can be minimized by using a pair of symmetric or anti-symmetric wavelet combined midpoint symmetry as boundary condition. This compression scheme was investigated on 3D data sequences obtained from clinical cardiac ultrasound sequences.

For the past few decades, 3D image acquisitions become popular in health care centers and research in imaging modalities. Compression of 3D images employ 3D wavelet transform to make use of the correlation

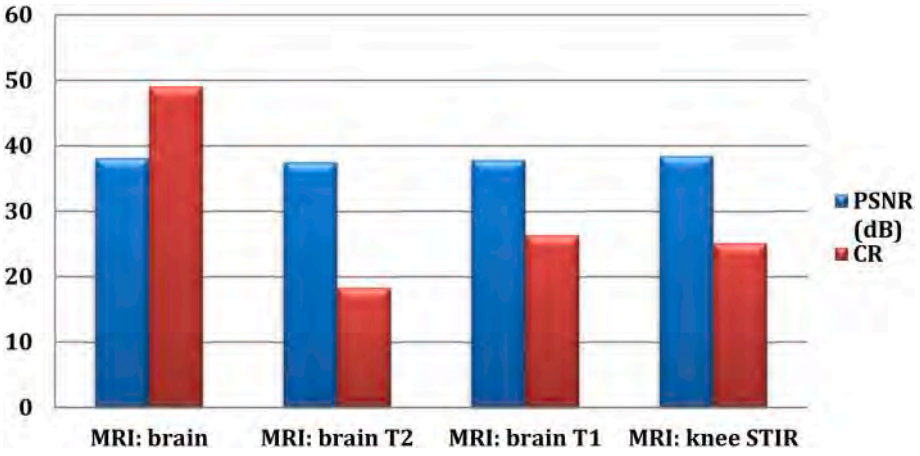


Fig. 6. Proposed PSNR and CR values.

present between slices to increase the performance. But, these methods are not based on symmetrical property in the region of interest. A new scalable lossless image compression method is proposed (Sanchez et al., 2009 [19]) for 3D medical data that exploits the inherent symmetry property of data. This work uses 2D wavelet transform for data decorrelation. Based on anatomical symmetry in the medical image, intra band prediction is done to reduce the sub band energy. Modified embedded block coding approach is used to encode the residual data obtained after prediction.

Performance analysis of several compression methods presecribed for medical images is investigated (Vibha Tiwari et al., 2012 [20]). Quality assessment parameter such as PSNR, NCCR, MD, SC and SSIM are evaluated for various medical images such as MRI, CT, and Ultrasound subjected to various compression. Improvement in the performance of JPEG2000 is investigated (Tim Bruylants et al., 2015 [21]) for the purpoes of volumetric compression. Image encoding technique using binary wavelet transform is proposed (Tirupathiraju Kanumuri et al., 2014 [22]) for medical images such as MRI and CT. Gray scale image is initially converted into eight binary plane. Then, every binary plane's information are transformed using binary wavelet transformation. Using quadtree based encoding method every bit plane is coded from the most significant plane. Significant enhancement in bitrate for the desired PSNR is achieved.

In region based compression scheme, the image is initially partitioned into several spatial regions. A region is distinguished from another region in terms of characteristics of intensity or salient details. Segmentation of image into several region is done for two reasons: to use appropriate encoding method tailored for the portion of the region and to have different resolution in different region. Commonly segmentation process can be done manually or automatically, thereafter it is possible to apply different compression scheme in different regions. By means of utilizing different compression method in various spatial region, it is possible to achieve higher compression by highly compressing the NROI.

This type of employing ROI based compression technique is termed as tunnable region specific method of compression. Compression of image in lossy fashion results in degradation of image quality and provides dianostic details loss. Similarly, applying lossless compression for the whole image is costly and expensive in case of low bandwidth environment such as tele radiology. Therefore, in accordance with salient features or diagnostic details present in different region, it is possible to compress image with high resolution in quality.

3. Need for medical image compression

Medical imaging is the study of human anatomy, physiology and pathology based on imaging techniques. In modern hospitals and diagnosis centers, voluminous health-related data are being acquired and processed. As the quantity and quality of the images increased extremely, the amount to store such images as well as the bandwidth for transmitting those images are also increased. Moreover, teleradiology in telemedicine applications are gaining momentum in extending the medical services regardless of distance and time. The picture archiving and communication system (PACS) is presently used in medical field which deals with medical image acquisition, display, transmission and archiving.

PACS facilitates the systematic utilization of medical imaging for patient care. It uses the international standard namely Digital Imaging and Communications in Medicine (DICOM) which comprise the compression technique and TCP/IP based communication protocol for transmission. Although DICOM standard have wide acceptance in medical practice, the quality of such images at high degree of compression rate is still poor.

Since, a massive amount of medical records is need to be processed and transmitted, the efficient compression techniques are essential in the field of telemedicine. Moreover, the evolution of medical imaging techniques shows its effect as producing 3D medical images over the collection of 2D images. Hence, the development of compression technique for 2D as well as 3D medical images are also indispensable. There are plenty of compression techniques available for 2D images and some of notable standard techniques works for 3D images too. The compression methods comprising the combination of certain features like lossless/near lossless, lossy and object-based are also available in the literature.

4. ROI based coding

The regions which are considered to be more important than other parts of medical images are known as a ROI. Our main ideology about the ROI method is to separate the foreground from the background of an image and the foreground is taken as ROI. The ROI based image compression method provides the best results than the normal whole image compression methods (Doukas and Maglogiannis, 2007 [23]). Aforementioned, the foreground is taken as ROI (i.e., P-ROI) and then the affected portion in a medical image is handled as a O-ROI where we apply the SPIHT method. In every medical image, the affected portion is cited as a diagnostically important region. Hence, there is a need to compress the entire affected area (O-ROI) with high fidelity. The process of finding the affected area and such high-fidelity compression on those portions are explained in further sections. However, our aim is to attain the best quality in the clinically important region such as tumors, traumatic brain injury, stone in the kidney MR images or any abnormality in the images of human organ. For that, we use the SPIHT method on the O-ROI and Vector Quantization in the remaining regions.

Quantization is a powerful lossy compression method that is utilized by employing a smaller set of code vectors for a larger set of data vectors. In vector Quantization, data points are coded with the nearest code vector (Gersho and Gray, 2011 [24]). Indeed, Quantization is the process of approximating the continuous data points with discrete values. The advantage of using Vector Quantization is only the small fraction of their vector space needed to represent the large black of data points. It is also called as block source codes and block quantizers (Linde et al., 1980 [25]). The significant challenge in the Vector Quantization method is the construction of an efficient codebook. There are three most widely used codebook generation algorithms namely K-Means based algorithm, Self-Organize Map (SOM) and the Linde–Buzo–Gray algorithm (LBG) (Lehn-Schiøler et al., 2005; Bal et al., 2019 [26]). Among these algorithms, LBG is a proven technique which used in most of the Vector Quantization methods. This work associates the LBG Vector Quantization as lossy compression to compress the Remaining Regions (RR) of the input image.

In the past few decades, the wavelet is noted as the best solution for image compression. Though DCT is used in JPEG for compressing images (Jamil et al., 2019 [27]), the wavelet's dominant quality for compressing images is a remarkable one. So, it is incorporated by JPEG2000 (Encyclopedia of Information Science and Technology, 2017). Later, some wavelet based image coders were also proposed such as Embedded Zerotree Wavelet (EZW) (Shapiro, 1993 [28]) and the Set Partitioning in Hierarchical Trees (SPIHT) (Said and Pearlman, 1996 [29]) which are the state-of-art techniques for image coders. Therefore, this paper presents a new compression algorithm for medical images. It uses the Bounding box and our own Abnormality Location Setting (ALS) algorithm to separate the P-ROI and O-ROI respectively. Unlike other ROI separation methods, these methods need only a very tiny amount of reconstruction details and it does not occupy many bits on the encoder as an overhead. Excluding of background and using the lossy compression namely vector quantization vastly increases the compression performance. The usage of SPIHT in the O-ROI portion preserves the quality on the clinically important region which leads to attain the near-lossless performance.

5. Proposed method

The flowchart of our proposed method is depicted in Fig. 1 and the step-by-step methodology of our proposed method is explained through the following steps.

The first step in the proposed method is to extract the primary ROI (P-ROI) from the input medical image. Our idea is to separate the background and foreground and this foreground part is taken as a P-ROI. The Bounding Box method is used to detect and extract the P-ROI in our proposed method. The ROI from an image is shown in Fig. 2. (highlighted with green). It is necessary to indicate the size of the image and details of the bounding box to the decoder for the reconstruction of the image.

In second step, normally in medical images, any abnormality in the human organs are depicted with high intensity values. This is the key to extract the O-ROI from any medical image as given in our Abnormality Location Setting (ALS) algorithm. The main objective of our proposed method is to use the SPIHT coding method in the affected portion in an image and this portion is known as Optimal ROI (O-ROI). There is a necessity to streamline the location of abnormality to derive the O-ROI. Our algorithm detects and extracts the O-ROI as follows. The problem of selecting P-ROI, O-ROI and RR in the randomly selected sample of tumorous brain MRI are illustrated in Fig. 3. According to the details obtained from the ALS process, our method processed further to extract O-ROI and RR. After extracting O-ROI, it is subject to apply SPIHT encoding. SPIHT gives better results when the size is optimal (i.e., sizes such as 64×64,128×128, 256×256 and so on). The Remaining pixels other than the O-ROI are collectively termed as Remaining Regions (RR) and coded using Vector Quantization. These O-ROI and RR are encoded separately. The red rectangle in Fig. 2 indicate the O-ROI that further use with the SPIHT encoding and other regions out of the green boundary and within the red rectangle is termed as Remaining Regions (RR). The process of separating the Regions are depicted in Fig. 4.

In the step three, the SPIHT encoding is applied on O-ROI part. For a decomposition purpose, the biorthogonal wavelet bior 5.5 is used and the resultant wavelet coefficients are encoded by the SPIHT encoding method. On the other hand, the RR part is simultaneously encoded by the Vector Quantization encoding method. In the case of brain tumor image, we can see from the Fig. 2 that the tumor region is located in O-ROI part and encoding it with the SPIHT method can improve the diagnostics accuracy. Finally, the compressed bit stream of both encoders with necessary bounding box details is sent to the decoder for the decompression process.

In the step four, the inverse operations of the encoding phase are done to decode ROI with the compressed bit stream. The SPIHT decoding and Vector Quantization are performed separately to get the O-ROI and RR respectively. Then, these two parts are merged accordingly to obtain the decompressed actual ROI (i.e., Foreground) of an image.

In the step five, as neglecting the background zero intensity pixels for encoding, we need to merge the foreground and background portions in order to reconstruct the decompressed image with the actual size. With the help of the size details and Bounding Box details which are already sent to the decoder, merge the decompressed ROI (Foreground) with the appropriate background. Thus, we obtain the whole decompressed image with an actual size.

6. Results and discussion

As described in the above section, the clinically important region in the input original image is cropped using the Bounding box for extracting foreground and background. The foreground part is taken as ROI and this ROI itself classified again as Optimal ROI (O-ROI) and Remaining Regions (RR). Then they are separately encoded using SPIHT and Vector Quantization encoding respectively. In the decoding process, the inverse process such as SPIHT decoding and Vector Quantization decoding are done separately and all the parts are merged accordingly for reconstructing the decompressed image with the original image size. While using SPIHT, a biorthogonal wavelet bior 5.5 used for decomposing the image (O-ROI in our case). This method is experimented with different types of wavelets such as haar, Daubechies wavelets, Symlets, Coiflets and other biorthogonal wavelets and finally the bior 5.5 yields better results than others and therefore, it is considered as a default wavelet for decomposition. The compression results of our proposed method on randomly selected sample images are depicted in Fig. 5. In Fig. 5, the images are the original medical images (Image 1: MR brain image; Image 2: MR-T2-weighted brain image; Image 3: MR T1-weighted brain image; Image 4: MR ST1R Knee image).

The corresponding decompressed images are given in the next column of Fig. 5. This proposed method gives better reconstruction quality of decompressed images. From the Fig. 5, it can be easily identified the performance of SPIHT and the lossy coding namely vector quantization. The P-ROI get a higher degree of reconstruction quality while other portions get slight distortion. Thus, our proposed method efficiently encodes the medical images and the numerical assessments of the proposed method and the comparative results are presented as follows.

The numeric results of the proposed method are given in Table .1 for the randomly selected images from various imaging modalities given in Figs. 5 and 6 shows the Proposed PSNR and CR values. As mentioned in earlier, the evaluation metrices such as CR, BPP and the quality metric PSNR are recorded for the tested medical images.

The tested medical images of various types are given in the first column of Table .1. The second column comprises the PSNR value of which the compression results are recorded. The corresponding CR and BPP values of the proposed method are given in the remaining columns.

It is evident from the Table .1 that the proposed method gives high degree of CR values and the probable very low BPP values at high quality of image reconstruction that indicating by the obtained PSNR range. Since the CR and BPP values are seems to give a lossy performance, this proposed method yields better compression outcomes accomplished with the high quality of reconstruction.

7. Conclusion

In this paper, SPIHT based ROI codec for medical image compression is proposed. The negation of the background (in many cases, zero intensity pixels) during the encoding process tremendously reduces the bit rate. Usage of SPIHT encoder in the O-ROI part is done its task very well and it gives dominating performance compared to the existing techniques. As SPIHT is an embedded coder, our method also inherited the property of embedded coding, so, the decoder can be truncated at any point when it meets the required visual quality and hence increases the accuracy of diagnostics process.

Declaration of competing interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

References

[1] Jiantao Zhou, Xianming Liu, C. Oscar, Yuan Yan Tang, Desigining an efficient image encryption - then - compression system via prediction error clustering and random permutation, IEEE Trans. Inf. Forensics Secur. 19 (1) (2014) 39–50.
[2] Ying Li, Khalid Sayood, Lossless sequence compression using adaptive prediction, IEEE Trans. Image Process. 16 (4) (2007) 997–1007.
[3] Nishat Kanvel, S. Letitia, Elwin Chandra Monie, Adaptive lifting based image compression scheme with particle swarm optimization technique, Int. J. Eng. Technol. 2 (9) (2010) 4886–4895.
[4] PrabhjeetKaur, Parminder Singh, Intergrated particle swarm optimization and genetic algorithm based compression for reduction of blocking artifacts, Int. J. Comput. Sci. Inf. Technol. 6 (5) (2015) 4520–4527.

[5] P.V. Kranthi Kumar, M.S.R. Naidu, An automated threshold selection using wavelet based Particle swarm optimization for image compression, IOSR J. Electron. Commun. Eng. 9 (5) (2014) 26–31.
[6] Dania Brunello, Giancarlo Calvagno, Gian Antonio Mian, Roberto Rinaldo, Lossless Compression of video using temporal information, IEEE Trans. Image Process. 12 (2) (2003) 132–139.
[7] Jacob Ziv, The Universal LZ77 compression algorithm Is essential optimal for individual finite length N-Blocks, IEEE Trans. Inf. Theor. 55 (5) (2009) 1941–1944.
[8] Christopher H. Messom, Gourab Sen Gupta, N. Serge, Hough Transform run length encoding for real time image processing, IEEE Trans. Instrum. Meas. 56 (3) (2007) 962–967.
[9] Junhui Hou, Lap-Pui Chau, Minqi Zhang, Nadia Magnenat-Thalmann, Ying He, A highly efficient compression framework for time varying 3D facial Expressions, IEEE Trans. Circ. Syst. Video Technol. 24 (9) (2014) 154–1553.
[10] Anmin Liu, Weisi Lin, Manoranjan Paul, Fan Zhang, Chenwei Deng, Optimal compression plane for efficient video coding, IEEE Trans. Image Process. 20 (10) (2011) 2788–2799.
[11] Yin-Tsung Hwang, Ming - Wei Lyu, Cheng-Chen Lin, A low complexity embedded compression codec design with rate control for high definition video, IEEE Trans. Circ. Syst. Video Technol. 25 (4) (2015) 674–687.
[12] Yun Chin Li, Yung Chang Chen, A Hybrid model based image coding systems for very low bit rate coding, IEEE Trans. Selected Areas Commun. 16 (1) (1998) 28–41.
[13] M. Penedo, W.A. Pearlman, P.G. Tahoces, M. Souto, J.J. Vidal, Region based wavelet coding methods for Digital Mammography, IEEE Trans. Med. Imag. 22 (10) (2003) 1288–1296.
[14] Peter Schelkens, Adrian Munteanu, Jan Cornelis, Waveletbased compression of medical images: protocols to improve resolution and quality scalability and region-of-interest coding, Elsevier Future generation computer 15 (1999) 171–184.
[15] Jonathan Taquet, Claude Labit, Hierarchical oriented predictions for resolution scalable lossless and near lossless compression of CT and MRI biomedical images, IEEE Trans. Image Process. 21 (5) (2012).
[16] Zhongmin Liu, Zixiang Xiong, Qiang Wu, Yu-Ping Wang, Kenneth Castleman, Cascaded Differential and wavelet compression of chromosome images, IEEE (Inst. Electr. Electron. Eng.) Trans. Biomed. Eng. 49 (4) (2002) 372–383.
[17] Zuo-Dian Chen, Ruey - Feng Chang, Wen - Jia Kuo, Adaptive predictive multiplicative auto regressive model for medical image compression, IEEE Trans. Med. Imag. 18 (2) (1999) 181–184.
[18] L.I. Zeng, Christian P. Jansen, Stephen Marsch, Michael Unser, R Hunziker Patrick, Four Dimensional wavelet compression of Arbitary sized Echocardiographic data, IEEE Trans. Med. Imag. 21 (9) (2002) 1179–1187.
[19] V. Sanchez, R. Abugharbieh, P. Nasiopoulous, Symmetry based scalable lossless compression of 3D medical image data, IEEE Trans. Med. Imag. 28 (7) (2009) 1062–1072.
[20] Vibha Tiwari, P. Bonsod, Abhay Kumar, Performance evaluation of various compression techniques on medical images, Int. J. Adv. Electron. Commun. Syst. 1 (2) (2012) 1–10.
[21] Tim Bruylants, Adrain Munteanu, Peter Schelkens, Wavelet based volumetric image compression, Elsevier Signal processing: Image Commun. 31 (2015) 112–133.
[22] Tirupathiraju Kanumuri, M.L. Dewal, R.S. Anand, Progressive medical image coding using binary wavelet transform, Springer Signal, Image and Video processing 8 (5) (2014) 883–899.
[23] C. Doukas, I. Maglogiannis, Region of interest coding techniques for medical image compression, IEEE Eng. Med. Biol. Mag. 26 (5) (2007) 29–35.
[24] A. Gersho, R.M. Gray, Vector quantization and signal compression, in: Vector Quantization and Signal Compression, 2011, https://doi.org/10.1007/978-1-4615-3626-0.
[25] Y. Linde, A. Buzo, R.M. Gray, An algorithm for vector quantizer design, IEEE Trans. Commun. 28 (1) (1980) 84–95, https://doi.org/10.1109/TCOM.1980.1094577.
[26] T. Lehn-Schiøler, A. Hegde, D. Erdogmus, J.C. Principe, Vector quantization using information theoretic concepts, Nat. Comput. 4 (1) (2005) 39–51, https://doi.org/10.1007/s11047-004-9619-8.
[27] A. Jamil, M. Majid, S.M. Anwar, An optimal codebook for content-based image retrieval in JPEG compressed domain, Arabian J. Sci. Eng. (2019), https://doi.org/10.1007/s13369-019-03880-0.
[28] J.M. Shapiro, Embedded image coding using zerotrees of wavelet coefficients, IEEE Trans. Signal Process. 41 (12) (1993) 3445–3462, https://doi.org/10.1109/78.258085.
[29] A. Said, W.A. Pearlman, A new, fast, and efficient image codec based on set partitioning in hierarchical trees, IEEE Trans. Circ. Syst. Video Technol. 6 (3) (1996) 243–250, https://doi.org/10.1109/76.499834.

FOREX Publication
Open Access | Rapid and quality publishing

International Journal of
Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

Human Emotion Recognition using Deep Learning with Special Emphasis on Infant's Face

Parismita Sarma[1*], Takrim UL Islam Laskar[2], Dankan Gowda V[3] and Ramesha M[4]

[1]*Assistant Professor, Department of Information Technology, Gauhati University, Guwahati, Assam, India, pari@gauhati.ac.in*
[2]*Department of Information Technology, Gauhati University, Guwahati, Assam, India, takrimulislam@gmail.com*
[3]*Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India, dankan.v@bmsit.in*
[4]*Assistant Professor, Department of Electronics and Communication Engineering, GITAM School of Technology, GITAM (DEEMED TO BE UNIVERSITY), Bengaluru, Karnataka, India, rameshmalur037@gmail.com*

***Correspondence:** Parismita Sarma; pari@gauhati.ac.in

ABSTRACT- This paper discusses a deep learning-based image processing method to recognize human emotion from their facial expression with special concentration on infant's face between one to five years of age. The work has importance because most of the time it becomes necessary to understand need of a child from their facial expression and behavior. This work is still a challenge in the field of Human Facial Emotion Recognition due to confusing facial expression that sometimes found in some of the samples. We have tried to recognize any facial expression into one of the mostly understood human mood namely Angry, Disgust, Fear, Happy, Sad, Surprise and Neutral. For this purpose, we have trained an image classifier with Convolutional Neural Network with Kaggle's Fer2013 Dataset. After the completion of the project, we achieved good accuracy in most of the prominent emotions by testing with 20 random images for each emotion.

Keywords: Convolution Neural Network, Emotion detection, Deep Neural Network, Haar Cascade Frontal Face Detector.

ARTICLE INFORMATION

Author(s): Parismita Sarma, Takrim UL Islam Laskar, Dankan Gowda V and Ramesha M;

Received: 17/11/2022; **Accepted**: 15/12/2022; **Published**: 20/12/2022;
e-ISSN: 2347-470X;
Paper Id: IJEER 1711-07;
Citation: 10.37391/IJEER.100466
Webpage-link:
www.ijeer.forexjournal.co.in/archive/volume-10/ijeer-100466.html

Publisher's Note: FOREX Publication stays neutral with regard to Jurisdictional claims in Published maps and institutional affiliations.

1. INTRODUCTION

Emotion recognition from facial expression is an important area of research now a days. Human are capable of recognizing emotion or mood of a person from his/her speech and facial expression. In recent years sentiment analysis research is progressing with the help of advanced technology like machine learning, deep learning and transfer learning [1]. This research field has many practical implications and basically used in the field of machine learning *i.e.*, Human Computer Interaction, systems relating to Driver Assistance, Cognitive Science *etc.*

Facial emotion recognition can be done using different techniques. A big dataset can help us to use dense network and it is one step above deep neural network that gives more accurate result compared to other machine learning techniques. In general, for all techniques face based model is used where different features of human faces are extracted for emotion detection.

1.1 Deep Neural Network

Neurons are the fundamental units used for construction of neural networks. A number of neurons form a layer are linked to the neurons of subsequent layer.

Together they form the data path which travels throughout the whole network. Every neuron acts upon a mathematical estimation, transmit the result to all the connected neurons [2]. Deeper the layers, the nodes of the last few layers become more learned. One of the special designs of Artificial Neural Network used in present days for training is CNN (Convolutional Neural Network). The layers in CNN can be assumed as a type of feature extractor. This network exhibits extra ordinary strength to any type of classification.

1.2 Emotion Classification

Human being poses different types of emotions. We are emotional by habit, starting from infants to older ones, we express our emotions by shaping different parts on our face like eyes, chick, chin etc. In this work we have tried to classify mostly known seven human emotions to seven groups. We are trying to focus on correct emotion recognition for the toddlers as they are more expressive from their facial expression. But there are a few more emotions that are a bit difficult to recognize, for example doubting and confusing state of a mind. Accuracy of these emotions shows better result when using both tune of their speech and facial expression. That is why in this work we are considering only seven mostly understandable emotions. More emphasize is put on to detect child's emotions from their facial expressions, we have filtered our dataset accordingly. The samples of children's facial images are of mainly age group between one to five. The emotion classes which we have considered are 'angry', 'disgust', 'fear', 'happy', 'sad', 'surprise', 'neutral' also shown in the *table 1*.

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

Table 1: Classification of Emotions

Classes	Emotions
0	Angry
1	Disgust
2	Fear
3	Happy
4	Sad
5	Surprise
6	Neutral

1.3 Motivation

The primary objective of this study is to develop a neural network-based intelligent system capable of identifying children's emotions. Most of the time human expresses feelings spontaneously that causes automatic change in shape and size of different semi movable parts of our face like chick, eyebrow, tongue, lips, eye size *etc*. A precise recognition system greatly helps doctors and parents to understand physical and mental health of the child. According to surveys 7% meaning of a message is understandable through verbal statement, 38% is understood from vocal analysis and most amazingly remaining 55% is attained from facial expression [2].

1.4 Objective

One of the primary goals of this study is to correctly recognize human precisely toddler's emotions from real time videos and photographs. We have used here Deep Convolution Network; input frames are taken from real-time video as input image. This study uses the kaggle fer 2013 labelled dataset. We are using this dataset to classify above mentioned seven emotional classes.

1.5 Contribution of this Work

This system has potentiality in health care sector like computerized scrutiny of mental and physiological state of children or adult by recognizing facial expressions in real-time. Toddler express their emotions mostly by their facial expression. They cry when they are hurt. Though child's facial expression is real to their feelings, sample collection is challenging task to collect these samples. Yet with the limited dataset we were able to get good accuracy for each emotion class but there is scope to improve. Several methods for recognising human emotions have been explored in this paper. For the purpose of creating the FER method, we analysed the FER approaches and other processes. Based on this body of research, it is clear that a combination of cutting-edge methods is required to build an autonomous facial expression detection mechanism that would be both efficient and accurate.

2. LITERATURE SURVEY

To formulate and achieve maximum accuracy we have studied a number of recent journal papers and decided to use deep learning as proposed architecture. A short overall description is mentioned here. Researchers are focusing on new architectures which are mostly based on CNN and Deep learning. Among them a researcher AndreeaPascu under guidance of Prof. Ross King found a high accuracy of 86% using machine learning technique [3]. From source [4] we come to know that that image structure and quality of the samples influence the recognition rate. They experimented with different deep learning techniques and CNN constructions like VGG-16 and ResNet50, Support Vector Machine (SVM) as the classifier. They achieved only 31.8% accuracy with KDFF datasets. Later they were able to improve it more by combining the output of two neural network layers and accuracy was 67.2% on kaggle dataset and 78.3% for KDFF dataset [4]. In the research paper [5] by H. Qin et al. a new approach is observed where a system is proposed and was trained jointly with R-CNN and RPN. At first stage they trained Risk Priority Number (RPN) then adjust the model. ROIs are generated from RPN training phase and used for training the joint network with R-CNN. A precision of 98.22% is achieved in this model. In Another paper [6] by N. Zeng et. al. a new approach for facial expression recognition is discussed. Three descriptors HOG (Histogram of Oriented Gradients), LBR (Local Binary Pattern) and gray values were used in that approach. By using Deep SparseAuto-encoders on CK+ (Extended Cohn-Kanade) dataset anger, disgust and contempt expression were evaluated with 99-99% accuracy. But for fear and sadness it was 86-87%, happy and surprise mood exhibited most 98-100% accuracy. According to paper [7] by M.H Siddiqi et al. curvelet transform can be used to extract different features from digital images and they have used it in their proposed work. The system is able to extract the curves from different prominent features of human face taken from digital images. This concept is used for image reconstruction purpose also. They used HMM for labeling different human expressions and achieved accuracy of about 99%. A paper by S.K.AKamarol, et al. [8] implemented a method called Spatio Temporal Texture Map (STTM) which was able to construct a three dimensional (3D) texture map. According to them this method is much better than any other look-based feature extraction techniques like CK+, AFEW dataset known for lesser computational cost. By using sensors, 3D facial recognition is able to record the face's contours with greater accuracy. Unlike more conventional approaches, 3D face recognition scans may be performed in low- or no-light conditions without sacrificing accuracy.

3. SYSTEM SPECIFICATION

3.1 Tools

In this work we have used python 2.7 which is an open-source high level language. We are using Open-source computer vision library (open CV) that is free for academic purpose. It takes help of multi core processing and written in C/C++. Open CV has interface of java, python and C++ [9].

Tensor flow is open source software libraries that are basically used through a series of tasks for dataflow purpose [10]. It is the mostly used and perfect math library for deep neural network. It can run on multiple CPUs as well as GPUs. It can be used in mobile operating system also. For recording of real time images our system has in built camera attached with the laptop. Keras and Jupiter Notebook are two tools used in this work.

3.2 Dataset

As already mentioned, we have used here fer2013.csv (Kaggle Facial Expression Recognition Dataset) consisting of 48x48

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

size 35888 digital images of different faces in gray scale. There are 28,709 photos in the training set, and 3,582 in the testing set [11].For classification purpose we have used Haar-cascade classifier. It is an object detection algorithm mostly used in machine learning. This strategy involves the training of a cascade function for identification of objects from many constructive and negative images. This algorithm is known for its capability to detect faces and body components from digital images. In the first stage Haar-cascade algorithm needs huge amount of positive images (with faces) and negative images (not faces) for training purpose. The link to the dataset is https://www.kaggle.com/c/challenges-in-representation-learning-facial-expression-recognition-challenge/data.

3.3 Environment

Google Colaboratory is a free interactive environment based on the Jupyter notebook, which we are using. It runs on cloud and does not need any set up [12]. It allows to write code and executes in computer languages like python. It has facility to run GPUs for massive computations as training dataset amid Deep learning.

4. METHODOLOGY

Our proposed system will identify seven basic emotions. The prominent stages of the whole system are as follows:

- First stage is image capturing. We are using this tool for real time human emotion detection. So, the frame taken from real-time video is input image.
- Then we have used Haar Cascade frontal face detector and it will detect emotion of the face plotted in the image. In this case Region of Interest (ROI) is the face itself.
- The dataset used here is kaggle fer2013 labelled dataset. It is used to train a data model with CNN that classifies the above-mentioned dataset into seven emotional classes. This model is now usable for detection of emotion of the face in the image.

After classification is completed probability of detection for individual emotion classes is plotted. The class with highest percentage is tagged as resultant emotion. In *figure 1* we can see a block schematic of the system we propose.

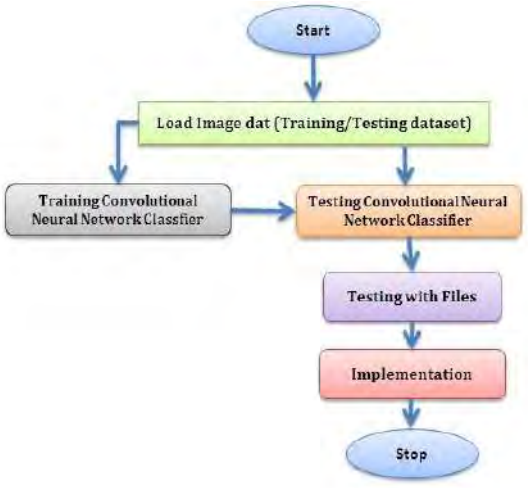


Figure 1: Schematic Representation of the Proposed System's Main Components

4.1 Loading Image Dataset

According to our work plan, the first phase consists of training and testing that performs on our repository. The image dataset is uploaded to the google drive. A jupyter notebook has been created in the Google colab with Python 3. Google colab can access these data from google drive as it is authorized for the same. The dataset used here is in form of three columns: Emotion, Pixel and Usage. The Emotion column has values from 0 to 6, it represents seven emotion classes. Pixel column correspond to pixel values of the 48 X 48-pixel images. Last the usage column gives the purpose of every data's individual instances. The rows have pixel values of one instance for one image. The *figure 2* shows a snapshot of emotion column and pixel values.

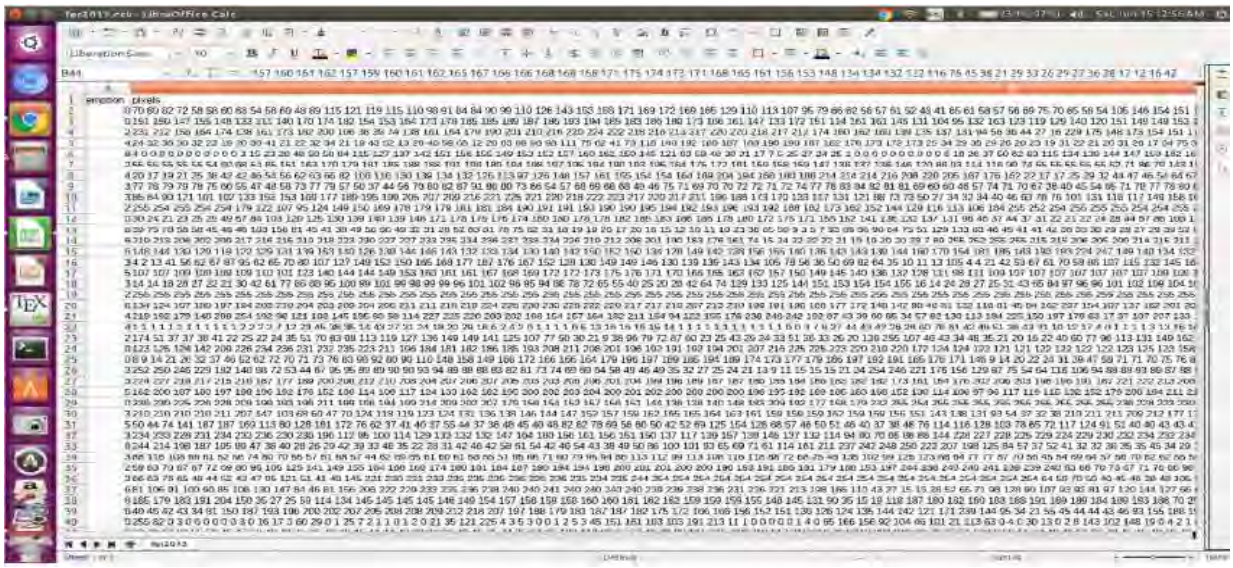


Figure 2: Snapshot of Emotion and Pixel of the Dataset

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

4.2 Storing Data in .csv File

All the data stored in fer2013.csv file are read and stored in an array 'content'. Next it is converted to numpy array and stored into a new array 'lines'. We have calculated the number of occurrences. Next the length of occurrence is calculated. The strategies we have used are

Arrays x_train, y_train, x_test, y_test and all are initialized. For each instance at a time the line is split into emotion class, image data and usage as shown below.

emotion, img, usage = lines[*i*].split(",")

Then the pixel values are split with reference to an empty space " " and stored in an array "val" and converted to numpy array "pixels" that is shown below.

val = img.split(" ")
pixels = np.array(val, 'float32')

The emotion is then categorised to its class with keras. Similarly, for testing dataset emotion is appended to y test and pixels to x test. y test.append(emotion)

x test.append(pixels)

The above assigned values are then converted to numpy arrays with 'float32' data type x_train. x_test pixel values are divided by 255 and reassigned the respective values to normalize input between [0,1]. For Training dataset, the emotion class is appended to the y_train and respective pixels are appended to x_train.

y_train.append(emotion)
x_train.append(pixels)

4.3 EPOCHS and BATCHES

We have taken 25 epochs to train the dataset. One Epoch means a complete process of passing the dataset through the neural network. We have assigned 256 batch size to each epoch. That means each epoch will divide the whole dataset into batches with each batch processed 256 instances.

4.4 Constructing Convolutional Neural Network Structure

In the sequential CNN model structure, there are four layers [13]. It is called sequential because the model has to know the proper input shape it is supposed to expect. The specification of each layer is as follows:

1st Layer -The first layer added is 2D CNN layer with 64 filters/kernals, kernel size (5, 5), activation function 'relu' and input shape 48 X 48 in grayscale. Then we added an average pooling layer of pool size (5,5) and stride size (2,2).
2nd Layer - The second layer is added with 64 filters/kernals, kernal size (3,3) and activation function 'relu'. Then we added an average pooling layer of pool size (3,3) and stride size (2,2).
3rd Layer - We have increased the filters/kernels from 64 to 128 in the third layer. Then by using flatten operation, which have converted the array to have 1-D shape.
4th Layer - Next the final dense layer is added with 1024 inputs from all the previous neurons. We have added Dropout value 0.2 and activation function 'relu'. There are two such neurons. Final dense layer is added with the total number of classes and activation function 'softmax'.

Rectified Linear Unit (ReLU): If the activation function is given a negative input, it will return 0, but it will return the same result if the input is positive [14]. That is,

f(n) = max(0,x)

Softmax Activation Function: It produces multiple outputs for 1 input array [15]. This helps to build a model which can classify more than 2 classes.

4.4.1 Epochs, Batch Processing and Model Generation

ImageDataGenerator() Function has been initialized and data flow has been defined with x_train, y_train and batch size are assigned to variable 'train_generator'.
The model is then compiled to calculate the categorical_crossentropy (loss) and accuracy.

The model generation is then started with train generator as **model.fit_generator(train generator, steps per epoch=batch size, epochs=epochs)** with batch size and epochs as parameter.

The system has been run to obtain the trained model. The image of the training process has been shown in the *figure 3*.

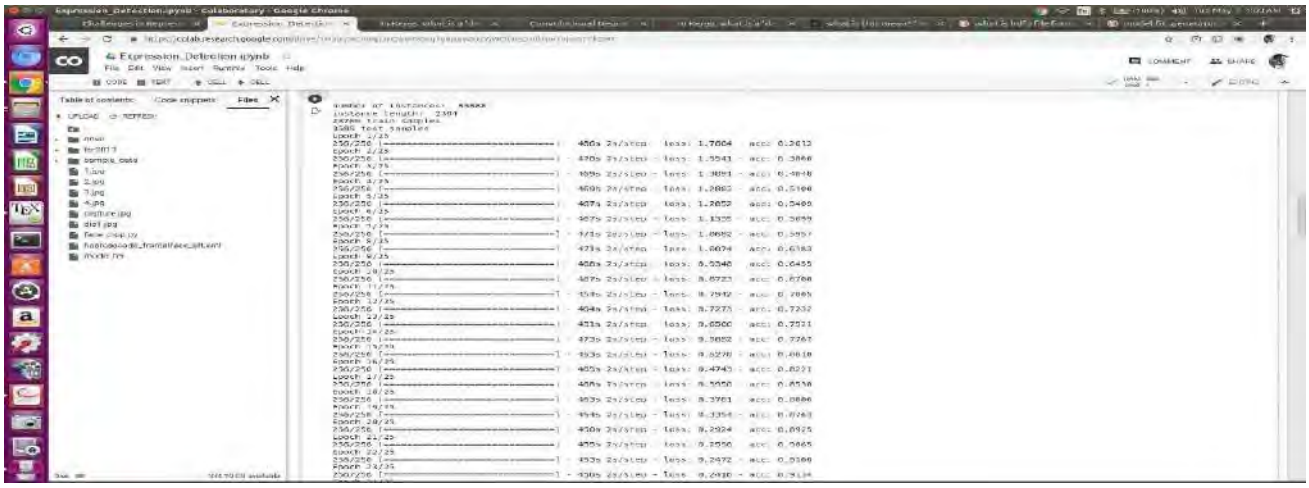


Figure 3: Training Data Set to Obtain a Model

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

The trained model is then saved in a Hierarchical Data Format file with extension "h5" as "model.h5". HDF5 / H5 (Hierarchical Data Format) is a file format designed to organize and store huge amount of data. It contains two major objects dataset and group.

- Datasets of multidimensional array with homogeneous type.
- Groups are the container structures which are capable of storing dataset and other groups.

4.5 Training Time on GPU and CPU

When we trained the image classifier, we tried different configurations. Firstly, we used only 56 CPUs and found that the time required was 3 hours and 15 minutes. Then we tried adding a GPU and the time consumption was only 4 minutes. And then the result of testing with 1 GPU and 1 CPU was shocking, as the time was 3 minutes, which was less by 1 (approx.) minute than that of the with 1 GPU and 56 CPUs. *Table 2* clearly describes this scenario. In presence of GPU if more CPUs are assigned process will consume more time. It is because more processes assigned to CPUs for parallel processing and as a result increases time taken by CPUs.

Table 2: Time required by CPUs and GPU

GPU	CPU	TIME (hh:mm)
0	56	03:15
1	56	00:04
1	1	00:03

4.6 Evaluation

Figure 4 shows the evaluation of train, test loss and accuracy of our system.

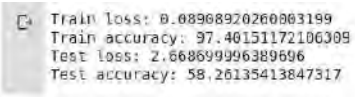


Figure 4: Evaluation of Testing, Training loss and Accuracy

4.7 Monitoring Test Set Results

Table 3 below shows the mapping of different emotion class to the Emotion Index used in our work.

Table 3: Mapping of Emotion Class to Emotion Index in the Array

Class number	Emotion	Emotion index in array
0	Angry	0
1	Disgust	1
2	Fear	2
3	Happy	3
4	Sad	4
5	Surprise	5
6	Neutral	6

For visualizing and monitoring the test set results, x_test (pixels) has been passed through predict function.

predictions = model.predict(x_test)

The pixel values from 20 to 30 instances have been retrieved and image have been reconstructed. *Figure 5* shows reconstruction of Twentieth test dataset and plotted along with the percentage of accuracy graph. The bar graph in the *figure 5* detects an angry face.

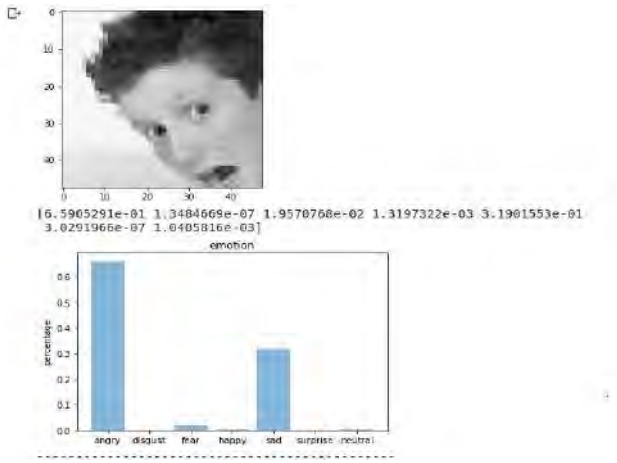


Figure 5: Images with Reconstruction of Twentieth Test Dataset

We have used Haar-Cascade frontal face detector for source image face detection [16]. More focus is given on Haar-like features compared to pixel intensity. This method considers adjacent rectangular region and search for total differences to classify the feature region (subsections) of the image [16].

Cascade classifier is initialized by loading the pre-trained cascade file 'haarcascade_frontalface_alt.xml' then passing through openCV cascade classifier. Next the image is passed through multi scale detector and hence all the faces in the image are detected. Faces are next bounded by rectangular boxes, cropped and stored. These cropped images are then converted to gray scale for emotion detection. *Figure 6* is the screen shot of a detected face with Haar_Cascade classifier.

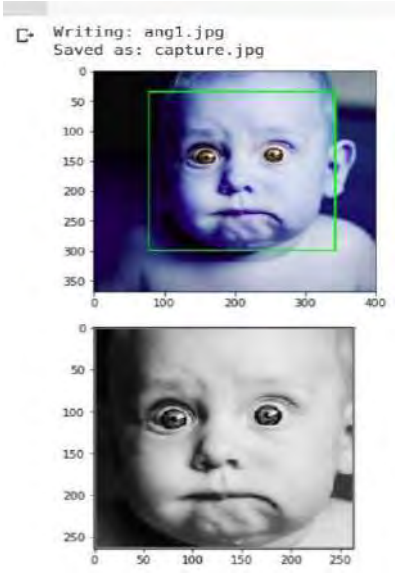


Figure 6: Face detection with Haar_Cascade classifier

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

4.8 Defining the Emotion Analysis Function

An array of emotions is created, for each emotion there is its index number that corresponds to emotion class number. That is:

objects = ('angry', 'disgust', 'fear', 'happy', 'sad', 'surprise', 'neutral')

4.9 Prediction of Emotions

The stored or real-time images are loaded in gray scale with 48 X 48 resolution and finally converted to a linear array containing all the pixel values. It is also made sure that no pixel value exceeds 255 bits.

This array is then passed through the prediction function and stored in an array 'custom'. This array 'custom' contains seven classes with the probability of every emotion. This is then passed to the emotion analysis function defined earlier for plotting using matplotlib.

5. RESULTS AND CONCLUSION

We have shown below a number of figures (screenshots) each one recognizing the appropriate emotion from the facial expression predicted from our system. From the analysis of the report, we have seen that Disgust and Fear produced accuracy less than 70% as these emotions resemble with each other most often. Disgust resembles with anger and sadness; fear resembles with surprise.

Model: "sequential_12"

Layer (type)	Output Shape
conv2d_60 (Conv2D)	(None, 44, 44, 64)
max_pooling2d_12 (MaxPooling	(None, 20, 20, 64)
conv2d_61 (Conv2D)	(None, 18, 18, 64)
conv2d_62 (Conv2D)	(None, 16, 16, 64)
average_pooling2d_24 (Averag	(None, 7, 7, 64)
conv2d_63 (Conv2D)	(None, 5, 5, 128)
conv2d_64 (Conv2D)	(None, 3, 3, 128)
average_pooling2d_25 (Averag	(None, 1, 1, 128)
flatten_12 (Flatten)	(None, 128)
dense_36 (Dense)	(None, 1024)
dropout_24 (Dropout)	(None, 1024)
dense_37 (Dense)	(None, 1024)
dropout_25 (Dropout)	(None, 1024)
dense_38 (Dense)	(None, 7)

Figure 7: Screenshot of classification model with output parameter

Some of our output samples with snapshots are shown in *figure 7*, to *figure 18*. The bar graph shown in every emotion is predicted emotion class, the highest valued emotion class is the resultant one. From those bar graphs we can conclude that figure… shows the angry emotion, figure… for fear, figure … for happy, figure … for sad, figure … for surprise and figure…. for neutral emotion. All the screenshots are shown below.

5.1 Classification Report

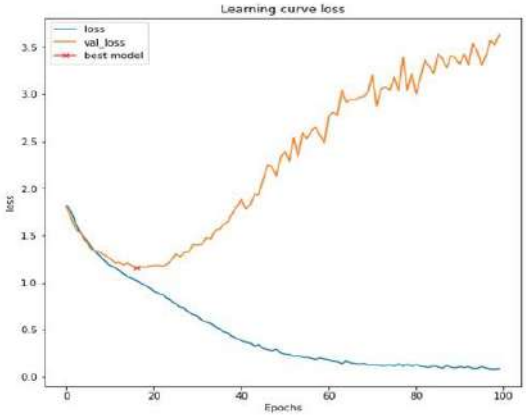


Figure 8: Learning curve showing the best model

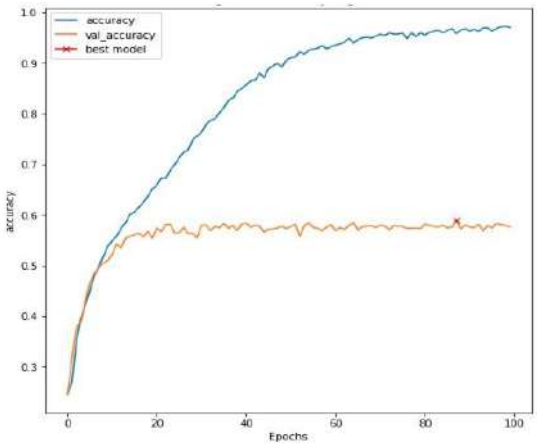


Figure 9: Accuracy curve of the best model

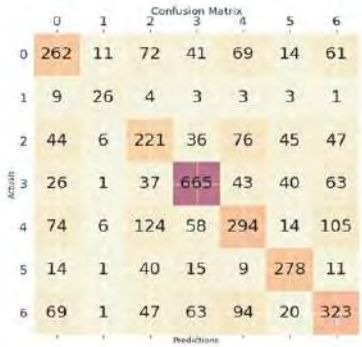


Figure 10: Confusion Matrix

	Precision	recall	f1-score	support
0	0.49	0.53	0.51	498
1	0.53	0.50	0.51	52
2	0.47	0.41	0.43	545
3	0.76	0.75	0.76	881
4	0.44	0.50	0.47	588
5	0.76	0.67	0.71	414
6	0.52	0.53	0.53	611
accuracy			0.58	3589
macro avg	0.57	0.56	0.56	3589
Weighted avg	0.58	0.58	0.58	3589

Figure 11: Classification Report

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

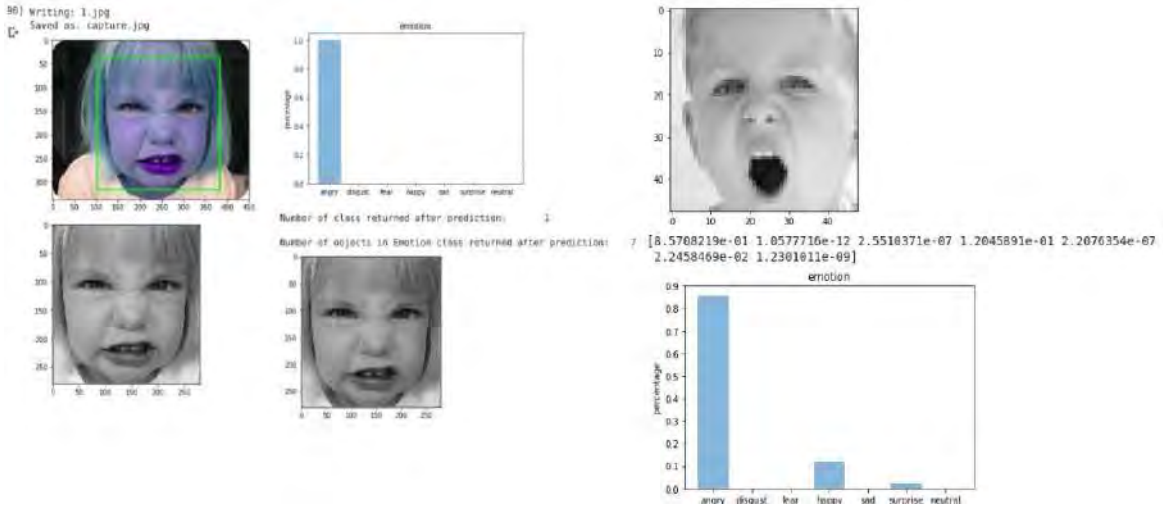


Figure 12: Output of random test for angry emotion

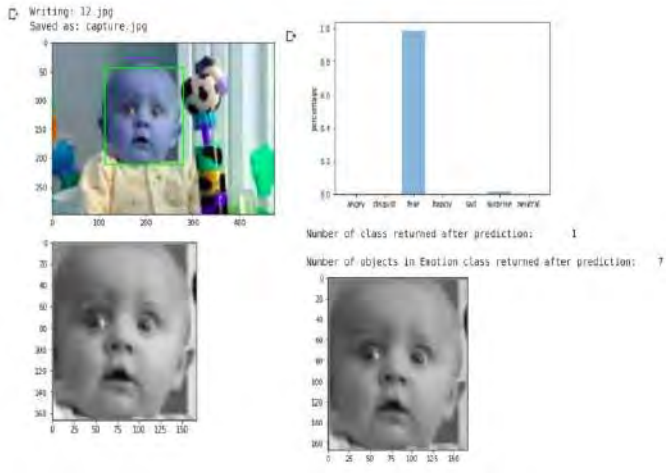


Figure 13: Output of random test for Fear emotion

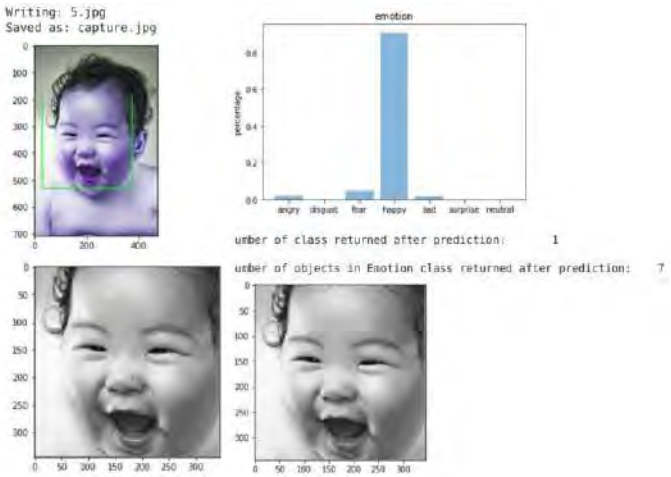


Figure 15: Screenshot of Happy emotion

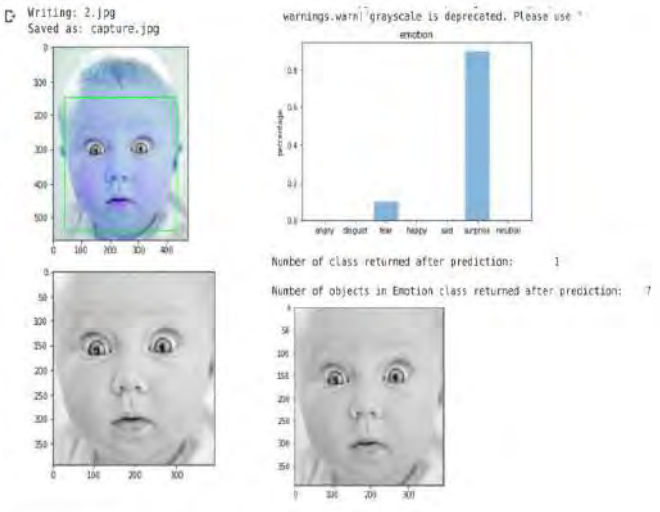


Figure 14: Output of random test for Surprise emotion

Figure 16: Screenshot of sad emotion

FOREX Publication
Open Access | Rapid and quality publishing
International Journal of Electrical and Electronics Research (IJEER)
Research Article | Volume 10, Issue 4 | Pages 1176-1183 | e-ISSN: 2347-470X

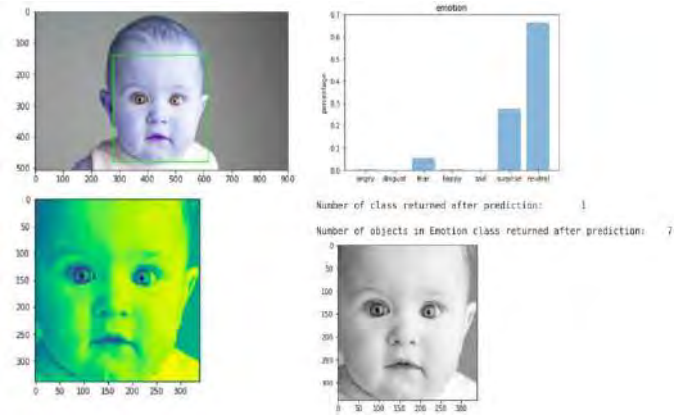


Figure 17: Screenshot of Neutral Emotion

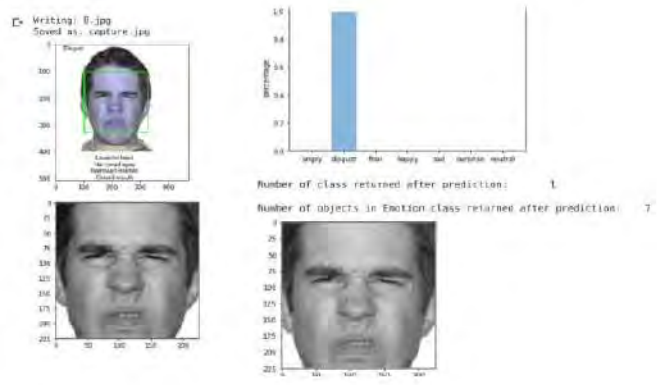


Figure 18: Screenshot of Disgust emotion

5. CONCLUSION

In this research work, more emphasis is given to emotion classification in the context of facial images, speech, and brain signals. Several deep learning models are evaluated, including Convolutional and Recurrent neural networks. An evaluation clearly shows that deep learning models are superior to traditional models for image and signal classification. This work focuses on developing a system that improves the accuracy of Deep Learning models for emotion recognition. Experimental results for each technique are analysed to highlight the most significant aspects of each approach.

REFERENCES

[1] Avinash Sharma, and A. Azhagu Jaisudhan Pazhani, A novel method of data compression using ROI for biomedical 2D images, Measurement: Sensors, Volume 24, 2022, 100439, ISSN 2665-9174.

[2] Mayur Rahul, Namita Tiwari, Rati Shukla, Devvrat Tyagi and Vikash Yadav (2022), A New Hybrid Approach for Efficient Emotion Recognition using Deep Learning. IJEER 10(1), 18-22. DOI: 10.37391/IJEER.100103.

[3] Shashank M Gowda and H N Suresh (2022), Facial Expression Analysis and Estimation Based on Facial Salient Points and Action Unit (AUs). IJEER 10(1), 7-17. DOI: 10.37391/IJEER.100102.

[4] I. Goodfellow, D. Erhan, P.-L. Carrier, A. Courville, M. Mirza, B. Hamner, W. Cukierski, Y. Tang, D. Thaler, D.-H. Lee, Y. Zhou, C. Ramaiah, F. Feng, R. Li, X. Wang, D. Athanasakis, J. Shawe-Taylor, M. Milakov, J. Park, R. Ionescu, M. Popescu, C. Grozea, J. Bergstra, J. Xie, L. Romaszko, B. Xu, Z. Chuang, and Y. Bengio, "Challenges in representation learning: A report on three machine learning contests," Available at: http://arxiv.org/abs/1307.0414, Accessed on 11 March, 2021.

[5] T. Nath, A. Mathis, A. C. Chen, A. Patel, M. Bethge and M. W. Mathis "Using DeepLabCut for 3D markerless pose estimation across species and behaviors", Nature protocols, Vol. 14, No. 7, pp. 2152-2176, 2019.

[6] H. Sadr, M. M. Pedram and M. Teshnehlab "A robust sentiment analysis method based on sequential combination of convolutional and recursive neural networks", Neural Processing Letters, Vol. 50, No.3, pp.2745-2761, 2019.

[7] M. Khan, S. Chakraborty, R. Astya & S. Khepra "Face Detection and Recognition Using OpenCV", proceedings in 2019 International Conference on Computing, Communication, and Intelligent Systems (ICCCIS), IEEE, pp. 116-119, 2019.

[8] Z. Qin, H. Ye, G. Y. Li and B. H. F. Juang "Deep learning in physical layer communications", IEEE Wireless Communications, Vol. 26, No. 2, pp. 93-99, 2019.

[9] X. Jiang, Y. Pang, X. Li, J. Pan and Y. Xie "Deep neural networks with elastic rectified linear units for object recognition", Neurocomputing, Vol. 275, pp. 1132-1139, 2018.

[10] Agarap, A. F. "Deep Learning using Rectified Linear Units (ReLU", Cornell University, arXiv preprint arXiv: 1803.08375, 2018.

[11] N. Zeng, H. Zhang, B. Song, W. Liu, Y. Li, and A. M. Dobaie, "Facial expression recognition via learning deep sparse autoencoders", Neurocomputing, Vol. 273, pp. 643-649, 2018.

[12] A. Savoiu and J. Wong, "Recognizing facial expressions using deep learning," Stanford University, 2017.

[13] S. K. A. Kamarol, M. H. Jaward, J. Parkkinen, and R. Parthiban, "Spatiotemporal feature extraction for facial expression recognition", IET Image Processing, Vol. 10, No. 7, pp. 534-541, 2016.

[14] M. Rabiei and A. Gasparetto, "System and method for recognizing human emotion state based on analysis of speech and facial feature extraction; applications to human-robot interaction," in 2016 4th International Conference on Robotics and Mechatronics (ICROM). IEEE, pp. 266-271, 2016.

[15] Pascu, A., Intelligence, B. A., & King, R. "Facial Expression Recognition System", University of Manchester, 2015.

[16] D. Lapakko "Communication is 93% nonverbal: An urban legend proliferates", Communication and Theater Association of Minnesota Journal, Vol. 34, No. 1, 2007.

Bilingual text detection from natural scene images using faster R-CNN and extended histogram of oriented gradients

Industrial and Commercial Application Published: 06 April 2022 **25**, 1001–1013 (2022)

Pattern Analysis and Applications

Aims and scope

Submit manuscript

Alex Noel Joseph Raj, Chen Junmin, Ruban Nersisson, Vijayalakshmi G. V. Mahesh & Zhemin Zhuang

Abstract

In today's world, scene text detection is important for a wide range of scientific and industrial processes. Compared with text detection in documents, text detection in natural scenes is challenging since they are subjected to different orientations, scaling, brightness variations, and complex backgrounds. Further scenes can contain multiple scripts which limits the performance of detection algorithms. In this paper, we propose a state-of-the-art algorithm for text detection for a bilingual natural scene dataset. The framework consists of

(a) Faster R-CNN employed to extract probable text regions within the scene images, (b) rearrangement of the text region as consecutive frames along the time axis and extraction of global and local shape features from the three orthogonal planes and (c) use of simple and effective classifier to predict the features extracted from regions as text or non-text region. The proposed frame when compared to other text detection techniques improves the overall text detection accuracy. The validity of the algorithm is verified on the bilingual text detection dataset MSRA-TD500, and a promising F1 score of 0.70 is reported.

References

1. Yan C, Xie H, Liu S, Yin J, Zhang Y, Dai Q (2017) Effective Uyghur language text detection in complex background images for traffic prompt identification. IEEE Trans Intell Transp Syst 19(1):220–229

Article Google Scholar

2. Jung K, Kim KI, Jain AK (2004) Text information extraction in images and video: a survey. Pattern Recognit 37(5):977–997

Article Google Scholar

3. Zheng Y, Li Q, Liu J, Liu H, Li G, Zhang S (2017) A cascaded method for text detection in natural scene images. Neurocomputing 238:307–315

Article Google Scholar

4. Zhang Z, Shen W, Yao C and Bai X (2015) Symmetry-based text line detection in natural scenes. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 2558–2567

5. Zhang X, Zhang Z, Zhang C and Bai X (2016) Symmetry-based object proposal for text detection. In: 2016 23rd international conference on pattern recognition (ICPR), IEEE, pp 709–714

6. Ma J, Shao W, Ye H, Wang L, Wang H, Zheng Y, Xue X (2018) Arbitrary-oriented scene text detection via rotation proposals. IEEE Trans Multimed 20(11):3111–3122

Article Google Scholar

7. Jiang Y, Zhu X, Wang X, Yang S et al (2017) R2cnn: rotational region cnn for orientation robust scene text detection.*arXiv preprint* arXiv:1706.09579

8. Zhong Z, Jin L, Zhang S and Feng Z (2016) Deeptext: a unified framework for text proposal generation and text detection in natural images. *arXiv preprint arXiv:1605.07314*

9. Girshick R, Donahue J, Darrell T and Malik J (2014) Rich feature hierarchies for accurate object detection and semantic segmentation. In: Proceedings of the IEEE conference on computer vision and pattern recognition, pp 580–587

10. Girshick R (2015) Fast r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 1440–1448

11. Ren S, He K, Girshick R and Su, J (2015) Faster r-cnn: towards real-time object detection with region proposal networks. In: Advances in neural information processing systems, pp 91–99

12. He K, Gkioxari G, Dollár P and Girshick R (2017) Mask r-cnn. In: Proceedings of the IEEE international conference on computer vision, pp 2961–2969

13. Shao F, Wang X, Meng F, Zhu J, Wang D, Dai J (2019) Improved faster R-CNN traffic sign detection based on a second region of interest and highly possible regions proposal network. Sens (Basel, Switz) 19(10):2288. https://doi.org/10.3390/s19102288

Article Google Scholar

14. Maheshwari K, Joseph Raj AN, Mahesh VG, Zhuang Z, Rufus E, Shivakumara P, Naik GR (2019) Bilingual text detection in natural scene images using invariant moments. J Intell & Fuzzy Syst 37(5):6773–6784

Article Google Scholar

15. Bosch A, Zisserman A and Munoz X (2007) Representing shape with a spatial pyramid kernel. In: Proceedings of the 6th ACM international conference on Image and video retrieval, pp 401–408

16. Fan X, Tjahjadi T (2015) A spatial-temporal framework based on histogram of gradients and optical flow for facial expression recognition in video sequences. Pattern Recognit 48(11):3407–3416

Article Google Scholar

17. Yao C, Bai X, Liu W, Ma Y et al (2012) Detecting texts of arbitrary orientations in natural images. In: 2012 IEEE conference on computer vision and pattern recognition, IEEE, pp 1083–1090

18. Liu X, Meng G, Pan C (2019) Scene text detection and recognition with advances in deep learning: a survey. Int J Doc Anal Recognit (IJDAR) 22(2):143–162

Article Google Scholar

19. Kumuda T and Basavaraj L (2015) Detection and localization of text from natural scene images using texture features. In 2015 IEEE international conference on computational intelligence and computing research (ICCIC), IEEE, pp 1–4

20. Jain AK, Yu B (1998) Automatic text location in images and video frames. Pattern Recogn 31(12):2055–2076

Article Google Scholar

21. Jaderberg M, Simonyan K, Vedaldi A, Zisserman A (2016) Reading text in the wild with convolutional neural networks. Int J Comput Vision 116(1):1–20

Article MathSciNet Google Scholar

22. Pan YF, Hou X, Liu CL (2010) A hybrid approach to detect and localize texts in natural scene images. IEEE Trans Image Process 20(3):800–813

MathSciNet MATH Google Scholar

23. Liu Y, Goto S, Ikenaga T (2006) A contour-based robust algorithm for text detection in color images. IEICE Trans Inf Syst 89(3):1221–1230

Article Google Scholar

24. Simonyan K and Zisserman A (2014) Very deep convolutional networks for large-scale image recognition. *arXiv preprint arXiv:1409.1556*

25. Nair V and Hinton GE (2010) Rectified linear units improve restricted boltzmann machines. In: Proceedings of the 27th international conference on machine learning (ICML-10), pp 807–814

26. Chorowski JK, Bahdanau D, Serdyuk D, Cho K et al (2015) Attention-based models for speech recognition. In: Advances in neural information processing systems, pp 577–585

27. Felzenszwalb PF, Girshick RB, McAllester D, Ramanan D (2009) Object detection with discriminatively trained part-based models. IEEE Trans Pattern Anal Mach Intell 32(9):1627–1645

Article Google Scholar

28. Kumbhar P, Mali M (2016) A survey on feature selection techniques and classification algorithms for efficient text classification. Int J Sci Res 5(5):9

Google Scholar

29. Dalal N and Triggs B (2005) Histograms of oriented gradients for human detection. In: 2005 IEEE computer society conference on computer vision and pattern recognition (CVPR'05), vol 1, IEEE, pp 886–893

30. Grauman K and Darrell T (2005) The pyramid match kernel: discriminative classification with sets of image features. In: Tenth IEEE international conference on computer vision (ICCV'05) volume 1, vol 2, IEEE, pp 1458–1465

31. Zhao G, Pietikainen M (2007) Dynamic texture recognition using local binary patterns with an application to facial expressions. IEEE Trans Pattern Anal Mach Intell 29(6):915–928

Article Google Scholar

32. Larose DT, Larose CD (2014) Discovering knowledge in data: an introduction to data mining. John Wiley & Sons, USA

MATH Google Scholar

33. Yin XC, Yin X, Huang K, Hao HW (2013) Robust text detection in natural scene images. IEEE Trans Pattern Anal Mach Intell 36(5):970–983

Google Scholar

34. Epshtein B, Ofek E and Wexler Y (2010). Detecting text in natural scenes with stroke width transform. In: 2010 IEEE computer society conference on computer vision and pattern recognition, IEEE, pp 2963–2970

35. Chen H, Tsai SS, Schroth G, Chen DM et al (2011) Robust text detection in natural images with edge-enhanced maximally stable extremal regions. In: 2011 18th IEEE international conference on image processing, IEEE, pp 2609–2612

Funding

Shantou University, NTF17016, Alex Noel Joseph Raj , Department of Science and Technology, Ministry of Science and Technology, INTRUSRFBR382, Ruban Nersisson. Basic and Applied Basic Research Foundation of Guangdong Province, No. 2020B1515120061, Zhemin Zhuang.

Author information

Authors and Affiliations

Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Department of Electronic Engineering, College of Engineering, Shantou University, Shantou, 515063, China
Alex Noel Joseph Raj, Chen Junmin & Zhemin Zhuang

School of Electrical Engineering, Vellore Institute of Technology, Vellore, 632014, India
Ruban Nersisson

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India
Vijayalakshmi G. V. Mahesh

Corresponding author

Correspondence to Ruban Nersisson.

Ethics declarations

Conflict of interest

The authors state that there is no conflict of interest/competing interest with anything or anybody.

Additional information

Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Rights and permissions

Reprints and Permissions

About this article

Cite this article

Joseph Raj, A.N., Junmin, C., Nersisson, R. *et al.* Bilingual text detection from natural scene images using faster R-CNN and extended histogram of oriented gradients. *Pattern Anal Applic* **25**, 1001–1013 (2022). https://doi.org/10.1007/s10044-022-01066-3

Received
11 June 2020

Accepted
12 March 2022

Published
06 April 2022

Issue Date
November 2022

DOI
https://doi.org/10.1007/s10044-022-01066-3

Keywords

Text detection Faster region-convolutional neural network (R-CNN)

Histogram of oriented gradients (HOG) MSRA-TD500 Natural scene image

Deep learning

Design and Implementation of a High-Speed D Flip Flop using CMOS Inverter Logic

JAYADEVA G. S., NIKHIL MURALI, MEGHANA S., RAKSHA K. KUMAR, NITHIN ANIL NAIR
Deptartment of Electronics and Communication Engineering,
BMS Institute of Technology & Management,
Bangalore,
INDIA

Abstract— This paper proposes an improvised D- flip flop configuration based on tristate inverter logic, which reduces the power dissipation, decrease the transition time from the input to output as well as reduced time to reach rail to rail voltage. The flip flop uses transmission gate instead of pass transistor to achieve this requirement. The design is simulated using 90nm CMOS technology and data is propagated at 50% duty cycle. The circuit is simulated using Cadence tools to assess the performance with respect to delay and power. These D-flip flops have numerous applications such as buffers, registers, digital VLSI clocking systems, microprocessors etc.

Key-words— D Flip flop, Cadence, Power consumption, Propagation delay.

Received: June 17, 2021. Revised: October 8, 2022. Accepted: November 11, 2022. Published: December 19, 2022.

1 Introduction

In the field of Very Large Scale Integrated (VLSI) circuits, sequential circuits are expansively used and serve a vital role in digital logic design. The purpose of recent circuit designs is to retain or improve the characteristics feature necessary for VLSI industry, while reducing power consumption. There are two major types of power dissipation viz, static and dynamic. The dynamic power dissipation is mainly due to switching activity of the load capacitor. Internal leakage is caused in devices when it is not in the working condition which constitutes the static dissipation [6,7].
Various attempts have been made to reduce power dissipation using different technologies, with different topologies. Specifically, the dynamic power. Among them CMOS technology with single edge triggering flip-flop topology is popular [1,2]. Further the devices kept on scaling to reduce individual device power dissipation. At the same time the density of components in a chip increases to accommodate more functionality, which increased the power dissipation in a given chip area.

The CMOS Technology node, in which we essentially deal with the physical dimensions of the device can accommodate more transistors in the given area and also can switch faster and consume less power compared with other technologies, hence require less energy and the chip runs at lower temperature.
The power dissipation and time delay variations depend on various parameters such as supply voltage, aspect ratio, oxide thickness, load capacitance, threshold voltage, which vary as the technology scales down.

In a given CMOS circuit power dissipation can be from three main sources. Namely, Signal transition, short circuit current flowing from supply to the ground terminal, and leakage currents. The short circuit power becomes comparable to the dynamic power dissipation as the technology is scaled down. Furthermore, the leakage power also plays a hugely important role in a digital CMOS circuit. As we further reduce the channel length, gate oxide thickness and threshold voltage in CMOS circuits, the significant contributor to power dissipation turns out to be the high leakage current [3][4][5]. Therefore, to estimate and reduce leakage power, relevant modelling and accurate identification of different leakage components is principally important in high-speed and low-power applications. The load capacitance is reduced from 10fF to 1fF to reduce switching capacitance and power dissipation.
CMOS technologies can be used to describe various systems in which CMOS flip-flops are extensively used such as personal computers,

servers, other portable systems etc. This work simulates modified D-ff using 90nm technology with single edge triggering flip-flop topology node on Cadence tools using CMOS technology. The earlier work on D-ff maintains functionality up to 100 Mhz [8] and 1 GHz [3]. Our modified D-ff with single edge trigger works up to 2.5 GHz with much less power dissipation, uses only 1V power supply.

In the design of an analog circuit, generally the power dissipation increases with increase in frequency of operation. The aim of this work is to show that with the introduction of transmission gate instead of Pass transistor in D-flip flop (D-ff) results in less power dissipation even with higher frequency of operation.

2 Related Work

The power dissipation can be reduced by reducing capacitance and the supply voltage in CMOS. This may result in the low performance of device. For low power consumption and low power dissipation, various models have been proposed with different topologies and their effect varies on power dissipation and propagation delay [1][2], one such topology is the Single Edge-Triggered Flip-Flop (SET).

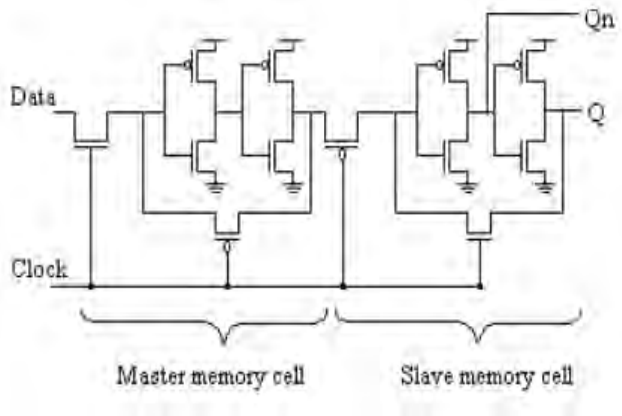


Fig. 1: SETFF Configuration

The advantages of SET flip flop are:

- SET has the fastest delay among any flip flop considered along with large amounts of negative set up time.
- SET uses the simplest flip flop design wherein the sampling data is used only on one clock edge.
- SET shows better results with respect to power dissipation.
- SETs are the best lower power flip flops.

An improvised version of the SET (Single Edge Triggered) Flip Flop is proposed in this paper which includes a resettable functionality. The proposed design uses tristate inverters instead of pass inverters as shown in Fig.1. In transmission gate (TG) arrangement, the combination of both a PMOS and NMOS not only counteracts reduced noise margin, but also decreases switching resistance and static power dissipation which is due to increased threshold voltage, but requires the clock signal and its complement. The proposed design aims to reduce the power dissipation due to both leakage and switching action, while keeping the propagation delay in check.

In the published D-ff design using SET flip flop [1], the pass transistors have higher leakage current, high switching resistance which results in high power dissipation. which can be minimized using transmission gates. Also, transmission gates give better noise margin. This results in improved performance in frequency of operation and power dissipation.

3 Methodology

3.1 CMOS Inverter Technology

The basic working of the project resembles the functionality of a D-ff. To reduce the waste of energy in flip-flops with low data activity, SET flip-flops has been developed. The proposed flip flop has the least power dissipation among all the designs for low switching activities

Fig. 2: Flip flop block diagram

The Delay(D) flip-flop is a clocked flip-flop which has a single input ‘D’. The output follows the state of ‘D’ each time a D-ff is clocked. D and Clock are the two inputs for D flip flop.

The given Fig. 3 shows the timing diagram of the proposed design of the D-ff.

Fig. 3: Timing diagram

As seen in Fig.3, data is passed from the master stage to the slave stage when clock is “0”, and output is seen across Q when clock is “1”.

3.2 Proposed Design

The High-Speed D-ff is proposed to be designed as shown in the Fig.4

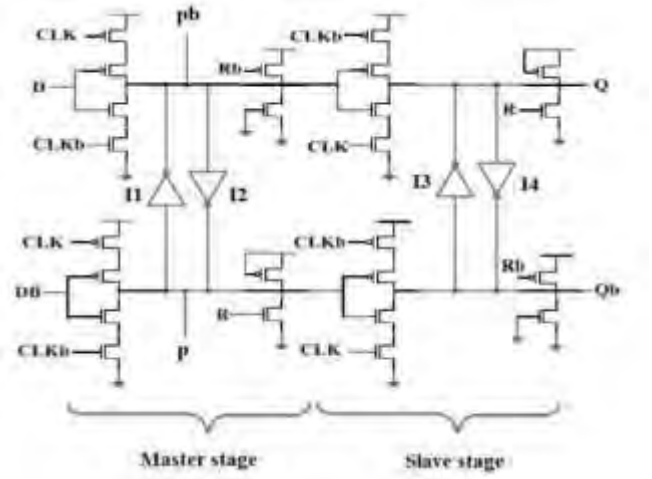


Fig. 4: Proposed design

CLK and CLKb are the clock signals which are complements of each other. D and DB are the inputs, complements of each other. R and Rb are reset signals, complements of each other. Q and Qb represent the outputs.

When CLK is '0', left stage tristate inverters (master stage) are ON and right stage tristate inverters (slave stage) are off. The master stage tristate inverter has been used to implement a switch. At this point, outputs Q and Qb hold the previous value. When CLK rises to '1', the master stage tristate inverters are turned OFF and the slave stage tristate inverters are turned ON, therefore the data is transferred from the master stage to the output on the positive edge of the clock. Meanwhile, the master stage holds the previous value using the back-to-back inverters I1 and I2. So, at the rising edge of CLK, the input data (D) is sent to the output (Q).

The back-to-back inverters act as a memory latch, storing the data when the respective stage is inactive. The reset functionality is added to the D-ff which deactivates the respective master/slave block in that clock signal.
Fig.5 represents the test bench, which include the Design Under Test (DUT) block,

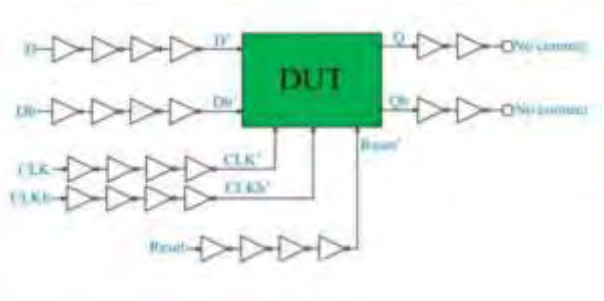


Fig. 5: Test bench

The 4 stage inverters act as buffers for the input and 2 stages at the output signals. The inverters can be used to increase driving strength by changing the strength of the transistors.

4 Results and Discussion

The proposed structure of the positive edge triggered D-ff constructed with the help of Cadence tool. Fig. 6 shows the sole DUT of the D-ff, and Fig.7 denotes the testbench along with its respective input and output environments.

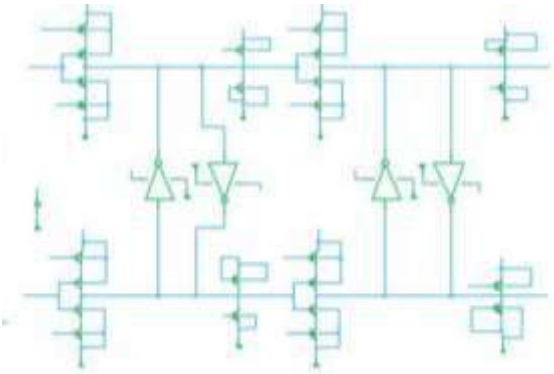


Fig. 6: Schematic of D-ff in cadence

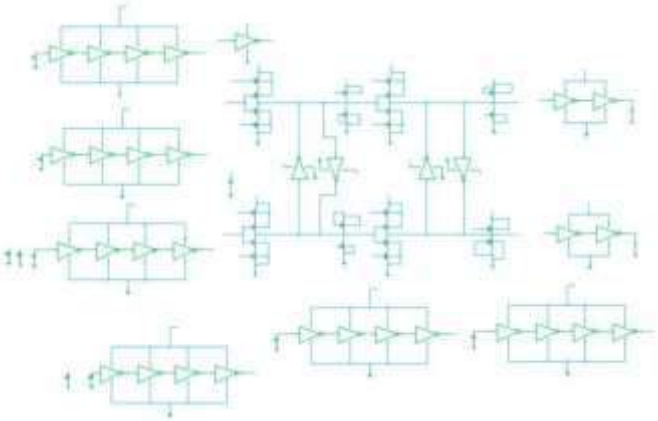


Fig. 7: Schematic of test bench

The pass transistor SET design [1] shown in Fig.1 yielded the following results in Cadence Virtuoso. Fig 8 shows the waveform of the referred SETFF design using 1Ghz clock frequency.

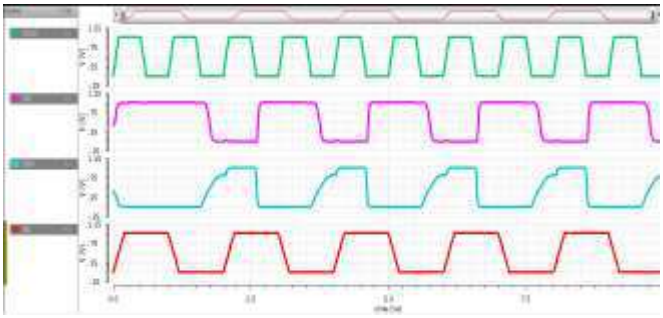


Fig. 8: Waveform of referred design at 1 GHz [1]

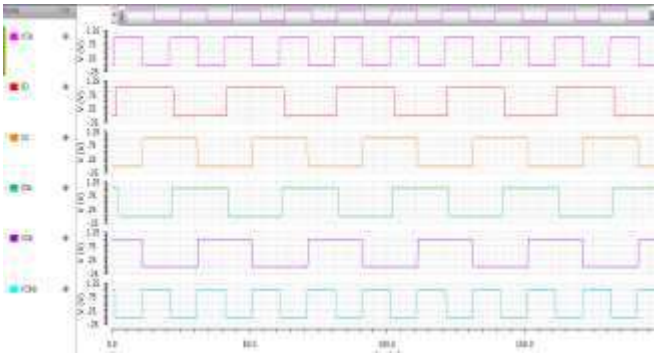


Fig. 9: Waveform of proposed design at 1 GHz

From the Table 1, we see that, the proposed design and the earlier work [1] shows similar propagation delay, but in the proposed design, the power dissipation is reduced approximately by 50% at 1 Ghz frequency. The output waveform shown in Fig. 8, gives the output obtained from re-simulating the design [1], which shows the further degradation as the clock frequency increases above 1 GHz.

Table 1. Performance Comparison of Referred and Proposed designs

90nm, 1V 1Ghz Clock frequency	**Referred Design**	**Proposed Design**
Clock to Output Delay(Tcq)(s)	1.074n	1.076n
Avg Power(uW)	42.32	22.38

The proposed design shows stable results in 2.5Ghz clock frequency as well. The test bench is simulated in nominal voltage temperature conditions i.e., 25°C, 1V as stipulated for 90nm CMOS technology. The functionality is observed in given Fig.10.

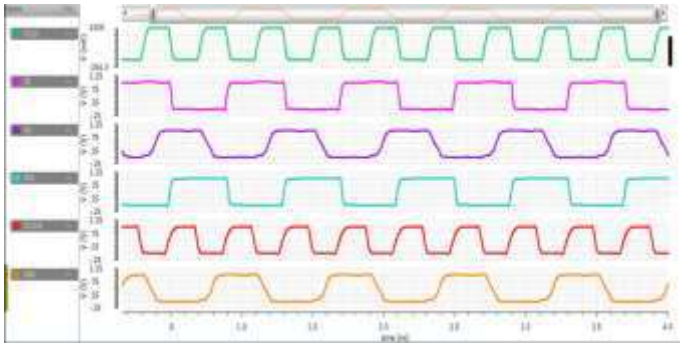


Fig. 10: Functionality check at 2.5 GHz

Table 2. Performance of Proposed Design at Higher Clock Frequencies

Clock Frequency (2.5 GHz) 90nm, 1 V, 25°C		
Signals	**Rise Time (ps)**	**Fall Time (ps)**
D	26.45	26.45
Clk (Clock)	42.99	42.99
Q	72.47	74.05
Delay(Tcq)(s)	475.8 p	
Average power (W)	50.73 u	

5 Conclusion

In this paper we implemented and constructed an improvised low power design using transmission gate instead of pass transistors for implementing D-ff. The results obtained from the implementation show that there is an improvement in stability in higher frequency of operation along with 50% lower power dissipation at 1Ghz frequency. Further, the frequency of operation is enhanced up to 2.5 Ghz giving stabilised output with only an average power dissipation of 50.73uW. The results obtained is promising for sequential circuits. The results were verified using Cadence Tools with 90nm CMOS Technology.

The proposed D-ff design using transmission gate concept can be extended to other sequential circuits and can be further analysed, verified for similar improvement in frequency and power dissipation

References:

[1] Rishikesh V. Tambat and Sonal A.Lakhotiya, "Design of Flip-Flops for High Performance VLSI Applications using Deep Submicron CMOS Technology", International Journal of Current Engineering and Technology, Vol.4, No.2 (April 2014)

[2] Jahangir Shaikh and and Hafizur Rahman, "High speed and low power preset-able modified TSPC D flip-flop design and performance comparison with TSPC D flip-flop", International Symposium on Devices, Circuits and Systems (ISDCS), IEEE Xplore 2018

[3] Uming Ko, and Balsara P.T., "High-performance energy-efficient D-flip-flop circuits," IEEE Transactions on Very Large-Scale Integration Systemstion Systems, Vol. 8, Issue 1, Feb. 2000.

[4] Johnson T.A., Kourtev I.S., "A single latch, high speed double-edge triggered flip-flop (DETFF)," The 8th IEEE International Conference on Electronics, Circuits and Systems, Volume 1, 2-5 Sept. 2001.

[5] B.Chinnarao, B.Francis & Y.Apparao, (2012), Design of a Low Power Flip-Flop Using CMOS Deep Submicron Technology, International Conference on Electrical and Electronics Engineering (ICEEE) -9th Sept, 2012, Guntur- ISBN: 978-93- 82208-21-1.

[6] Akash Agarwal, Tarun Kumar Gupta, Ajay Kumar Dadoria, "Ultra low power adiabatic logic using diode connected DC biased PFAL logic", Vol 55, pp.46-47, March 2017.

[7] Manoj Sharma and Arti Noor, "Positive Feed Back Adiabatic Logic: PFAL Single Edge Triggered Semi-Adiabatic D Flip Flop", African Journal of basic and Applied sciences 5(1):pp.42-46, 2013.

[8] Dahlan Abdullah," Design of Set Based D Flip-Flop for High-Speed Applications" Vol. 2 No. 2 (2020): JVCS

Contribution of Individual Authors to the Creation of a Scientific Article (Ghostwriting Policy)

-Dr. Jayadeva G. S. has supervised and advised the research activity planning and execution, including mentorship

-Nikhil Murali designed the Methodology and carried out the Simulation in Cadence Tools.

-Meghana S. has organized the Investigations and Visualization of the paper.

-Raksha K. Kumar carried out Management and Coordination responsibility for the research.

-Nithin Anil Nair was responsible for the Formal Analysis of the design.

Design and FPGA Implementation of High Throughput and Low Latency Machine Learning based Approximate Multiplier for Image Processing Applications

ANIL KUMAR D.
ECE department, BMS Institute of Technology and Management, Bengaluru,
INDIA

Abstract:- One of the uses of approximate circuits is machine learning (ML) and with the help of inexact logic minimization as well as through probabilistic pruning, these approximate computing circuits can be implemented. Nowadays, these approximate circuits have been widely explored due to their essential factors such as compact silicon areas well as low power consumption in movable devices. This research work shows how a 4:2 compressor can be designed using inexact logic minimization and thereby reversing a few bits of the output to ensure efficiency as well as accuracy. The average area, propagation delay as well as the average power of the proposed 4:2 compressor is been calculated and are employed in the 8 × 8 and 16x16 Dadda multiplier and truncation and rounding-based scalable approximate multiplier (TOSAM). Using Vivado Design Software Systems in 45nm technology, all the simulations were carried out and the MATLAB tool make use of error analysis to distinguish between precise as well as approximate proposed circuits. This work is mainly concentrated on the design of exact and approximate multipliers and measures the error between them and minimization of this error using the Machine Learning approach and finally validated the results on the Artix-7 FPGA development board of part XCA7CSG324_110t, the partial products which are generated by multipliers are added using 4:2 compressor adder. In the case of digital processing at nano-metric scales, approximate or inexact computing is considered one of the important examples. For computer arithmetic designs, inexact computation plays a significant role, and the new approximate 4:2 compressors are used in a multiplier that is based on TOSAM. These architectures mainly depend on various compression aspects to enable inaccuracy in computing which is described as error rate and is also referred to as normalized error distance which is used to satisfy circuit-based figures of merit, the number of transistors, delay as well as consumption of power. For a Dadda multiplier, four distinct approaches for exploiting the suggested approximation compressors are designed as well as evaluated. The usage of approximate multipliers for image processing, as well as a wide range of simulation results, are presented in this work. When contrasted to an accurate design, the proposed designs achieve a substantial reduction in the number of transistors, power dissipation as well as delay. Furthermore, the presented multiplier models exhibit outstanding image multiplication capabilities in terms of average normalized error distance as well as peak signal-to-noise ratio that is more than 50dB for the analyzed image samples. The proposed ML-based digital system has been developed in Vivado Design Suite and synthesized which is designed using Verilog HDL. Based on obtained results, 17% reduction in power, 21% reduction in latency, and 33% improvement in throughput.

Key-words: Machine learning, 4:2 Compressor based adder, Dadda Multiplier, TOSAM and FPGA

Received: June 21, 2021. Revised: April 22, 2022. Accepted: May 21, 2022. Published: July 5, 2022.

1 Introduction

In multiplier designs, compressors have been considered the major elements for partial product reduction. Previously, various adders were utilized for partial product reduction and one of such adders is carry-save adders. However, these adders were substituted by different order compressors like 3:2, 4:2, as well as 5:2 because of their low power requirement as well as their compact structure. Compact size, as well as reduced power consumption, are the essential factors that are considered for the development of portable devices application. As a result, approximate computation must be used in digital systems to achieve the required power. Approximate computing has received considerable attention whenever precision is

not important in applications such as signal as well as video processing hardware, accelerators that are used in a specific application, machine learning, and so on. 4:2 compressors that consist of XOR-XNOR circuits, as well as transmission gate multiplexers with complemented as well as un-complemented outputs, were presented by Veeramachaneni et al [1]. The research work in [2] illustrates various XOR-XNOR circuits with varying numbers of transistors and each of these transistors overcomes the drawback of another. The suggested XOR-XNOR circuit is being used in compressors such as 4:2 as well as 5:2. The work in [3] utilized a transmission gate multiplexer along with CMOS logic as well as a buffer to increase the current-carrying capacity that is to be used in a 4:2 compressor. The idea of approximate computing has been developed in VLSI architecture because of the increase in the application of portable devices as well as the reduction in power consumption and supply voltage. Many of the applications make use of approximate computing which is illustrated in [4]. Approximate compressors were introduced in [5] and were used in the Discrete Cosine Transform. These compressors used spintronic devices which include magnetic domain wall motion stripe as well as magnetic channel junction. Probabilistic pruning, as well as inexact logic minimization, are the two approximate functional equivalence relaxation approaches introduced by Avinash et al., and further works are presented in [6]. The work proposed in [7] produced only 50% of the output with approximate compressors when contrasted to the inputs. Whereas, the output weights produced were the same as the input weights with the proposed approximate compressor. The research work suggested in [8] illustrates two 4:2 compressors with design1 and design2 showing 20 as well as 12 correct outputs out of 32 outputs and 16 outputs respectively. Dadda multipliers employed this type of compressor and error analysis was carried out by using the applications of image processing.

Error distance (ED), mean error distance (MED), as well as normalized error distances (NED), are some of the error metrics that are presented in [9-10] and these metrics can be employed in the compressor as well as multiplier error analysis. An inaccurate multiplier with reduced performance measures such as area, time as well as power can achieve desired precision by making use of a tool which is presented in [11] wherein a user may specify the maximum error rate that the circuit can tolerate. The work presented in [12] describes the use of machine learning on existing approaches of CAD to develop unique devices. In [13], an approach for automatically synthesizing approximate circuits using Verilog code was presented, which was later turned into a tool called ABACUS. A mechanism for automatically generating approximate circuits was introduced in [14] and was accomplished through the use of a unit in the ASIC design process, which resulted in significant power, area, as well as latency reductions. The majority of applications of computer arithmetic are accomplished using digital logic circuits which offer high-level accuracy as well as reliability. Multimedia, as well as image processing, are some of the applications that can tolerate computation errors and inaccuracies and thereby generating significant as well as effective output. Algorithms and techniques that are accurate as well as reliable are often not desirable or effective for using these applications. Consider an instance that whenever developing an energy-efficient system, the inexact computation approach mainly concentrates on reducing accurate as well as fully predictable structures. This enables approximate computation to reconfigure the current design process of digital circuits, systems and thereby resulting in reduced complexity, and cost, as well as a probable increase in power efficiency and performance. When contrasted to accurate (precise) logic circuits, approximate or imprecise computing mainly depends on the fact to develop approximate circuits that operate at high performance and/or reduced power consumption [1]. In computer arithmetic, operations such as addition as well as multiplication are commonly employed. In the case of approximate computation, full-adder cells have been extensively examined for addition operation [2-4]. The research work in [1] examined various adders and also introduced several new models for measuring approximate as well as probabilistic adders in terms of integrated figures of merit for inexact computing applications. By taking into account the averaging impact of several inputs and the normalization of multiple-bit adders, the mean error distance (MED, as well as normalized error distance (NED), were presented. Since this normalized error distance is generally constant with size, it can be used to examine the reliability of a certain model. In the work presented in [1] the difference between accuracy as well as power has been critically

examined. The architecture of approximate multipliers on the other hand has received minimal consideration. Multiplication can be performed as the repetitive addition of partial products. Furthermore, the usage of approximate adders to develop an approximate multiplier is not feasible due to inefficiency in terms of accuracy, computational complexity, as well as other performance measures. In the research work of [4] [5] [6] [7], various approximate multipliers have been presented. The majority of these approaches employ the truncated multiplication technique in which the least significant columns of partial products are approximated as a constant. In the case of neural network applications, an approximate array multiplier is being used and is presented in [4] where some of the least significant bits in partial products are excluded and thereby eliminate some of the adders in an array. The work in [5] introduces a truncated multiplier along with a correction constant. To minimize error distance, the correction constant n+k bits are chosen which is found to be as close as possible to the estimated sum of these errors. Whenever the partial products are all 1's or all 0's in the n-k least significant columns then a truncated multiplier with constant correction will be having a maximum error. In the work presented in [6,] a variable correction truncated multiplier is been introduced which is based on column n-k-1 and this technique modifies the corrective term. The correction term is increased as well as decreased only when all the partial products in column-k-1 are one and zero respectively.

To develop large multiplier arrays, a reduced and imprecise 2x2 multiplier block is been introduced in [7]. Compressors are extensively employed in the design of fast multipliers which are presented in [8-10] to enable faster partial product reduction tree and power consumption reduction. In [8, 11 - 16], 4-2 accurate compressors design that is optimized were suggested. Approximate multiplication compression is been addressed in [17] [18]. An approximate signed multiplier for arithmetic data value speculation (AVDS) was presented in [17] and the Baugh-Wooley algorithm is used for multiplication. Therefore, for approximate computation, no unique design for compressors is been suggested. In [18], estimated compressor models were designed and these models do not aim at multiplication. It is important to remember that the work presented in [7] is more efficient than that proposed in [17] [18] by implementing a reduced multiplier block that can handle imprecise multiplication.

In DSP algorithms, addition operations are most widely employed and this algorithm includes filters, transforms, as well as predictions. This type of digital signal processing algorithm is commonly used in audio as well as video processing mobile devices that are powered by batteries. Four additions can be carried out at the same by using an effective 4-2 adder compressor. Generally, the critical path, as well as internal glitching, is been reduced by this higher order of parallelism which further minimizes the dynamic power dissipation. Topologies that are based on two CMOS+ gates to minimize power, area, as well as the delay of the 4-2 adder compressor are proposed in this research work. Using the Cadence Virtuoso tool, the suggested CMOS+ 4-2 adder compressor circuit topologies are implemented and simulated at electric as well as layout levels by using 45 nm technology. When contrasted to the research work in [19], the proposed 4-2 adder compressor realization minimizes power by 22.41%, delay by 32.45% as well as area by 7.4% respectively [19].

Designers have been compelled to explore alternative sources of calculating efficiency because the advantages of technological scalability have been reduced. The architectures that are based on multicores, as well as heterogeneous accelerators, are the outcomes of this work which further enhances the computing efficiency by reducing the power budgets. We have also examined approximate computing which is considered one of the major areas of interest over the previous few years. The basic concept behind approximate computing is to produce efficient outcomes in terms of quality and these concepts must be accepted in several areas like designing an algorithm for networks as well as distributed systems The methodologies of precise computing have been developed from ad hoc as well as applications that are specific and are extensively validated by systematic design approaches. At last, the development of applications like identification, processing, discovery, data analytics, inference, as well as vision is significantly extending the possibilities for imprecise computing. A unified cross-layer architecture for approximate computing [20] is been presented in this research work and also describes the concepts as well as important ideas that have influenced our work in this field.

Approximate computing has been widely analyzed for digital signal processing applications to maximize

accuracy as well as to improve other circuit characteristics like power, area as well as efficiency. With the help of upcoming nanoscale spintronic devices, we propose an approximate arithmetic circuit in this work. Primarily we describe a hybrid spin-CMOS majority gate model that is based on a hybrid spintronic device structure that comprises of magnetic domain wall motion stripe as well as a magnetic tunnel junction and thereby exploits the advantage of the spintronic device's intrinsic current-mode thresholding function. We also present a majority gate-based compact as well as energy-efficient accuracy-configurable adder architecture. However, when compared to earlier approximate circuit designs, the proposed architecture adapts to the intrinsic robustness in a variety of applications to various degrees of precision which in turn were fixed to a constant degree of approximation in existing circuit designs. Two unique approximate compressors were presented in this work and fast multiplier systems employ these types of compressors. When contrasted to the recently developed domain wall motion that is based on full adder design, the simulation results of the proposed device-circuit SPICE indicate 34.58 percent and 66 percent reduction in power for precise as well as approximate modes of the accuracy-configurable adder. Furthermore, the suggested accuracy-configurable adder, as well as approximate compressors, can be extensively used in digital image processing algorithms and the discrete cosine transform (DCT). The final results demonstrate that inverse discrete cosine transform and discrete cosine transform that employs approximate multiplier accomplish approximately 2x energy as well as speed up 3x by maintaining comparable output quality when compared to that of the precisely designed circuit [21].

One of the latest trending concepts in digital design is approximate computing which reduces the necessity for exact computation to enhance efficiency as well as speed. This research work presents a unique approximate compressor as well as an approach for these compressors to design effective approximate multipliers. With the help of the 40-nm library, approximate multipliers for multiple operand lengths have been synthesized using this approach. The suggested circuits offer enhanced speed and power for a target precision when compared to that of the existing approximate multipliers. This research work also includes image filtering as well as adaptive least mean squares filtering applications [22]. For digital processing at nanometric scales, inexact computing is considered one of the desirable techniques. In the case of computer arithmetic designs, inexact computing plays a major role in this work. Design, as well as analysis of two unique approximate 4-2 compressors that are employed in a multiplier, are the main objectives of this research work. These architectures mainly depend on various compression aspects to enable inaccuracy in computing which is described as error rate and is also referred to as normalized error distance which is used to satisfy circuit-based figures of merit, the number of transistors, delay as well as consumption of power. For a Dadda multiplier, four distinct approaches for exploiting the suggested approximation compressors are designed as well as evaluated. The usage of approximate multipliers for image processing, as well as a wide range of simulation results, are presented in this research work. When contrasted to an accurate design, the proposed designs achieve a substantial reduction in the number of transistors, power dissipation as well as delay. Furthermore, the presented multiplier models exhibit outstanding image multiplication capabilities in terms of average normalized error distance as well as peak signal-to-noise ratio that is more than 50dB for the analyzed image samples [23].

To minimize the consumption of power of error-tolerant intrinsic applications, approximate computing is been used which has received considerable attention. An approach for partial product perforation is described in this research work for developing approximate multiplication which mainly concentrates on hardware-level approximation. We demonstrate that in partial product perforation, the induced errors are finite as well as predictable and are dependent on the input distribution in a precise mathematical approach. A partial product perforation approach is implemented for various multiplier designs as well as used for identifying effective architecture-perforation configuration combinations for varied error conditions by detailed experimental analysis. It is demonstrated that the suggested partial product perforation reduces power, area as well as critical delay by 50%, 45%, and 35% respectively when compared to the precise model. Furthermore, the product perforation technique outperforms the conventional approximation techniques like truncation, the voltage over scaling, as well as a logic

approximation concerning error and power dissipation [24].

2 Proposed ML-Based Low Power and High Throughput 4:2 Compressors Adder for Tosam Multiplier

Probabilistic logic minimization on a precise 4:2 compressor is utilized in this work. Here the minterms bit flipping of Boolean functions of Sum, Carry as well as *Cout* is computed to reduce operands and thereby minimize delay, area, and power consumption of the circuit. Various combinations have been examined and are found to be proportional to the number of bits that are flipped to determine the desirable bit flips. The complexity of the circuit should be reduced with a reduced error rate only after the implementation of the above procedure on any precise circuit. Since the most significant bit is very much important in achieving the outcome than the least significant bit, the lower order bits of the 4:2 compressor are flipped from one's to zero's and thus resulting in a 25% error rate (i.e., the number of inaccurate outputs to correct outputs) with no reduction in the number of inputs as well as outputs. The bits can be reversed from 'one to zero' or even from 'zero to one,' but 'one to zero' is carried out in this work to minimize the circuitry size. Carry is flipped by 4, Cout and SUM are given in Equation 1 and Equation 2 illustrates the k-maps of Sum, Carry, as well as Cout.

Two unique approximate 4-2 compressors are suggested as well as examined in this research work. When compared to the existing improved (accurate) 4-2 compressor designs illustrated in [8], the proposed simplified compressors show improved results in terms of delay, as well as power, consumed [8]. Further, the restoration unit of the Dadda multiplier makes use of the proposed approximate compressors. The approximate multiplication makes use of four distinct techniques. In the case of CMOS feature sizes like 32, 22 as well as 16 nm, the detailed simulation results are presented at the circuit level for figures of merit like delay, power dissipation, transistor count, and error rate, as well as normalized error distance. Further, the use of these multipliers in image processing is been discussed. The outcomes of two instances of image multiplication are presented. Furthermore, the presented multiplier models exhibit outstanding image multiplication capabilities in terms of average normalized error distance as well as peak signal-to-noise ratio that is more than 50dB for the analyzed image samples. The suggested approximate architectures for the compressor as well as multiplier are feasible alternatives for inexact computing as per the analysis and simulation findings. Minimizing n numbers to two numbers is the primary objective of multi-operand carry-save addition or parallel multiplication. As a result, n-2 compressors or n-2 counters are frequently employed in computer arithmetic. An n-2 compressor shown in Fig.1 is a circuit slice that reduces n integers to two numbers when correctly reproduced. The n-2 compressor accepts n bits in position as well as one or more carry bits from the places to the right i.e., $i - 1$ or $i - 2$, in slice i of the circuit. Two output bits in i and $i + 1$ places, as well as one or more carry bits in higher positions like $i + 1$, are generated. When compared to the precise compressor in [1] as well as approximated 4:2 compressors, the proposed architecture has more elements in the form of AND and OR gates. However, the precise compressor's XOR, XNOR, as well as MUX circuits use more transistors, which improves the structure of the multiplier and power consumption of this circuit as shown in Fig.2.

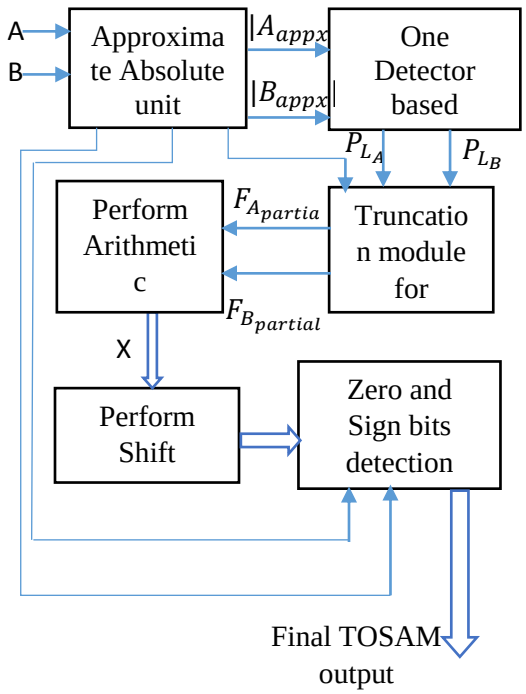


Fig. 1: Proposed design flow diagram of TOSAM for both signed and unsigned operations

In Fig.1, P_{L_A} and P_{L_B} are one detection-based position leadings, and $|A_{appx}|$ and $|B_{appx}|$ are truncated bits in both input operands. The $F_{A_{partial}}$ and $F_{B_{partial}}$ are fractional of both operands after performing truncation. X is the output of arithmetic operation i.e to perform $1 + F_{A_{partial_t}} + F_{B_{partial_t}} + F_{A_{appx}} x F_{B_{appx}}$

$$AxB \cong AxB_{appx} = 2^{P_{L_A}} x 2^{P_{L_B}} (1 + F_{A_{partial_t}} + F_{B_{partial_t}} + F_{A_{appx}} x F_{B_{appx}}) \text{---------(1)}$$

The remainder of Eq.(1) is stored as fractional part (F) of both operands A and B derived from Eq.(1). From the Eq.(1), the proposed 4:2 compressor for proposed TOSAM multiplier is shown in Fig.3.

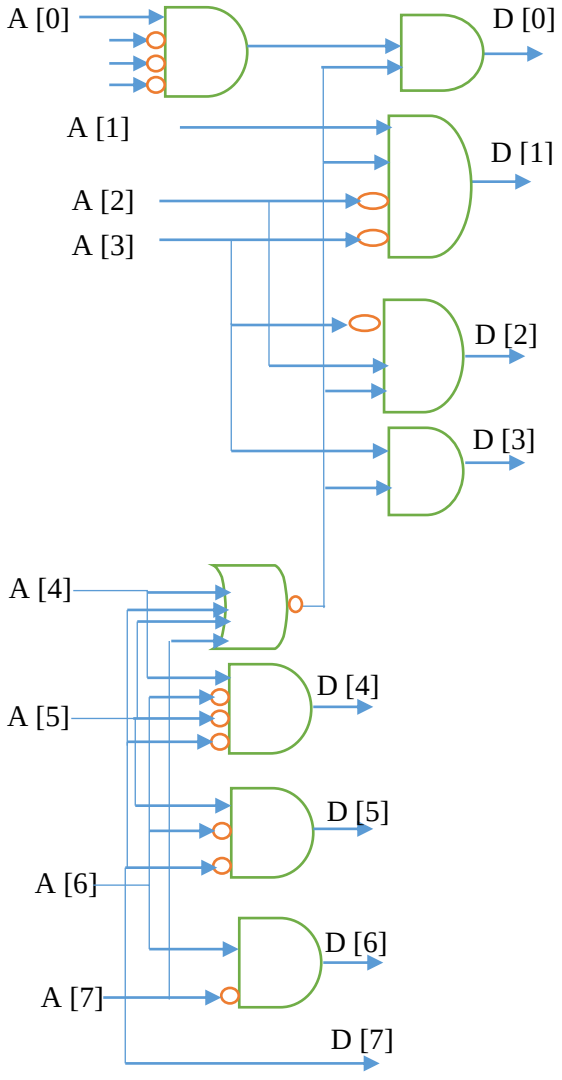


Fig. 2: Gate level schematic of one detector leading module for 8-bit input data.

The Fig.2, the sign is the output bit which will be set in line with the sign of the proposed multiplier operands and the zero-detector module set the output bits to zero when any bit of input is zero. In unsigned multipliers, the TOSAM should be neglected the sign and it replaced with a Zero detector module to optimize the delay and area. The suggested 4:2 compressor's power consumption, as well as the size of the circuit, have been lowered because of the application of modules such as AND, 2T MUX, OR, and 6T XOR & XNOR that are based on pass transistor logic, as well as the transistor count has been reduced to 34 from 52 [2] and 50 [1]. 25% is the error rate for the proposed design because the total number of bit flips is approximately 8 as shown in Fig.3. The error rate of design1 is 37.5 percent and the error rate of design 2 is 37.5 when compared to the work presented in [8]. In the suggested design error rate as well as design 2 [8] error rate, there are only four inputs due to the absence of carrying Cin. The suggested architecture, as well as the precise compressor, consists of the same number of inputs, but slightly higher transistor counts, as well as power, are consumed than the work in [8]. Dadda multiplier is the quickest multiplier among all other parallel multipliers. As a result, this approach is used in this research work to execute multiplication operations. In this paper, simple AND gates were utilized to generate the partial products for the 3 stages of multiplication, whereas Carry Save is utilized to generate the partial products for the second stage of multiplication as well as Carry Propagate Adders were used for the third stage of multiplication. Among all the three stages, the second stage of multiplication is considered the most crucial stage in the design of a multiplier because it significantly reduces the operation and efficiency. As a result, to minimize power consumption and to enable faster functioning of the entire multiplier circuit, compressors are been employed in the second stage. Fig 6 depicts the simulated results of multiplier architecture with precise compressors. In this paper, AND gates are used to design an n=8 Dadda multiplier, with approximated 4:2 compressor, as well as an accurate carry, propagate adder in the second and third phases respectively.

Five Dadda multipliers are simulated in this research work.

- All the precise 4:2 compressors of [2] make use of the initial multiplier.

- All the precise 4:2 compressors of [1] make use of the next multiplier.
- During the initial phase of the Dadda multiplier, the precise multipliers of [2] are employed by the 3rd multiplier and in the second phase, the suggested approximate 4:2 compressor is been used. In the initial phase, precise compressors are used to allow only estimates in the second phase.
- During the initial and second phases of the Dadda multiplier, precise multipliers of [1] are employed by the 4th multiplier, and suggested approximate 4:2 compressors are employed in the second phase.
- During the first as well as the second phase of the Dadda multiplier, all approximate 4:2 compressors are employed by the 5th multiplier.

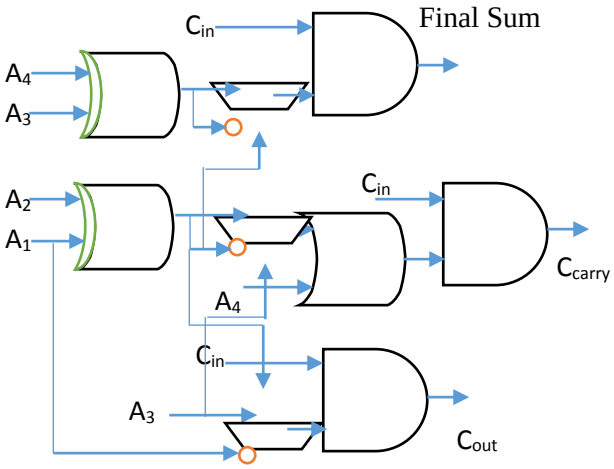


Fig. 3: Proposed 4:2 compressor adder for TOSAM multiplier

The proposed 4:2 compressor is the main block adder for adding the partial products generated by the TODAM multiplier and Fig.3 shown only XOR and AND gates to minimize the number of transistors and power consumption optimization. The MUX blocks are to select the proper selection of carrying bits generated by XOR gates to optimize latency and increase throughput. The main advantages of this circuit are the reduction of the number of transistors from 22 to 34.

The initial two multipliers are precise multipliers with zero error distance, but with greater delay and power consumption. Since the approximate compressor is employed by only one phase of the multiplier, the third multiplier, as well as the fourth multiplier, are partially approximate multipliers. The fifth multiplier increases the error distance and thereby reduces the area as well as power consumed by incorporating all approximate multipliers. Therefore, it is a completely exact multiplier. The designs of the 3rd and 4th multipliers are not depicted because the design is similar to that of Fig. 3, however, the precise compressors in phase 1 have been substituted with an approximate 4:2 compressor and the 5^{th} multiplier consists of all suggested compressors. Since the number of inputs as well as outputs does not change, the multiplier structure also does not change. The significance of employing the suggested compressors for multiplication is been examined in this section. In most cases, a precise (exact) multiplier is made up of three components (or modules) [8].

- Generation of Partial product.
- A Carry Save Adder (CSA) tree is used to minimize the matrix of partial products to two addition operands.
- For the final estimation of the binary output, a Carry Propagation Adder (CPA) is used.

The second unit is very important for designing the multiplier in terms of circuit complexity, power consumption, and delay. To enable faster operation of the CSA tree, compressors have been frequently employed to reduce the power consumed and thereby resulting in quick and low-power operation [9, 10]. The approximate multiplier is obtained with the application of approximate compressors in a multiplier's CSA tree. To evaluate the effect of implementing the suggested compressors in approximate multipliers, an 8x8 unsigned Dadda tree multiplier is being used. In the initial phase, all the partial products using AND gates are generated by the suggested multiplier. The reduction architecture of a precise multiplier for n=8 is shown in Figure 9(a). Half-adders, full-adders, as well as 4-2 compressors, are used in the reduction phase of this design, and every partial product bit is indicated by a dot. Two half-adders, full-adders, as well as eight compressors, are used in the first phase to minimize partial products into four rows. The final two partial product rows are calculated using one half-adder, full-adder, as well as ten compressors in the second or final phase. In the reduction circuit of an 8x8 Dadda multiplier, two phases of reduction as well as three half-adders, full-adders, and 18 compressors are

required. When contrasted to an accurate multiplier, the first two estimated models aim to minimize delay and power usage. A large error distance, on the other hand, is predicted. To reduce the error distance, multiplier 3 and multiplier 4 are presented. The precise compressors in the critical path influence the delay in these systems. As a result, when compared to an accurate multiplier, there is no change in latency for these imprecise systems. The use of imprecise compressors in the least significant columns is intended to reduce power consumption as well as the number of transistors. The initially suggested two multipliers have improved delay as well as power consumption, whereas, the 3rd and 4th designs are predicted to have reduced error distances. By using CMOS logic in NOR and XNOR gates, as well as by using pass transistor logic, the approximate 4:2 compressor design 1 has been employed in distinct ways and the combinations are illustrated in Table I. The area, as well as power consumed by the multipliers, vary depending on the number of transistors of exactly 4:2 compressors and 4:2 approximate compressors. In all of the preceding multipliers, the first and 3rd phases are identical and precise; whereas in the second phase, the 4:2 compressors have been substituted as per the multiplier implementation type. The performance measures of multipliers are shown in Fig 4 (a & b).

Table 1. Comparison between existing and proposed TOSAM multiplier

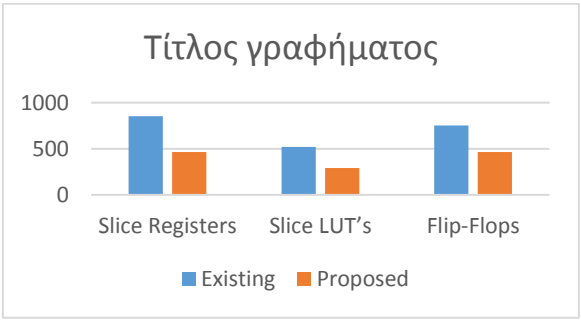
Parameters	Existing	Proposed
Slice Registers	852	466
Slice LUT's	520	290
Flip-Flops	752	465
Delay in ns	5.72	3.875
Power in mW	1.4	0.088

The complete TOSAM multiplier is synthesized in Vivado Design suite software and its design summary is shown in Table.1, the number of slice registers and LUTs is reduced by 45% and power is minimized by 41%.

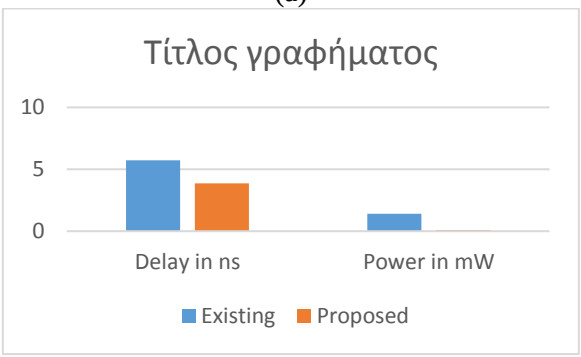
Table 2. Comparison between existing and proposed Dadda multiplier with 4:2 compressor adder

Parameters	Existing	Proposed
Slice Registers	852	48
Slice LUT's	520	103
Flip-Flops	752	103
Delay in ns	5.72	11.705
Power in mW	1.4	0.448

The Dadda multiplier with a 4:2 compressor adder consumed more power and produced more delay in partial products. So it concluded that the TOSAM multiplier with a 4:2 compressor adder has optimized power, area, delay, and hardware utilization. Therefore, the TOSAM with adder is applied for image processing applications as validations of results. To minimize an error between exact and approximate multipliers, the Machine Learning (ML) is applied, based on the number of iterations, ML has greatly reduced the error and it is nearly 0.02%, therefore, proposed TOSAM with ML and 4:2 compressor adder is best suitable for any complex applications like image and video processing applications with minimal hardware utilization and high throughput.



(a)



(b)

Fig. 4: Performance analysis of proposed TOSAM with 4:2 compressor adder.

The proposed design is compared with a recently published article in [24] about Resilient backpropagation Neural networks (RBPNN) as Machine learning and it has been applied for reduction of error produced between exact and approximate multipliers. The proposed

multiplier is validated for real-time data which is published and available at Rajasthan technical university as a standard dataset and discussed in [24], this dataset is about forecast weather. The neural network has compared two data sets errors after RBPNN is trained and accordingly generates an error to minimize this error. The RBPNN is the latest biological network and its operation is based on interconnected neutrals with biases and weights. This network is mainly to solve regression-related issues like forecasts related to weather.

Error Analysis: Measuring accuracy is been conducted for the specified binary inputs A[0-7] as well as B[0-7] along with 100ns stop time. A multiplier's error distance is defined as the absolute difference between the actual product (P) as well as the approximate product (P'), which is determined for every 5ns time ranging between 0 to 100ns. MATLAB is used to carry out all of these tasks. P, P' are depicted in Fig. 9 as an instance of ED, MED, as well as NED computations for the period of 10ns to 15ns.

Error Distance: To evaluate the error distance, four new approximate multipliers are computed. For n=8, the multiplier presented in [7] i.e., multiplier 5 is approximated. For n=8 as well as for k=1, the reduced multiplier with constant correction [5] (Multiplier 6) as well as the reduced multiplier with variable correction [6] (Multiplier 7) is approximated. Furthermore, to evaluate the effect of implementing the suggested approximate compressors with other approximate compressors, an approximate multiplier i.e., Multiplier 8 is simulated. The 4-2 compressors are designed using two precise full adders and are used in an 8x8 Dadda multiplier (Figure 3). The approximate multiplier utilizes the first full-adder design that is presented in [2]. 8 approximate multipliers are evaluated in this work, namely the four proposed architectures as well as the remaining four are the approximate multipliers which is been described in Table IX along with their significant characteristics. To evaluate these approximate multipliers, a normalized error distance (NED) is employed. The average error distance throughout all inputs and is further reduced by the maximum possible error is termed normalized error distance in [1]. As a result, the average NED and the normalized error distance described in [1] are all same. The normalized error distance is specified for every input in this work. The highest absolute value of NED is achieved when the incorrect output is more (less) than the precise outcome and is also known as the maximum high (low) NED. The average normalized error distance, maximum high and low normalized error distances, as well as the accurate number of approximate multiplier outcomes for n=8, is illustrated in Table X. The possibility of accuracy for any design is represented by the number of correct outputs out of the overall outputs. According to Table X, the possibility of accuracy in Multiplier 1 and Multiplier 4 is 0.16 percent (103 out of 65025), as well as 14.3 percent (9320 out of 65025) respectively. The suggested approximate multipliers will yield an incorrect output only if one of the inputs is 0, and the suggested approximate compressors will produce output when all the inputs are zero (row 1 in Tables II and III). Therefore, the multiplier can yield an accurate outcome by introducing a circuit to identify zero-valued inputs for n=8. As a result, the zero-valued input patterns are not taken into account anymore in the simulation for a valid comparison.

When compared to all estimated multipliers, a low average normalized error distance is achieved by multiplier 4. Multiplier 4's average NED is improved 18 times that of Multiplier 5, 2 times that of Multiplier 6, as well as 1.5 times that of Multiplier 7. Maximum correct outcomes are achieved by multiplier five and also possess reduced max high NED. The maximum high NED for this architecture is Zero because the estimated result is always less than the actual output. As a result, when compared to all the architectures, multiplier 5 possesses the worst maximum low NED. To evaluate the results of approximate multipliers, a graph of the NED distribution is made (Figure 10). The product in an 8x8 multiplier ranges between 0 to 65025 and is unsigned values. 127 intervals are used to classify all probable outcomes. The output varies between 0 to 512 during the first interval and ranges between 513 to 1024 in the second interval. The outcome varies from 64513 to 65025 during the last interval. The approximate multiplier is then calculated using the average NED for every interval. Figures 10a and 10b demonstrate that the average NED for Multipliers 1 and 2 increases only at extremely high or very small product values, indicating that these approximate multipliers result in a minimal output error on average when compared to the precise estimation.

Approximate Multipliers: For n=8, the suggested four approximate multipliers are been evaluated and these approximate strategies, as well as precise multipliers, are examined in terms of delay, power consumed as well transistor count. The comparison between other approximate multipliers as well as with the suggested multipliers' error distance is explored in this work.

Approximate Compressors: At a frequency of 1 GHz, the two estimated compressors, as well as the least precise compressor of [8] that are constructed with XOR-XNOR gates, are simulated in this work. In all simulations, a fan out of four is used. Table IV shows the simulation results for power consumption, delay, as well as power-delay product (PDP) employing PTMs at 16 nm, 32 nm as well as 22 nm.

Delay: Based on the reduction phases as well as the delay of every phase, the delay of the reduction circuitry i.e., the second module of a Dadda multiplier is computed. Approximate compressors are employed in all columns in the case of multipliers 1 and multiplier 2. As a result, approximate compressors' delay, as well as the delay of each phase, are the same. Therefore, approximate compressors delay and multiplier three, as well as multiplier 4 delay, are all same. On the other hand, there is no change in delay when using an approximate compressor in n/2 LSBs than using a precise multiplier. Table VI represents the delayed improvement in the reduction circuitry of every multiplier at 32 nm CMOS technology when contrasted with a precise adder.

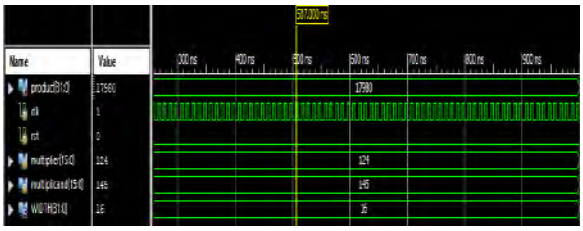


Fig. 5: Simulated results of Dadda multiplier with 4:2 Compressor for 16 bits.

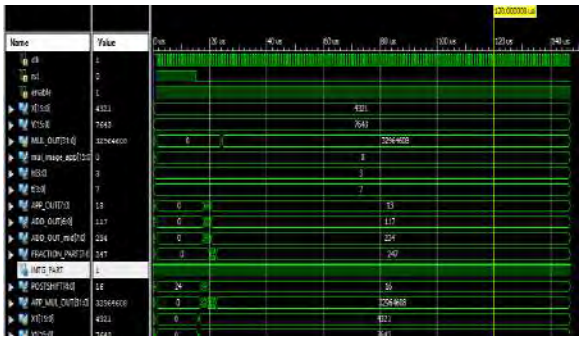


Fig. 6: Simulated results of TOSAM multiplier with 4:2 Compressor for 16 bits of inputs A= 4321 and B= 7643 and product is 32964608.

The simulated results of both Dadda and TOSAM multipliers are validated for different test vectors and based on products results as shown in Fig.5 and Fig.6, the error between exact i.e Dadda multiplier and approximate multiplier i.e TOSAM is 33025403 -32964608 =60795 so error rate is 0.6 and this error is acceptable for image and video processing applications.

3 Application of Proposed Digital Multipliers Using ML

Nowadays, ML has become the most important topic in almost all areas. In general, there are two ways to apply machine learning to VLSI. One way is by utilizing hardcoded coding where the program's exact constraints, as well as necessary outputs, are described and the instructions for the entire process are specified. Therefore, by employing ML algorithms the VLSI processes are made more precise as well as simple. In contrast, an architecture for a neural network is developed to understand how the task can be carried out by itself. Designing effective VLSI hardware designs that are based on neural networks enables the applications of Machine Learning. ML is a process that is carried out continuously by evaluating different alternatives until an appropriate solution is found. Minimizing the number of errors that are present between precise as well as approximate 4:2 compressors is the objective of this research work. By using inaccurate logic minimization, several combinations of imprecise compressors are developed until a lower error rate is achieved. Tensor Flow is used to integrate VLSI approximate circuits with Machine Learning. An algorithm with a reduced error rate is demonstrated by choosing a 4:2 compressor.

4 Manual Calculation of TOSAM Multiplier

Let 16 bit of X operand and 16 bit Y operand
Input Parameters (h, t), where h: height and t: fraction part
For Ex: 1011_1000_001 can be represented as
1.**011_1000_001** x 2^{-10}
Green – fraction part
0000_0010_0001_0 101 x 0001_1010_001 1_1100
2^{15}, 2^{14}, 2^{13},…… 2^{1}, 2^{0}
Check first 1 from MSB and find its binary location
Here for X first '1' comes at 2^{9} KA=9
Here for Y first '1' comes at 2^{12} KB=12
Next to KA fractional parts and t =7
(XA)t = **0000101** (YA)t = **1010001**
APX first 3 bits from (XA)t and pad '1' at LSB side
(XA)APX = **0001** (YA)APX = **1011**
Final computation = ((XA)APX**) x** (YA)APX + **1** = 8 bit output = 0000_1011 = 0000_1011
((XA)APX**) x** (YA)APX + 1 + (XA)t + (YA)t = 0000_1011
0000_1010 (pad 0 at LSB side)
1010_0010 (pad 0 at LSB side)
01 1011_0111
POST SHIFT >> KA+KB-(t+1)= (9+12)=21
1011_0111)<< 21 (already 8 bit fractional part)
(21-8)= 13 times <<
FINAL OUTPUT= 01 1011_0111
0000_0000_0000_0 = 35,96,288
EXACT OUTPUT= 11011010011111011101100 = 35,79,628

5 Conclusion

This research work provides the presentation of approximate 4:2 compressors using approximate logic minimization with error analysis. The 8x8 Dadda multiplier employs a modern 4:2 compressor. The analysis of error and simulations of these designs were carried out using MATLAB and Vivado Design Suite 2018.1 simulator. The circuit developed in this work achieves a lower error rate and average power by considering the transistor count. At last, this research work described the concept of integrating VLSI approximate circuits as well as machine learning. One of the unique approaches for computing nanoscale is approximate computing. For approximate computing, computer arithmetic provides huge operational benefits, and there exist many works on inexact adders. Furthermore, this research work primarily concentrates on compression in the context of a multiplier till now no work on this area has been presented as per the author's knowledge. The unique architectures of two approximate 4-2 compressors were proposed in this research. The compression unit comprising four approximate multipliers uses these approximate compressors. When compared to a precise design, proposed approximate compressors exhibit a reduced number of transistors, delay as well as power consumption.

- The number of transistors used in design 1 and design 2 is reduced by 46% and 50% respectively.
- For CMOS implementation at technology scaling of 32, 22 as well as 16 nm, the first design and second design offer a power reduction of 57% as well as 60 % respectively.
- Overall, for varied CMOS feature sizes of 32, 22 as well as 16 nm, the second design as well as the first design show 44% and 35% improvement in delay respectively.

This research work proposes four distinct approximate approaches to examine the performance of inexact compressors for the specified imprecise multiplication measures. The reduction unit of a Dadda multiplier employs inexact compressors. The simulation outcomes reported in this work are as follows:

- When contrasted to a precise multiplier, the first multiplier and second multiplier suggested in this work exhibit better results in terms of the number of transistors as well as power usage.
- The second multiplier and first multiplier exhibit significantly large normalized error distances as well as large PSNRs. The second multiplier also exhibits maximum delay by using a second inexact compressor for all bits.
- The 3rd multiplier and 4th multiplier exhibit reduced normalized error distance values because of the reduced number of transistors and power consumed and thereby indicating optimum trade-off for energy as well as precision

Furthermore, the presented imprecise multipliers are employed in many of the outstanding image processing applications that achieve a peak signal-to-noise ratio that is approximately 50dB. The comparison of the suggested four approximate

models against four other approximate models is been evaluated in Table XIII. In terms of all performance criteria for imprecise multiplication as well as the for two PSNR instances, the multiplier 4 is considered one of the best solutions. Concerning Max High NED as well as the number of accurate outputs, the multiplier 5 exhibits better results. Whenever the PSNR instances are taken into account, the other performance criteria that are low enable the PSNR to rank in the middle. Among all the strategies examined in this research work, multiplier 3 is the second most efficient design. Existing, as well as upcoming work, mainly concentrates on the trade-offs between various performance criteria. Physical architectures of approximate multipliers are explored to facilitate the analysis described in this work. At last, this report illustrates the multipliers that are used for approximate computing and can be designed using an appropriate approximation compressor architecture. These multipliers provide huge benefits in terms of error measurements as well as circuit-level. Analysis of error indicators is the present concept that is examined in this work.

References:

[1]. Veeramachaneni *et.al*., Novel architectures for high-speed and low power 3-2, 4-2 and 5-2 compressors, Proc. Int. Conf. on VLSI Design (VLSID), 2007, pp. 324-329.

[2]. Chang *et.al*, Ultra- voltage, low power CMOS 4-2 and 5-2 compressors for fast arithmetic circuits, IEEE Trans. Circuits Syst. I, Fundam. Theory Appl., 2004, 51, (10), pp. 19851997.

[3]. Raphael, D *et.al*, A Power-Efficient 4-2 Adder Compressor Topology, 15th IEEE (NEWCAS), Strasbourg, France, 2017, pp. 281-284.

[4]. Swagath, V *et.al*, Approximate Computing and the Quest for Computing Efficiency, 52nd (DAC), 2015, San Francisco, CA, USA.

[5]. Shaahin, A *et.al*, Majority-Based Spin-CMOS Primitives for Approximate Computing, IEEE Trans. on Nanotech., 17(4), 2018, pp. 795-806.

[6]. Avinash, L *et.al*, Parsimonious Circuits for Error-Tolerant Applications through Probabilistic Logic Minimization, Int. Workshop on PATMOS 2011, pp.204-213.

[7]. Darjnlow, E *et.al*, Approximate Multipliers Based on New Approximate Compressors, IEEE Trans. on CAS-I: Reg. Pap, PP(99), 2018, pp. 1-14.

[8]. Momeni, A *et.al*, Design, and analysis of approximate compressors for multiplication, IEEE Trans. on Comp., 64 (4), 2015, pp. 984994.

[9]. Liang, J *et.al*, New Metrics for the Reliability of Approximate and Probabilistic Adders, IEEE Trans. on Comp., 63(9), 2013, p. 1760-1771.

[10]. Zervakis, G *et.al*, Design-Efficient Approximate Multiplication Circuits Through Partial Product Perforation, IEEE Trans. on VLSI Systems, 24(10), 2016, pp. 3105-3117.

[11]. A. Betti, M. Gori, and G. Marra, "A Constrained-Based Approach to Machine Learning," 2018 14th International Conference on Signal-Image Technology & Internet-Based Systems (SITIS), 2018, pp. 737-746, DOI: 10.1109/SITIS.2018.00118.

[12]. Y. Cho and M. Lu, "A Reconfigurable Approximate Floating-Point Multiplier with CNN," 2020 International SoC Design Conference (ISOCC), 2020, pp. 117-118, DOI: 10.1109/ISOCC50952.2020.9332978.

[13]. M. Hajizadegan and P. Chen, "Harmonics-Based RFID Sensor Based on Graphene Frequency Multiplier and Machine Learning," 2018 IEEE International Symposium on Antennas and Propagation & USNC/URSI National Radio Science Meeting, 2018, pp. 1621-1622, DOI: 10.1109/APUSNCURSINRSM.2018.8608604.

[14]. D. G. Mahmoud, B. Shokry, A. ElRefaey, H. H. Amer and I. Adly, "Runtime Replacement of Machine Learning Modules in FPGA-Based Systems," 2021 10th Mediterranean Conference on Embedded Computing (MECO), 2021, pp. 1-4, DOI: 10.1109/MECO52532.2021.9460192.

[15]. Y. Ishiguchi, D. Isogai, T. Osawa and S. Nakatake, "A Perceptron Circuit with DAC-Based Multiplier for Sensor Analog Front-Ends," 2017 New Generation of CAS (NGCAS), 2017, pp. 93-96, DOI: 10.1109/NGCAS.2017.23.

[16]. Xiang-Jun Ji and Na Han, "Evaluation model on the building of real estate brand based on income multiplier," 2009 International Conference on Machine Learning and Cybernetics, 2009, pp. 2549-2554, DOI: 10.1109/ICMLC.2009.5212098.

[17]. R. Zhang and Q. Zhu, "Consensus-based transfer linear support vector machines for decentralized multi-task multi-agent learning," 2018 52nd Annual Conference on Information Sciences and Systems (CISS), 2018, pp. 1-6, DOI: 10.1109/CISS.2018.8362195.

[18]. H. Wang, Y. Gao, Y. Shi, and R. Wang, "Group-Based Alternating Direction Method of Multipliers for Distributed Linear Classification," in IEEE Transactions on Cybernetics, vol. 47, no. 11, pp. 3568-3582, Nov. 2017, DOI: 10.1109/TCYB.2016.2570808.

[19]. R. Dornelles, G. Paim, B. Silveira, M. Fonseca, E. Costa, and S. Bampi, "A power-efficient 4-2 Adder Compressor topology," 2017 15th IEEE International New Circuits and Systems Conference (NEWCAS), 2017, pp. 281-284, DOI: 10.1109/NEWCAS.2017.8010160

[20]. S. Venkataramani, S. T. Chakradhar, K. Roy and A. Raghunathan, "Approximate computing and the quest for computing efficiency," 2015 52nd ACM/EDAC/IEEE Design Automation Conference (DAC), 2015, pp. 1-6, DOI: 10.1145/2744769.2744904.

[21]. S. Angizi, H. Jiang, R. F. DeMara, J. Han and D. Fan, "Majority-Based Spin-CMOS Primitives for Approximate Computing," in IEEE Transactions on Nanotechnology, vol. 17, no. 4, pp. 795-806, July 2018, DOI: 10.1109/TNANO.2018.2836918

[22]. D. Esposito, A. G. M. Strollo, E. Napoli, D. De Caro, and N. Petra, "Approximate Multipliers Based on New Approximate Compressors," in IEEE Transactions on Circuits and Systems I: Regular Papers, vol. 65, no. 12, pp. 4169-4182, Dec. 2018, DOI: 10.1109/TCSI.2018.2839266.

[23]. A. Momeni, J. Han, P. Montuschi and F. Lombardi, "Design and Analysis of Approximate Compressors for Multiplication," in IEEE Transactions on Computers, vol. 64, no. 4, pp. 984-994, April 2015, DOI: 10.1109/TC.2014.2308214.

[24]. G. Zervakis, K. Tsoumanis, S. Xydis, D. Souris and K. Pekmestzi, "Design-Efficient Approximate Multiplication Circuits Through Partial Product Perforation," in IEEE Transactions on Very Large-Scale Integration (VLSI) Systems, vol. 24, no. 10, pp. 3105-3117, Oct. 2016, DOI: 10.1109/TVLSI.2016.2535398.

Received: 6 November 2021 | Revised: 30 May 2022 | Accepted: 13 August 2022

DOI: 10.1002/ett.4640

RESEARCH ARTICLE

WILEY

Hybrid optimization model for design and optimization of microstrip patch antenna

Ajitha Subhadra Sarngadharan[1] | **Rashmi Narasimhamurthy**[2] | **Balasubramaniam Sankaramoorthy**[3] | **Suryabhan Pratap Singh**[4] | **Charanjeet Singh**[5]

[1]Department of Electronics and Communication Engineering, TKM College of Engineering, Kollam, Kerala, India

[2]BMS Institute of Technology and Management, Bengaluru, Karanataka, India

[3]Department of Futures Studies, University of Kerala, Thiruvananthapuram, Kerala, India

[4]Department of Information Technology, Institute of Engineering and Technology, Deen Dayal Upadhyaya Gorakhpur University, Gorakhpur, Uttar Pradesh, India

[5]Department of Electronics and Communication, Deenbandhu Chhotu Ram University of Science and Technology, Murthal, Sonipat, Haryana, India

Correspondence
Ajitha Subhadra Sarngadharan, Department of Electronics and Communication Engineering, TKM College of Engineering, Kollam, 691005, Kerala, India.
Email: ajithasubhadra@tkmce.ac.in

Abstract

The antenna design for a specified resonant frequency necessitates the computation of optimal values of different sizes. This is a harder task for microstrip patch antenna (MPA) since there is no precise numerical formula that leads to accurate solutions for designing these antennas. Presently, bio-inspired approaches are widely deployed in numerous antenna designs and it has revealed an immense assurance in handling the rising necessities of antenna engineering for overall cost, reduced size, and enhanced performances. This work aims to introduce an optimal MPA design, where the antenna elements like patch length, patch height, substrate width, and substrate length are optimally tuned by a new hybrid optimization model. For this, a new hybridized model known as shark smell integrated EHO is proposed. Eventually, the primacy of the suggested model is scrutinized via varied assessments.

1 | INTRODUCTION

The ability to integrate an microstrip patch antenna (MPA) makes it better renowned & productive for appliances like-cellular, satellite or aerospace systems.[12] Together with several outstanding features such as light-weighted, lower-power, lower-cost, and lower profiles, the realistic deployment of MPA is hindered owing to narrowband features, predominantly in broadband appliances.[3,4] Enhanced bandwidths in the antenna is an excellent region of analysis in previous decades. Various such studies were done with different antenna geometries. Improved bandwidth of MPA is obtained by low permittivity with a thick substrate; nevertheless, the effectiveness of radiation gets corrupted owing to surface wave.[5-7] Machine learning techniques are currently a big part of ongoing research, and they are projected to be a big part of today's innovation.[8-11]

Accordingly, MPA is fed using several schemes, which were grouped under 2 types; (a) non-contacting and (b) contacting.[12] In the latter scheme, the RF power is directly offered to radiating patches via a microstrip line.[13,14] In the former method, EMF coupling is made for transmitting power among the radiating patches and microstrip lines.[15] The 4 renowned feeding approaches were the "microstrip line, coaxial probe (both contacting schemes), aperture coupling and proximity coupling (both non-contacting schemes)"; nevertheless, their patterns differ based on frequency.[16-18] Moreover, they reveal poorer back-radiation and XP features.[19-21]

Commonly, the most well-known approaches for examining MPA were the cavity model, transmitting line model, and full-wave schemes.[22-24] The full-wave system is very flexible, precise and it takes care of the infinite and finite arrays, distinct components, random components, stacked, and coupling components.[25,26] For the manipulation process computer vision is required.[9] The transmission-based system is uncomplicated and it offers enhanced physical insights; nevertheless, it incurs minimal accuracy. The cavity system offers superior physical insights and is much precise; however, it is intricate.[27]

The contributions are given below.

- Develops an optimal MPA design, in which the antenna constraints are tuned optimally via a novel hybrid optimization model.
- By optimally tuning the constraints of MPA, higher gain and minimal return loss are attained.
- For tuning purposes, a new hybridized model named shark smell integrated EHO (SSI-EHO) is introduced.

This article is sorted as: Section 2 reviews the concerned topic. Section 3 offers a brief explanation on the microstrip patch antenna and Section 4 defines the objectives of developed work. Proposed SSI-EHO based optimization for optimal MPA design is briefed in Section 5. Sections 6 and 7 illustrate the outcomes and conclusion.

2 | LITERATURE REVIEW

2.1 | Related works

In 2020, Dutta et al[28] proposed RMPA with a simpler design and superior XP transparency. The formulated method was extremely flexible for adjusting the frequency of the antenna to obtain optimal performances. Further, a wide-ranging design was presented for computing the frequency and for computing the relevant, resonant modes. In the end, the attained simulated resultants demonstrated the efficacy of the proposed theory.

In 2018, Qian et al[29] offered an innovative MPA design with high gain and fine beam-width with MTM. Consequently, an arithmetical design was modeled for evaluating MPA, which assisted in acquiring optimal performance by optimizing the antenna constraints. This method also presented noteworthy control for attaining a transmitting antenna system with high efficiency.

In 2020, Chandrashekar et al[30] developed a compacted MPA approach to realize wideband features. In addition, an "annular metal ring" was employed and it acted as the inductive load for rectangle patches. Also, the robust systematic technique was deployed for electing the size of rings and patches, wherein they vibrate nearer to one another. This strategic model guaranteed broad radiation. Eventually, the analyzed outcomes have exposed the development of an innovative scheme regarding better bandwidth.

In 2020, Rashmitha et al[31] modeled the MPA system, which was compatible with 5G communication. The antenna was operated at 43.7 GHz. Consequently, a substrate of FR4 epoxy was employed here and the antenna was scrutinized based on varied parameters such as gain and so on. In the end, the obtained outcomes were established to be deployed for diverse appliances.

In 2019, Hocini et al[32] modeled 5 THz MPA based upon "modified photonic bandgap substrate in the frequency range from 0.5 to 0.8 THz." Accordingly, the purpose here was to bring the antenna features closer to 0.65 THz with communiqué and sensing appliances. In addition, experiments were done depending upon different substrates.

In 2021, Susamay et al[33] introduced a new quantitative parameter, which recognized the vital feature for appropriate geometry selection in low XP machinery for MPA. The significances of PS and NPS geometries were determined depending upon FAR. On carrying out precise experimentation, the proposed model ensured lower XP radiations.

TABLE 1 Reviews on existing MPA models

Author	Techniques	Advantages	Drawbacks
Dutta et al[28]	XP based model	• Negligible polarization levels • Superior peak gain	• Impedance matching is corrupted
Tian et al[29]	FEM method	• High gain • Efficiency is large	• Radiation gain need to be considered more
Chandrashekar et al[30]	Microstrip feeding method	• Least XP • Obtains higher gain	• Need contemplation on bandwidth and radiation properties
Rashmitha et al[31]	Trial and error technique	• Condensed impedance • Higher gain.	• No execution on a realistic platform
Hocini et al[32]	FEM	• Reduced loss • High efficiency	• Shock waves might corrupt the antenna performance
Susamay et al[33]	FAR	• Minimal XP • High efficiency.	• Requires deliberation on diverse kinds of the feeding mechanism.
Shalini et al[34]	Dual-polarization scheme	• Negligible return losses • High peak gain.	• Polarization issues need much concern
Hossein al[35]	FEM	• Minimal cost • Higher gain	• Return loss is not concerned.

In 2019, Shalini et al[34] modeled a "microstrip fed graphene-based dual-polarized patch antenna that included single graphene patch fed by dual ports oriented orthogonally." Here, the patches have 3.98 THz frequency and were positioned upon silicon substrates. The adopted scheme provided a minimal loss, high gain, and high efficiency. Eventually, simulated outcomes have confirmed the supremacy of the developed model regarding gain, efficiency, and loss.

In 2019, Hossein et al[35] suggested a new MPA model with high efficiency for THz appliances. The presented design exhibited negligible price as the substrates were FR-4 and it has high effectiveness. Furthermore, on employing fractal structures, the radiation lengths of presented paths were increased and the antenna dimension was diminished. Finally, the resultants attained from simulations exposed the effectiveness of the developed scheme regarding cost, efficiency, and gain.

2.2 | Review

Table 1 reviews the extant MPA approaches. XP-oriented approach[28] has offered better gain and condensed polarization levels. On the other hand, impedance matching was degraded. Moreover, FEM used in Reference 29 has presented higher efficiency and augmented gain; however, it has to be a concern much on gain and radiation. Microstrip feeding method[30] poses effectual control of XP and accomplishes proficient gain. Nevertheless, it wants deliberation on radiation and bandwidth property. The trial and error technique was employed in Reference 31 presented condensed impedance and higher gain; however, there was no execution on the real platform. FEM[32] presents condensed return losses and high radiation efficiency. However, shock waves might corrupt the antenna performances. FAR[33] diminishes XP and offers superior radiation effectiveness. On the other hand, it necessitates contemplation on diverse types of feeding mechanisms. Dual-polarization used in Reference 34 assured high peak gain with negligible losses. Nonetheless, it has to be a concern much on XP issues. FEM used in Reference 35 offered reduced cost and superior gain, but return losses have to be concerned more.

3 | A BRIEF EXPLANATION OF MICROSTRIP PATCH ANTENNA

Generally, MPA encompasses a ground plane as well as a control patch. A substrate is available at the middle that reveals the accurate value of the dielectric constant. Typically, patch sizes are small when compared to the substrate and ground. On designing an MPA, a dielectric medium and resonance frequency are chosen to depend upon the dimension of MPA.

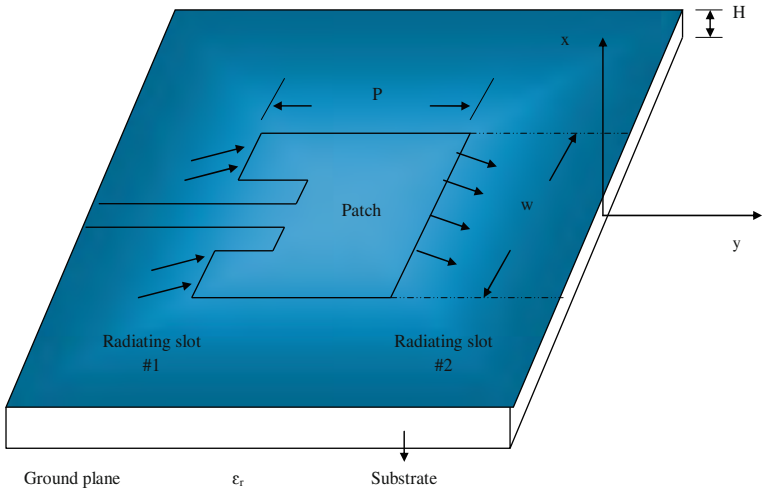


FIGURE 1 The radiation pattern for MPA design

Consequently, to develop the antenna performances, it is required to have a dielectric substrate, in which the effectiveness is made superior with higher radiation and bandwidth. Figure 1 discloses the method of MPA radiation.

4 | OBJECTIVE OF DEVELOPED WORK

The objective of the adopted scheme is revealed in Equation (1), wherein, W_1, W_2, and W_3 refers to variables lying between 0 and 1, Ga points out gain, Eff points out efficiency, BW points out bandwidth, and RL points out return loss.

$$\text{Obj} = Min\left[\frac{1}{|W_1 * \text{Ga} + \text{Eff} + W_2 * \text{BW} + W_3 * \text{RL}|}\right] \tag{1}$$

The variables in Equation (1) are depicted below.

Computation on bandwidth: Equation (2) shows the calculation of BW, in which, $\alpha = 3.77$, $\lambda = L/r^f\varepsilon_r$ points out dielectric value, w points out width, P points out patch length, and H points out height.

$$\text{BW} = (\alpha * ((\varepsilon_r - 1)/(\varepsilon_r * \varepsilon_r)) * (w/P) * (H/\lambda)) * 100. \tag{2}$$

Computation on gain: The part of antenna intensity to the input power is said to be MPA gain. Equation (3) demonstrates the MPA gain, in which Leff is designed as revealed in Equation (4).[36]

$$\text{Ga} = 10 * \log_{10}\left(\text{Eff} * 2 * \pi * \text{Leff} * \frac{w}{\lambda}\right), \tag{3}$$

$$\text{Leff} = \frac{\lambda}{2}. \tag{4}$$

Efficiency: It is calculated as revealed in Equation (5), in which, PSP points out space wave power and PSUR points out surface wave power. Here, PSP is modeled as exposed in Equation (6), in which, $\theta = 377$, ε^{re} implies an effectual refractive index. The K_0 and $K_0 d$ are modeled as exposed in Equations (7) and (8), where H implies the substrate's height. Moreover, PSUR is calculated as shown in Equation (9), wherein, x_0 is modeled as exposed in Equation (10), wherein, x_0^d and x_0^n are computed as exposed in Equations (11) and (12). In Equation (12), α_0 is evaluated as in Equation (13) and α_1 is formulated as in Equation (14).

$$\text{Eff} = \text{PSP}/(\text{PSP} + \text{PSUR}), \tag{5}$$

$$\mathrm{PSP} = \theta * K_0^2 * K_0 d^2/3/Pi * \left(1 - 1/\varepsilon^{re} + 2/5/\left(\varepsilon^{re}\right)^2\right), \tag{6}$$

$$K_0 = 2 * Pi * \mathrm{freq}/30, \tag{7}$$

$$K_0 d = K_0.H. \tag{8}$$

$$\mathrm{PSUR} = \theta * K_0^2/4 * \varepsilon^{re} * \left(x_0^2 - 1/\left[\begin{array}{l}\varepsilon^{re} * \left(1/\sqrt{x_0^2 - 1}\right)\\ + \left(x_0^2 - 1\right)/\varepsilon^{re} - x_0^2\end{array}\right]\right)$$
$$+K_0 d * \left[1 + \left(\varepsilon^{re}\right)^2 * \left(x_0^2 - 1\right)/\left(\varepsilon^{re} - x_0^2\right)\right], \tag{9}$$

$$x_0 = 1 + x_0^n/x_0^d, \tag{10}$$

$$x_0^d = \left(\varepsilon^{re}\right)^2 - \alpha_1^2, \tag{11}$$

$$x_0^n = \left(\varepsilon^{re}\right)^2 - \alpha_0 * \alpha_1 + \varepsilon^{re} * \sqrt{(\varepsilon^{re})^2 - 2\alpha_0 * \alpha_1 + * \alpha_0^2}, \tag{12}$$

$$\alpha_0 = S * \tan K_0 dS, \tag{13}$$

$$\alpha_1 = \left[\tan\left(K_0 dS\right) + K_0 dS/\cos\left(K_0 dS\right)^2\right], \tag{14}$$

$$K_0 dS = K_0 d * S, \tag{15}$$

$$S = \sqrt{\varepsilon^{re} - 1}. \tag{16}$$

Width calculation: Equation (17) illustrates the assessment of MPA width, in which, r^f implies resonant frequency, and L refers to speed of light. Equation (18) symbolizes the formula for r^f.[36]

$$w = \frac{L}{2r^f}\sqrt{\frac{2}{\varepsilon_r + 1}}, \tag{17}$$

$$r^f = \frac{L}{2P\sqrt{\varepsilon^{re}}}. \tag{18}$$

Height estimation: The MPA height (H) is calculated as formulated in Equation (19), wherein T implies substrate thickness.[36]

$$H = \frac{T}{1000 \times 2.54}. \tag{19}$$

ERI assessment: The constraint ε^{re} is assessed as in Equation (20).

$$\varepsilon^{re} = \frac{\varepsilon_r + 1}{2} + \frac{\varepsilon_r + 1}{2}\left[1 + 12\frac{H}{w}\right]^{-1/2}, \frac{w}{H} > 1. \tag{20}$$

Standardized length estimation: The standardized length (ΔP) is assessed as in Equation (21).

$$\frac{\Delta P}{H} = 0.412\frac{\left(\varepsilon^{re} + 0.3\right)\left(\frac{W}{H} + 0.264\right)}{\left(\varepsilon^{re} - 0.258\right)\left(\frac{W}{H} + 0.8\right)}. \tag{21}$$

Length evaluation: The MPA length signified by (P) is measured as in Equation (22).

$$P = \frac{L}{2r^f\sqrt{\varepsilon^{re}}} - 2\Delta P. \tag{22}$$

CI estimate: Equation (23) shows the CI of conversion factor, wherein, PI_i implies patch impedances and is formulated as in Equation (24).[36]

$$\mathrm{PI}_I = \sqrt{50 + Z_i}, \tag{23}$$

$$\mathrm{PI}_i = 90\frac{\varepsilon_r^2}{\varepsilon_r - 1}\left(\frac{P}{W}\right)^2. \tag{24}$$

Directivity: "Directivity of an antenna is defined as the ratio of the radiation intensity in a given direction from the antenna to the radiation intensity averaged over all directions. The average radiation intensity is equal to the total power radiated by the antenna divided by 4π."

Return loss: It is also known as power loss and is signified in dB. The return loss is evaluated as revealed in Equation (25), where P_{in} and P_{ref} implies incident and reflected power. The developed model is concerned on minimizing the return losses for better antenna performance.[36]

$$\mathrm{RL}(dB) = 10\log_{10}\frac{P_{\mathrm{in}}}{P_{\mathrm{ref}}}. \tag{25}$$

5 | PROPOSED SSI-EHO BASED OPTIMIZATION FOR OPTIMAL MPA DESIGN

5.1 | Solution encoding

The introduced scheme improves the parameters namely, gain, bandwidth, efficiency, and directivity as exposed in Equation (1). Consequently, the modeling constraints such as patch height (*PH*), patch length (*P*), substrate width (*w*), and substrate length (*SL*) of MPA are tuned optimally via the SSI-EHO model, thereby attaining the optimized microstrip patch antenna design. Figure 2 portrays the solutions given as inputs to SSI-EHO.

5.2 | Proposed M-EFO algorithm

Although extant EFO[37] includes varied improvements, it endures specified limits regarding convergence rate and speed. To prevail over the disadvantages of traditional EFO, the theory of SSO[38] is combined with it to set up a new algorithm termed as SSI-EHO. Hybrid optimization approaches are much more proficient for specified searching issues.[27,39-41]

At first, the electric fish population gets spread randomly using searching space by concerning on space bounds, in which X_{uv} implies uth individual position in N at D-dimension space, $X_{\min v}$ and $X_{\max v}$ implies lower and upper bounds for v dimension, $\phi \in (0, 1)$ implies random values.[37]

$$X_{uv} = X_{\min v} + \phi\left(X_{\max v} - X_{\min v}\right). \tag{26}$$

The frequency acts as the most significant part in SSI-EHO to balance exploration and exploitation and is employed to determine if an individual does active or passive electrolocation.

Active electrolocation: The active range of uth individual (R_i) is shown by their own amplitude value (B_i). The active range calculation in EFO is in Equation (27). Traditionally, Cartesian distance is evaluated depending upon X_{kv} and X_{uv},

FIGURE 2 Solution encoding

however, in the suggested SSI-EHO scheme, the Manhattan oriented distance assessment between two points (X, J) is done as shown in Equation (28). In addition, Cauchy's mutation takes place to identify the best individual.

$$R_i = (X_{\max v} - X_{\min v})\, B_i, \tag{27}$$

$$\text{Dis} = |X_1 - X_2| + |J_1 - J_2|\,. \tag{28}$$

If a single neighbor is available at active sensing area, EFO employs Equation (29); else, Equation (30) is applied, in which, k implies arbitrary elected individual from a neighbor set of uth individual.

$$X_{uv}^{can} = X_{uv} + \phi\,(X_{kv} - X_{uv})\,, \tag{29}$$

$$X_{uv}^{can} = X_{uv} + \phi R_i. \tag{30}$$

Here, $\phi \in (-1, 1)$ in Equation (29) refers to an arbitrary count produced from an even distribution and X_{uv}^{can} refers to the candidate location of uth individual.

Passive electrolocation: Traditionally, the novel position is updated as per X_{rv} and X_{uv}. However, in SSI-EHO based model, the position is updated depending upon SSO[38] update as exposed in Equation (31), where; Y_i^{k+1} symbolize the subsequent position of the shark with higher *Obj* value and G_i^{k+1} refers to a new position due to the shark's forward movement.

$$X_i^{k+1} = \arg\max \left\{ obj\left(G_i^{k+1}\right), Obj\left(Y_i^{k+1,i}\right), \ldots, Obj\left(Y_i^{k+1,M}\right) \right\}. \tag{31}$$

The final phase concerns on adjusting a parameter of uth individual for increasing the probability of a trait to be varied as in Equation (32), in which, $ran(0, 1)$ implies arbitrary values.

$$\begin{aligned} X_{uv}^{can} &= X_{\min v} + \phi\,(-X_{\min v} + X_{\max v}) \\ ran(0,1) &\le ran(0,1). \end{aligned} \tag{32}$$

If vth parameter value of uth individual exceeds search space bounds, it is reassigned to space limits that it exceeds.

$$X_{uv}^{can} = \begin{cases} X_{\min v} & X_{uv}^{can} < X_{\min v}, \\ X_{uv}^{can} & X_{\max v} > X_{uv}^{can} > X_{\min v}, \\ X_{\max v} & X_{uv}^{can} > X_{\max v}. \end{cases} \tag{33}$$

The pseudocode for SSI-EHO is specified by Algorithm 1.

Algorithm 1. SSI-EHO

Start.
Initializing the population.
Compute the quality of individuals.
Repeat.
  Divide the population into passive and active electrolocations based upon the frequency values of every individual
    Compute Manhattan based distance evaluation as shown in Equation (28)
    Apply Cauchy's mutation for identifying the best individual
    Perform active electrolocation as shown in Equations (29) and (30)
    Perform passive electrolocation based on the proposed SSO model as shown in Equation (31)
    Carry out Cauchy mutation for obtaining the best individual
    Update the amplitude and frequency values for every individual
Till termination condition is met.

6 | RESULTS

6.1 | Simulation set-up

The developed MPA design with SSI-EHO was executed in Matlab and outcomes were scrutinized in brief. Here, the analysis was performed based on antenna gain and efficiency. In addition, the advantage of developed antenna design was proven by comparing it with conventional models namely, DA,[42] CSO,[43] PSO,[44] SSO,[38] and EFO.[37] Furthermore, the examination was held regarding cost function for varied iterations from 0, 5, 10, 15, and 20. Moreover, the examination was done regarding gain, reflection coefficient, return loss, and VSWR. Accordingly, the analysis was held by varying the frequency (GHz) that ranges from 1×10^9, 1.25×10^9, 1.5×10^9 and 1.75×10^9.

The diagrammatic illustration of the 3D pattern of fractal carpet antenna obtained using the developed model and extant schemes like DA, CSO, PSO, SSO, and EFO is shown in Figure 3. At every pattern, the frequency, elevation, maximum value, azimuth, minimum value, and output are described.

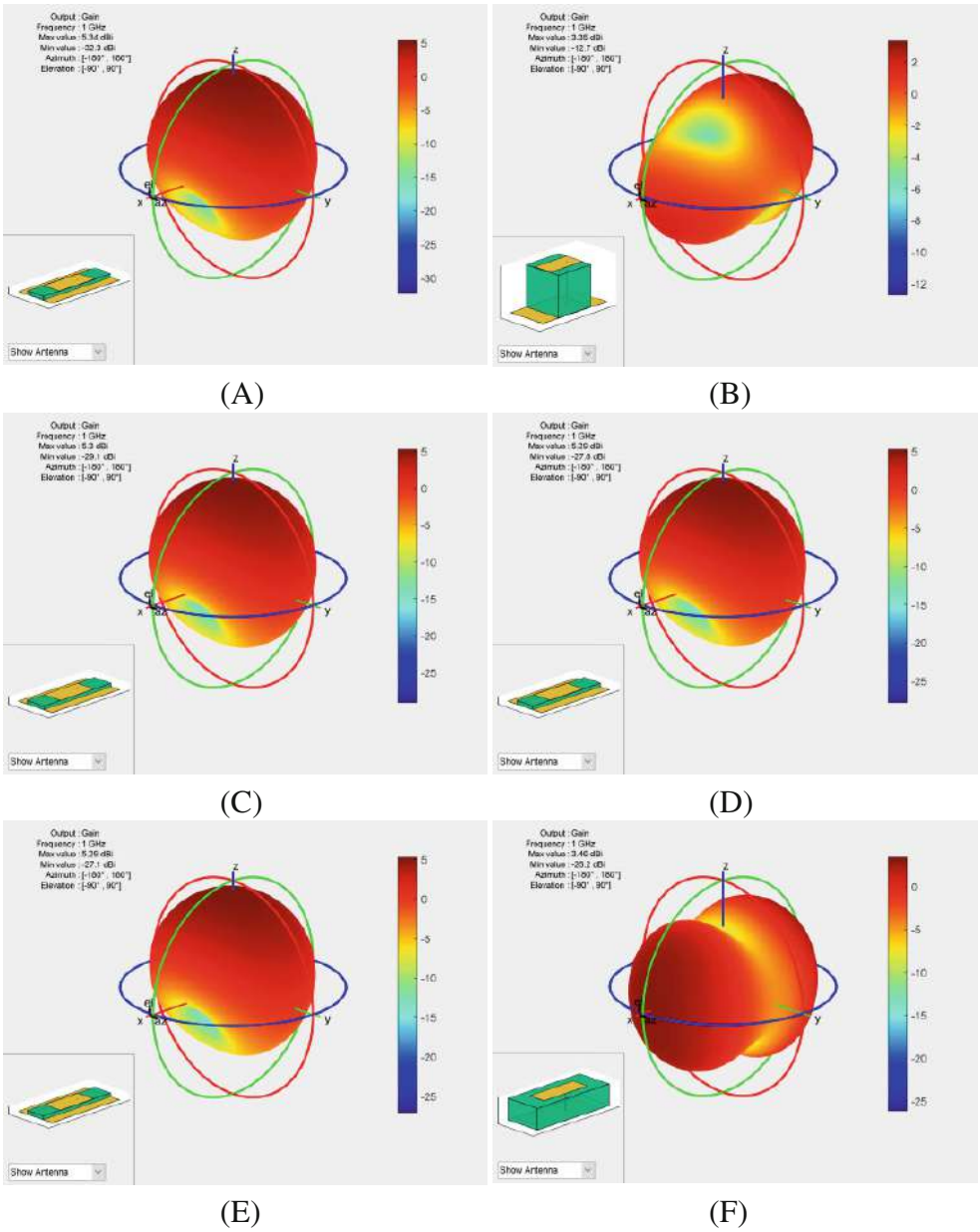


FIGURE 3 3D-Radiation pattern of MPA for (A) CSO, (B) DA, (C) EFO, (D) PSO, (E) SSO, and (F) SSI-EHO

6.2 | Directivity analysis

"Directivity (D) is the quantitative measure of an antenna's ability to concentrate radiated power per unit solid angle in a certain direction and it is highly dependent on the three-dimensional pattern of an antenna.[45]" It is arithmetically shown in Equation (34), wherein, $Ra_{sspecific}$ and $Ra_{reference}$ points out the radiation intensity of the specified antenna and reference antenna, respectively.

$$D = \frac{Ra_{specific}}{Ra_{reference}}. \tag{34}$$

The polar plot exhibiting the directivity of SSI-EHO over the conventional models such as DA, CSO, PSO, SSO, and EFO at the specific frequency of 1×10^9 is exposed in Figure 4.

6.3 | Charge distribution

The supreme value of charge on the antenna surface is evaluated and the resultants are graphically symbolized at a specified frequency of 1×10^9. The distribution of charge on the antenna surface using the SSI-EHO model over the conventional models such as DA, CSO, PSO, SSO, and EFO is revealed in Figure 5.

6.4 | Current distribution

The distribution of surface current in MPA holds a crucial role in obtaining the objective and as a result, it is necessary to scrutinize the distribution of surface current. Here, the distribution of current is evaluated at a frequency of 1×10^9.

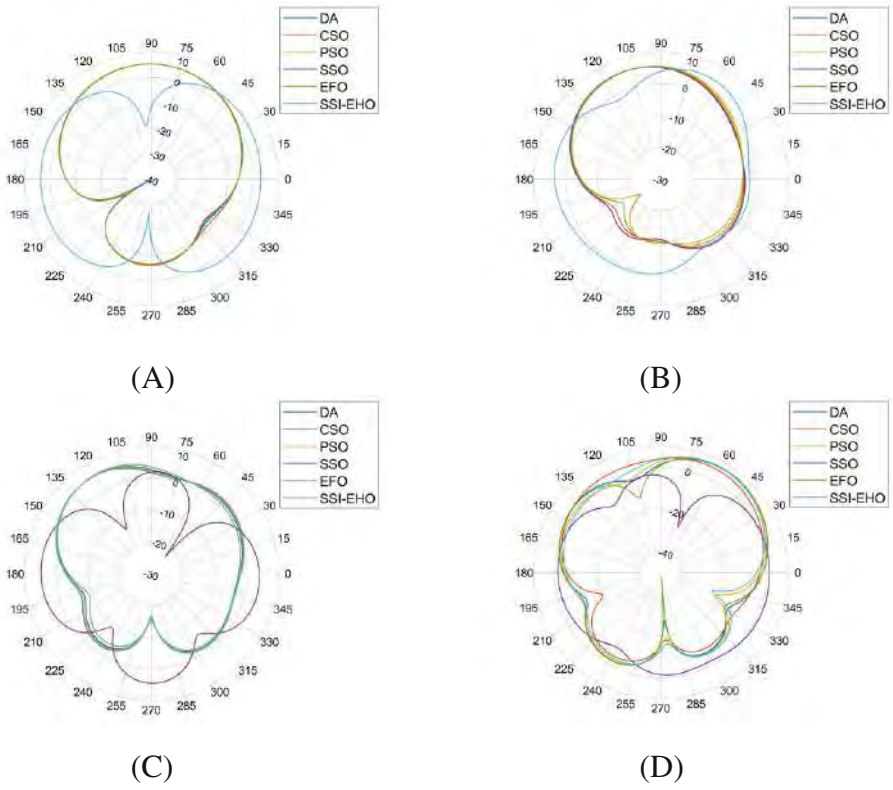


FIGURE 4 Polar plot revealing the directivity of optimized MPA for varied frequencies like (A) 1×10^9, (B) 1.25×10^9, (C) 1.5×10^9, and (D) 1.75×10^9

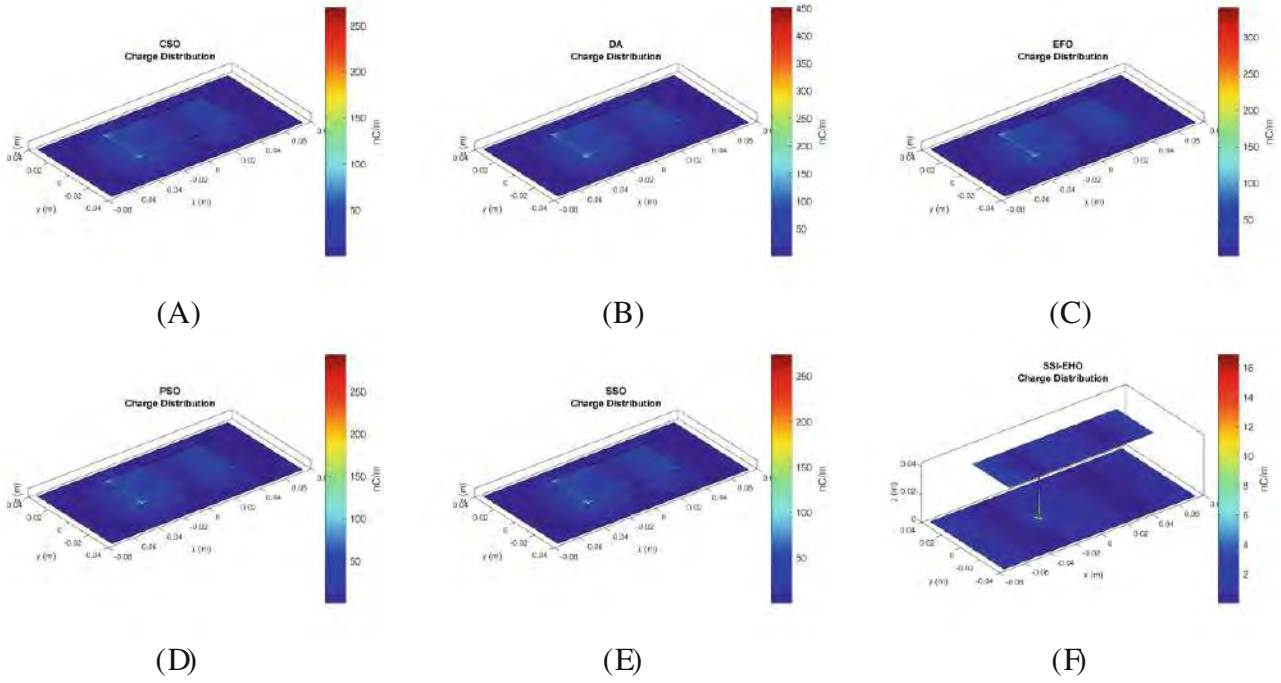


FIGURE 5 Charge distribution of optimized MPA for (A) CSO, (B) DA, (C) EFO, (D) PSO, (E) SSO, and (F) SSI-EHO

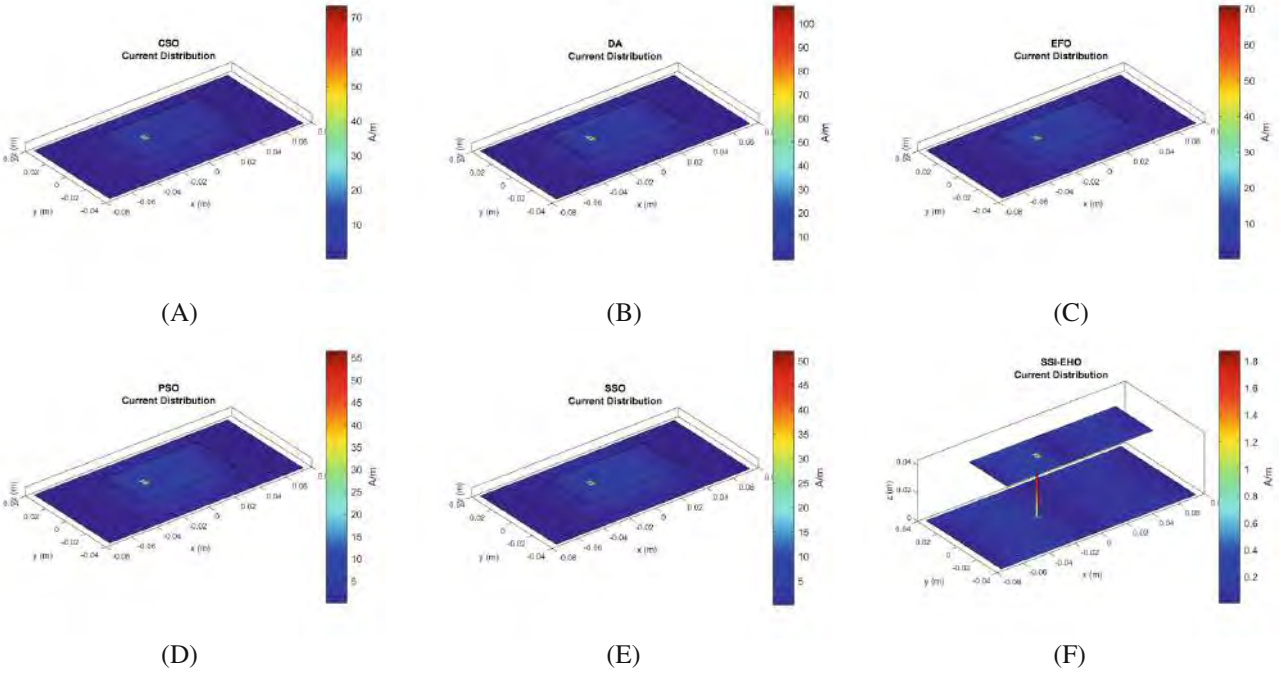


FIGURE 6 Current distribution of optimized MPA for (A) CSO, (B) DA, (C) EFO, (D) PSO, (E) SSO, and (F) SSI-EHO

Figure 6 exhibits the current distribution of the SSI-EHO model over the conventional models such as DA, CSO, PSO, SSO, and EFO for MPA design.

6.5 | Convergence analysis

The convergence analysis of the SSI-EHO scheme over extant schemes is exposed in Figures 7, 8, 9, 10, 11, and 12. The examination is done regarding cost for varied frequencies like, 77, 74, 72, 70, 68, and 65 GHz. Consequently, the cost is computed as: (1/gain = cost). On analyzing the outcomes, the convergence (cost) value acquired by the suggested SSI-EHO model is lesser than the existing ones for every frequency. Specifically, a smaller cost value of 44 790 000 is accomplished by the developed model from iteration 10 to iteration 20, when the frequency is at 65 GHZ. Moreover, the cost values accomplished by the developed scheme at 65 GHZ is much minimal than the cost values accomplished by the developed scheme at other frequencies such as 77, 74, 72, 70, and 68 GHz. Thus, the primacy of offered SSI-EHO technique was established from the convergence outcomes.

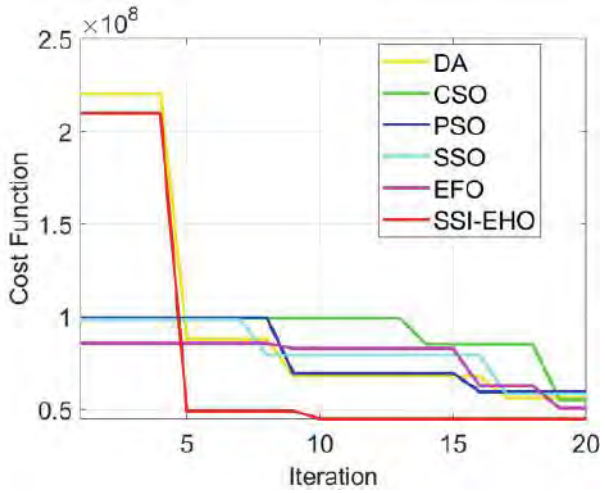


FIGURE 7 Convergence analysis of deployed approach over extant approaches for 65 GHz

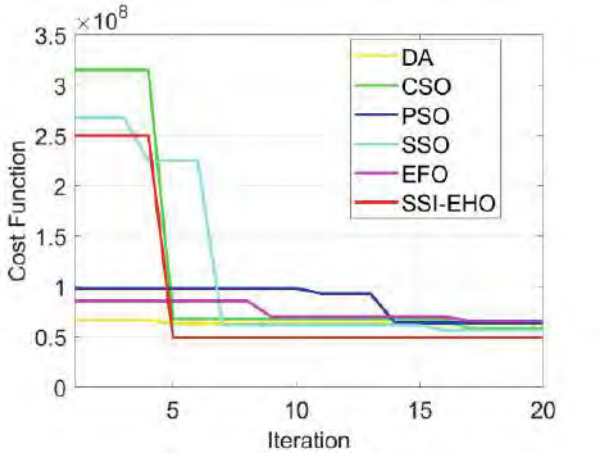


FIGURE 8 Convergence analysis of deployed approach over extant approaches for 68 GHz

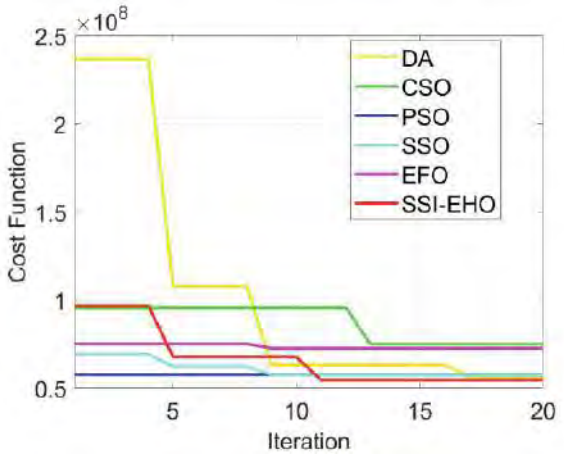


FIGURE 9 Convergence analysis of deployed approach over extant approaches for 70 GHz

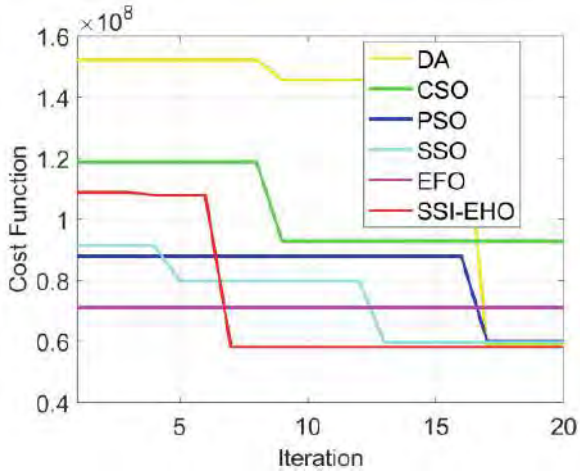


FIGURE 10 Convergence analysis of deployed approach over extant approaches for 72 GHz

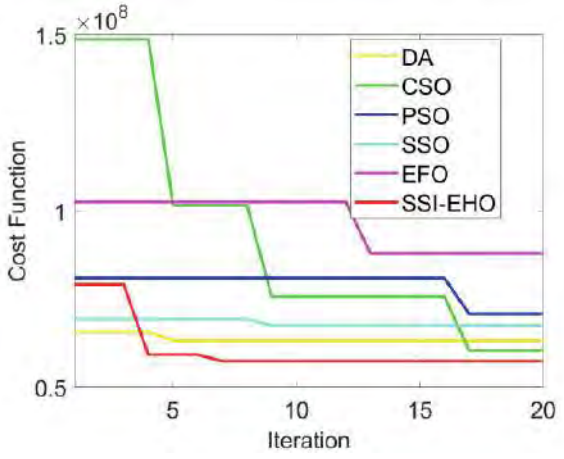


FIGURE 11 Convergence analysis of deployed approach over extant approaches for 74 GHz

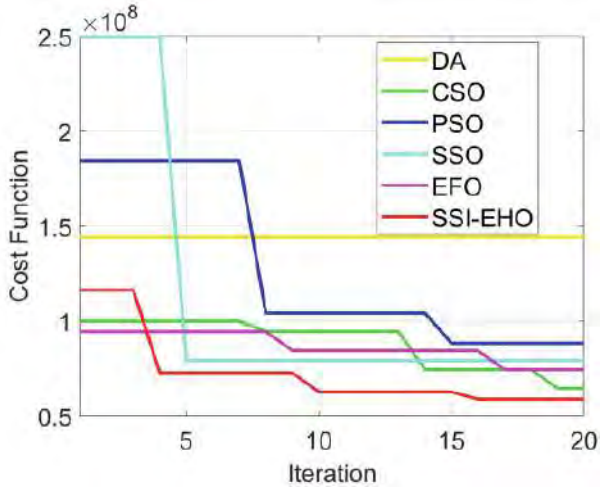


FIGURE 12 Convergence analysis of deployed approach over extant approaches for 77 GHz

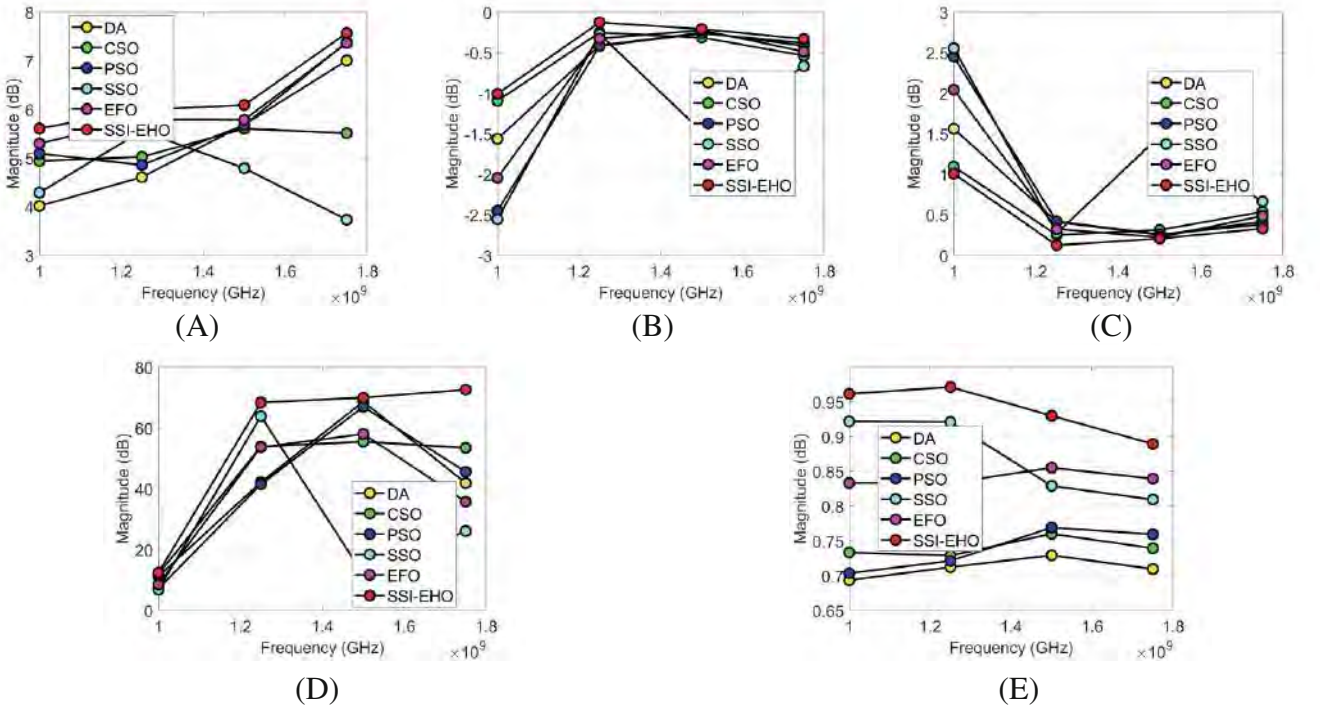


FIGURE 13 Performance analysis of the adopted scheme over extant schemes regarding (A) gain, (B) reflection coefficient, (C) return loss, (D) VSWR, and (E) efficiency

6.6 | Performance analysis

The analysis of varied constraints (efficiency, reflection coefficient, VSWR, return loss, and gain) by altering frequency is described in this section. Here, frequency is varied from 1×10^9, 1.2×10^9, 1.4×10^9, 1.6×10^9, 1.8×10^9, and 2×10^9 GHz. In fact, "Gain is defined as the ratio of the radiation intensity, in a given direction, to the radiation intensity that would be obtained if the power accepted by the antenna were radiated isotropically." Thus, gain has to be high to achieve an improved antenna system. On scrutinizing the graph in Figure 13A, the gain is increasing with an increase in frequency. Particularly, when the frequency reaches 1.75×10^9, the suggested SSI-EHO scheme has accomplished a higher gain of 7.5 dB. Likewise, analysis on the reflection coefficient is observed from Figure 13B. "The reflection coefficient is a

TABLE 2 Computational analysis of proposed method over existing methods

Frequency	DA	CSO	PSO	SSO	EFO	SSI-EHO
1×10^9	3664.6	4057.5	4048.1	4076.4	3626.6	3425.6
1.2×10^9	3940.1	3926	3984.5	4069.4	3776.9	3719.9
1.4×10^9	3592.8	3966.7	4070.6	3804.1	3858.1	3826.4
1.6×10^9	3719	3959.2	3914.2	3988.6	4069.6	3247.4
1.8×10^9	3600.2	3778.4	3594.8	3649.5	4073.4	3524.6
2×10^9	3626.4	3914	4013.8	3793.6	3657.6	3484.2

parameter that describes how much of an electromagnetic wave is reflected by an impedance discontinuity in the transmission medium. It is equal to the ratio of the amplitude of the reflected wave to the incident wave, with each expressed as phasors & is denoted by the symbol gamma Γ." It can be noted that there is a sudden rise in reflection coefficient values when the frequency reaches 1.23×10^9 GHz. Then, there is a gradual drop down of reflection coefficient values from frequency 1.24×10^9 to 1.75×10^9 GHz. However, the adopted SSI-EHO scheme has attained higher values when compared to extant schemes. The analysis on return loss is observed from Figure 13C. In fact, the return loss should be minimal for the enhanced performance of the MPA. Accordingly, a rapid fall of return loss values is noticed from 1×10^9 to 1.23×10^9 GHz. After that, a gradual rise in return loss is observed for both developed and compared models, however, the adopted SSI-EHO scheme is found to have attained much least loss values than the compared ones.

In Figure 13D, the analysis output based on VSWR is achieved. VSWR[46] is a "measure that numerically describes how well the antenna is impedance matched to the radio or transmission line it is connected to." The arithmetical formulation for VSWR is shown in Equation (35), wherein Γ denotes reflection coefficient.

$$\mathrm{VSWR} = \frac{1+|\Gamma|}{1-|\Gamma|}. \tag{35}$$

A higher VSWR value ensures the superior performance of the system. Here, when the frequency reaches 1.75×10^9 GHz, a high VSWR value (68 dB) is acquired by the developed SSI-EHO model, whereas, at other frequency ranges, the adopted model has acquired comparatively minimal values. The efficiency accomplished by the adopted scheme over other compared schemes is shown in Figure 7E. "Antenna efficiency is the ratio of power radiated by the antenna to the power supplied to the antenna." Here, it can be noticed that the efficiency decreases with an increase in frequency. However, at 1.23 GHz, a high-efficiency value (0.97 dB) is attained by the adopted scheme. Nevertheless, the introduced scheme poses higher efficiency than all the compared schemes. Thus, the enhancement of deployed SSI-EHO method is ensured by the integrated optimization concept.

6.7 | Computational analysis

This section discusses the computational analysis of adopted and extant schemes, which is shown in Table 2. For analysis, the frequency is varied from 1×10^9, 1.2×10^9, 1.4×10^9, 1.6×10^9, 1.8×10^9 and 2×10^9 GHz. The proposed model is 6.52%, 15.57%, 15.37%, 15.96%, and 5.54% better than DA, CSO, PSO, SSO, EFO, and SSI-EHO, algorithms. From Table 2, we can conclude that the SSI-EHO method takes very lesser time to calculate over other methods.

7 | CONCLUSION

This work developed a novel MPA model by incorporating the optimization theory named as SSI-EHO. Consequently, the MPA constraints like patch height, substrate width, patch length, and substrate length were optimized via developed SSI-EHO for acquiring amplified gain and minimal return loss. Finally, an examination was done that proved the effectiveness of the developed work. On scrutinizing the graph, the convergence (cost) value acquired by the suggested SSI-EHO model was lesser than the existing ones for every frequency. Specifically, a smaller cost value of 44 790 000 was accomplished by the developed model from iteration 10 to iteration 20, when the frequency is at 65 GHZ. Moreover, the cost

values accomplished by the developed scheme at 65 GHZ was much more minimal than the cost values accomplished by the developed scheme at other frequencies. In the future, the antenna geometry and power handling capability should be concerned more with enhancing the system's performance.

NOMENCLATURE

Acronyms	Descriptions
CSO	Cat swarm optimizer
CI	Characteristic impedance
XP	Cross-polarization
DA	Dragonfly algorithm
ERI	Effective refractive index
EFO	Electric fish optimization
EMF	Electromagnetic field
EHF	Extremely high frequency
FEM	Finite element method
FAR	Field asymmetry ratio
FR4	Fire resistant 4
MPA	Microstrip patch antenna
MIS	Modified-I-shaped
MTM	MIS electromagnetic metamaterials
NPS	Non-point-symmetric
PSO	Particle swarm optimization
PS	Point-symmetric
RMPA	Rectangular microstrip patch antenna
SSO	Shark smell optimization
SSI-EHO	Shark smell integrated EHO
VSWR	Voltage standing wave ratio

DATA AVAILABILITY STATEMENT

Data sharing is not applicable to this article as no new data were created or analyzed in this study.

REFERENCES

1. Jain S Mishra PK Mishra J. Design and analysis of moisture content of Hevea latex rubber using microstrip patch antenna with DGS. *Mater Today Proc*. 2020;29(2):556–560.
2. Koziel S, Ogurtsov S, Zieniutycz W, Sorokosz L. Expedited design of microstrip antenna subarrays using surrogate-based optimization. *IEEE Antennas Wirel Propag Lett*. 2014;13:635-638.
3. Kushwaha RK, Karuppanan P, Malviya LD. Design and analysis of novel microstrip patch antenna on photonic crystal in THz. *Physica B Condens Matter*. 2018;545:107–112.
4. Khandelwal MK, Kanaujia BK, Dwari S, Sachin K, Gautam AK. Analysis and design of wide band microstrip-line-fed antenna with defected ground structure for Ku band applications. *AEU - Int J Electron Commun*. 2014;68(10):951–957.
5. Sharma N, Sharma V. A design of microstrip patch antenna using hybrid fractal slot for wideband applications. *Ain Shams Eng J*. 2018;9:2491-2497.
6. Prema N, Kumar A. Design of multiband microstrip patch antenna for C and X band. *Optik*. 2016;127(20):8812-8818.
7. Mathur D, Bhatnagar SK, Sahula V. Quick estimation of rectangular patch antenna dimensions based on equivalent design concept. *IEEE Antennas Wirel Propag Lett*. 2014;13:1469-1472.
8. Kaviyaraj R, Uma M. A survey on future of augmented reality with AI in education. Paper presented at: 2021 International Conference on Artificial Intelligence and Smart Systems (ICAIS); 2021:47-52; IEEE.
9. El Misilmani HM, Naous T. Machine learning in antenna design: an overview on machine learning concept and algorithms. Paper presented at: 2019 International Conference on High Performance Computing & Simulation (HPCS); 2019:600-607; IEEE.
10. Khatiwada P, Subedi M, Chatterjee A, Gerdes MW. Automated human activity recognition by colliding bodies optimization-based optimal feature selection with recurrent neural network. arXiv preprint. arXiv:2010.03324; 2020.
11. Tran K, Nguyen T. Preliminary research on the social attitudes toward the AI's involved Christian education in Vietnam: promoting AI Technology for Religious Education. *Religions*. 2021;12(3):208.
12. Tang Y, Bilodeau GA. Evaluation of trackers for Pan-Tilt-Zoom Scenarios. arXiv preprint. arXiv:1711.04260; 2017.

13. Yang X, Geyi W, Sun H. Optimum Design of Wireless Power Transmission System Using Microstrip Patch Antenna Arrays. *IEEE Antennas Wirel Propag Lett*. 2017;16:1824-1827.
14. Bhongale SR, Ingavale HR, Shinde TJ, Vasambekar PN. Microwave sintered Mg-Cd ferrite substrates for microstrip patch antennas in X-band. *AEU - Int J Electron Commun*. 2018;96:246-251.
15. Saxena NK, Kumar N, Pourush PKS. Radiation characteristics of microstrip rectangular patch antenna fabricated on LiTiMg ferrite substrate. *AEU - Int J Electron Commun*. 2015;69(12):1741-1744.
16. Di Mauro M, Galatro G, Longo M, et al. Availability evaluation of a virtualized IP multimedia subsystem for 5G network architectures. *Safety Reliab Theory Appl*. 2017;2203-2210.
17. Di Mauro M, Galatro G, Longo M, Postiglione F, Tambasco M. Comparative performability assessment of SFCs: the case of containerized IP multimedia subsystem. *IEEE Trans Netw Serv Manag*. 2020;18(1):258-272.
18. Kumhar SH, Kirmani MM, Sheetlani J, Hassan M. Sentiment analysis of Urdu language on different social media platforms using Word2vec and LSTM. *Turk J Comput Math Educ*. 2020;11(3):1439-1447.
19. Gupta M, Mathur V. Koch boundary on the square patch microstrip antenna for ultra wideband applications. *Alex Eng J*. 2018;57(3):2113-2122.
20. Emadeddin A, Shad S, Rahimian Z, Hassani HR. High mutual coupling reduction between microstrip patch antennas using novel structure. *AEU – Int J Electron Commun*. 2017;71:152-156.
21. Nuangpirom P, Klinbumrung K, Tangthong N, Akatimagool S. Wave iterative computation for fractal microstrip patch antenna. *Procedia Comput Sci*. 2016;86:39-42.
22. Anantha B, Merugu L, Rao PVDS. A novel single feed frequency and polarization reconfigurable microstrip patch antenna. *AEU – Int J Electron Commun*. 2017;72:8-16.
23. Agarwal A, Mehta SN. Design and performance analysis of MIMO-OFDM system using different antenna configurations. Paper presented at: 2016 International Conference on Electrical, Electronics, and Optimization Techniques (ICEEOT); 2016:1373-1377; IEEE.
24. Misra G, Agarwal A, Agarwal K. Design and performance evaluation of microstrip antenna for ultra-wideband applications using microstrip feed. *Am J Electr Electron Eng*. 2015;3(4):93-99.
25. Naderi M, Zarrabi FB. Fereshteh Sadat Jafari, Speideh Ebrahimi, "fractal EBG structure for shielding and reducing the mutual coupling in microstrip patch antenna array". *AEU - Int J Electron Commun*. 2018;93:261-267.
26. Bharathi A, Merugu Lakshminarayana PVD, Rao S. A quad-polarization and frequency reconfigurable square ring slot loaded microstrip patch antenna for WLAN applications. *AEU - Int J Electron Commun*. 2017;78:15-23.
27. Marsaline Beno M, Valarmathi IR, Swamy SM, Rajakumar BR. Threshold prediction for segmenting tumour from brain MRI scans. *Int J Imaging Syst Technol*. 2014;24(2):129-137. doi:10.1002/ima.22087
28. Poornima S, Dutta K, Chandramma S. Flexible and miniaturized design of microstrip patch antenna with improved cross-polarized radiation. *AEU - Int J Electron Commun*. 2020;116:153083.
29. Li T-q, Ma B, Lv W-r. A novel design of microstrip patch antenna array with modified-I-shaped electromagnetic metamaterials applied in microwave wireless power transmission. *Optik*. 2018;173:193-205.
30. Chandrashekar KS, Dutta K, Chandramma S. An analytical approach of designing compact microstrip patch antenna using metal-ring superstrate for wideband and broadside radiations. *AEU – Int J Electron Commun*. 2020;127:153437.
31. Rashmitha R, Niran N, Ahmed MR. Microstrip patch antenna design for fixed mobile and satellite 5G communications. *Procedia Comput Sci*. 2020;171:2073-2079.
32. Hocini A, Temmar MN, Zamani M. Novel approach for the design and analysis of a terahertz microstrip patch antenna based on photonic crystals. *Photonics Nanostruct Fundam Appl*. 2019;36:100723.
33. Samanta S, Reddy PS, Mandal K. Field asymmetry ratio: a new quantitative parameter to select microstrip antenna geometries for low cross-polarization application. *AEU – Int J Electron Commun*. 2020;128:153519.
34. Shalini MM, Madhan G. Design and analysis of a dual-polarized graphene based microstrip patch antenna for terahertz applications. *Optik*. 2019;194:163050.
35. Davoudabadifarahani H, Ghalamkari B. High efficiency miniaturized microstrip patch antenna for wideband terahertz communications applications. *Optik*. 2019;194:163118.
36. Guttula R, Nandanavanam VR. Mutation probability-based lion algorithm for design and optimization of microstrip patch antenna. *Evol Intell*. 2020;13:331-344. doi:10.1007/s12065-019-00292-9
37. Yilmaz S, Sen S. Electric fish optimization: a new heuristic algorithm inspired by electrolocation. *Neural Comput Appl*. 2020;32:11543-11578.
38. Mohammad-Azari S, Bozorg-Haddad O, Chu X. Shark smell optimization (SSO) algorithm. In: Bozorg-Haddad O, ed. *Advanced Optimization by Nature-Inspired Algorithms. Studies in Computational Intelligence*. Vol 720. Singapore: Springer; 2018. doi:10.1007/978-981-10-5221-7_10
39. Thomas R, Rangachar MJS. Hybrid optimization based DBN for face recognition using low-resolution images. *Multimed Res*. 2018;1(1):33-43.
40. Mukhedkar MM, Kolekar U. Hybrid PSGWO algorithm for trust-based secure routing in MANET. *J Netw Commun Syst*. 2019;2(3):1-10.
41. Roy RG. Rescheduling based congestion management method using hybrid Grey wolf optimization - grasshopper optimization algorithm in power system. *J Comput Mech Power Syst Control*. 2019;2(1):9-18.
42. Jafari M, Chaleshtari MHB. Using dragonfly algorithm for optimization of orthotropic infinite plates with a quasi-triangular cut-out. *Eur J Mech A Solids*. 2017;6:1-146.

43. Chu SC, Tsai PW, Pan JS. Cat swarm optimization. Paper presented at: Conference in Lecture Notes in Computer Science; 2014.
44. Zhang J, Xia P. An improved PSO algorithm for parameter identification of nonlinear dynamic hysteretic models. *J Sound Vib*. 2017;389:153-167.
45. Rao KN, Meshram V, Suresh HN. Optimization of ultra wide band communication TCA and AVA by using PSO and IGWO-FU algorithm; 2021. doi:10.21203/rs.3.rs-413477/v1.
46. Kamo B, Cakaj S, Kolici V, Mulla E. Simulation and measurements of VSWR for microwave communication systems. *Int J Commun Netw Syst Sci*. 2012;5:767-773. doi:10.4236/ijcns.2012.511080

How to cite this article: Subhadra Sarngadharan A, Narasimhamurthy R, Sankaramoorthy B, Singh SP, Singh C. Hybrid optimization model for design and optimization of microstrip patch antenna. *Trans Emerging Tel Tech*. 2022;e4640. doi: 10.1002/ett.4640

The peak-to-average power ratio reduction using hybrid scheme with companding and discrete Hartley transform for orthogonal frequency division multiplexing system

N. Rashmi*

Department of Electronics and Communication Engineering,
BMS Institute of Technology and Management,
Doddaballapur Main Road, Avalahalli,
Yelahanka, Bengaluru-560064, India
Email: rashmiswamy@bmsit.in
*Corresponding author

Mrinal Sarvagya

Department of Electronics and Communication Engineering,
REVA University, India
Email: mrinalsarvagya@reva.edu.in

Abstract: Orthogonal frequency division multiplexing is a multicarrier and high data rate system. Increasing data rate through such modulation techniques increases amplitude variation with large dynamic range usually referred as peak-to-average power ratio (PAPR) in OFDM systems. As a consequence, power amplifier need to operate in linear region to avoid nonlinear distortion. In this article, to combat the nonlinearity effect in communication systems we propose a novel hybrid scheme, combined with μ-law companding and discrete Hartley transform (DHT) spread. We present the detailed simulation results for PAPR reduction for OFDM and superposition coded modulation-OFDM. It is observed that SCM-OFDM performance is much superior to OFDM. In this article we proposed a scheme for SCM-OFDM system that reduces PAPR to 1.7949 dB. The article also compares the obtained results with clipping technique, partial transmit sequence, selective mapping. The computational complexity of the system using SCM-OFDM is reduced by 50% by applying hybrid scheme.

Keywords: clipping technique; selective mapping; partial transmit sequence; discrete Hartley transform; DHT; superposition coded modulation; SCM-OFDM; modified μ-law companding; MMC; peak-to-average power ratio; PAPR; bit error rate; BER; complementary commulative distribution function; CCDF.

Reference to this paper should be made as follows: Rashmi, N. and Sarvagya, M. (2022) 'The peak-to-average power ratio reduction using hybrid scheme with companding and discrete Hartley transform for orthogonal frequency division multiplexing system', *Int. J. Intelligent Enterprise*, Vol. 9, No. 3, pp.303–317.

Biographical notes: N. Rashmi is working as an Assistant Professor in the Department of Electronics and Communication Engineering at the BMS Institute of Technology and Management, Doddaballapur Main Road, Avalahalli, Yelahanka, Bengaluru-560064, India.

Mrinal Sarvagya is a Professor in the School of ECE at the REVA University. Her research interests are wireless communication systems, communication networks, communication theory and systems.

1 Introduction

Orthogonal frequency division multiplexing (OFDM) is an assured technique for systems that require a high data transmission rate, such as 4G LTE-advanced and WiMAX. Several researchers and academicians have studied the OFDM technique, which can cater to high data rate applications. OFDM offers several advantages over channels that experience frequency selectivity and time variations. Furthermore, the OFDM technique allows each subcarrier to independently select the constellation and coding scheme. OFDM offers robustness to the multipath fading channel and has a low-implementation complexity. Although OFDM exhibits several advantages, OFDM suffers from various technical challenges, such as high peak-to-average power ratio (PAPR) and carrier frequency offset (CFO). The complexity of digital-to-analogue converter (DAC) increases with PAPR, that leads to the degradation of power amplifier efficiency in OFDM system. The transmission of high-PAPR OFDM signals through nonlinear power amplifier causes spectral broadening, that will expand the dynamic range of the DAC. Superposition coded modulation (SCM) is emerging as non-bijective, bandwidth and power efficient coded modulation with Gaussian quadrature components. The shaping gain in SCM is inherent without any additional complexity at the transmitter side. At the receiver side the complexity is low as $O(K)$ for K-layer SCM when equal power is allocated to the symbols. Hybrid combination of multi layer SCM and OFDM can achieve greater data rates compared to single carrier OFDM system.

The main aim of this study is to decrease the PAPR in OFDM system to reduce the range of DAC and thereby decrease the cost of the system. We propose the use of a novel technique called hybrid precoding using DHT and μ companding for SCM-OFDM system. SCM-OFDM system increases the data rate of OFDM system complementing the advantages of OFDM. Implementing SCM coding scheme for MIMO system reaches Shannon channel capacity. But, using SCM in OFDM system further increases the PAPR problem in OFDM but increases the data rate and capacity of the OFDM system. Several techniques for PAPR reduction have been proposed in literature, and are categorised into three categories: multiplicative schemes, coding techniques and additive schemes. In multiplicative schemes, phase sequences are multiplied with OFDM sequences, the popular techniques using multiplicative schemes such as selective mapping (SM) and partial transmit sequence (PTS) (Vittal and Naidu, 2017; Liang et al., 2014; El-Hamed et al., 2018; Hassan, 2016). In Satyavathi and Rao (2019), a low-complexity PAPR reduction with modified linear SM scheme was suggested for the OFDM system. Lee and Chung (2018) proposed a scheme to reduce hardware complexity by analysing loglikelihood ratio for channel coding there by reducing PAPR. The performance for

modified linear SM scheme for PAPR reduction in Satyavathi and Rao (2019) was insignificant compared with conventional schemes. Additive schemes include tone reservation (TR) and clipping and filtering techniques. Nagaraju et al. (2018) analysed the SM and PTS technique for PAPR reduction and performance analysis of modified SM technique based on BER and complexity of the system for PAPR were reported. TR technique with low-complexity tones with null subcarriers was proposed in Wang et al. (2016) to decrease the PAPR in the filter bank multi carrier systems combined with OFDM system. In Azim et al. (2017), decision directed method proposed for PAPR reduction for optical OFDM. In Nadal et al. (2014), clipping and quantisation noise cancellation is proposed for low-complexity PAPR reduction. Clipping PAPR reduction technique causes in- and out-band interferences that take out symmetry among subcarriers of OFDM that result in inter carrier interference (ICI). A PAPR reduction system including coding technique anticipated in Chen and Liang (2007) implemented Reed Muller codes are implemented with PTS scheme. All these schemes reduces PAPR at cost of computational complexity and with additional information need to be transmitted to receiver, that increases bandwidth requirement. Motivated by the limitations of clipping and other techniques, we propose modified μ-law companding techniques with discrete Hartley transform (DHT) as precoding technique for PAPR reduction in SCM-OFDM system. The proposed technique does not require any side information as compared to multiplicative methods, precoding techniques scrambles the signal phase to reduce the PAPR of the OFDM system. Precoding with DHT demonstrates higher PAPR reduction compare to discrete Fourier transform (DFT) technique. DHT distorts phase of the signal there by reducing PAPR. DHT is less complex compared to DFT as indicated in Table 2. DHT brings down the multiplication complexity by a factor of two contrasts with DFT, since DHT matrix has only real elements. DHT provides higher spectral efficiency, since it does not require side information compared to various different PAPR reduction systems like PTS and SLM techniques.

This paper is organised as follows. In Section 2, the SCM-OFDM system is discussed, and the PAPR and cumulative complementary distribution function (CCDF) are formulated. In Section 3, the clipping technique for PAPR reduction for SCM-OFDM system, PTS and SM techniques were discussed. Section 4 is the simulation results were discussed and Section 5 is the conclusion.

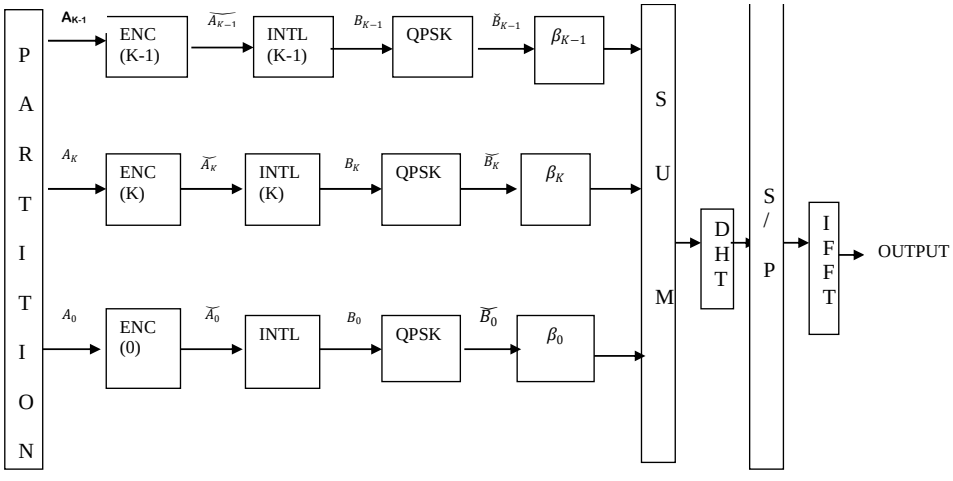
2 SCM-OFDM system

2.1 Characterisation of the SCM-OFDM system

OFDM system has high bandwidth efficiency compared to other multiple access system. However, there is continuous demand for high data rate and high capacity to meet the requirements of various evolving technologies and applications. Therefore there is strong motivation to come up with systems and methods that achieve even higher bandwidth efficiency compared to traditional OFDM system which inherits various advantages of traditional OFDM system. SCM is the multicarrier system, which has the advantage of adaptive modulation and coded techniques. Combining SCM with OFDM system further increases channel capacity and the data rate of the OFDM system, preserving the

advantages of traditional OFDM system. However use of SCM system for OFDM further increases the PAPR problem in OFDM system.

Figure 1 The proposed SCM-OFDM system for PAPR reduction



To reduce PAPR in SCM-OFDM system we propose DHT which provides spectral efficiency to SCM-OFDM system. In SCM-OFDM system, K-layer SCM is defined over a 2^K constellation, the binary data sequence P is partitioned into K sub-sequences $\{P_k\}$. Rashmi and Sarvagya (2016) proposed a transmitter for SCM-OFDM system. The k^{th} layer, resulting in a coded bit sequence $c_k = \{c_k(j)\}$ of length $2J$. Where $c_k(j) \in \{0, 1\}$and J is the frame length. Coded sequence c_k is then mapped to QPSK sequence $x_k(j)$.

$$x_k^K = x_{re,j}^K + ix_{im,j}^K \tag{1}$$

The superscripts 'Re' and 'Im' denote the real and imaginary components, of complex numbers, respectively.

QPSK sequences are linearly superimposed to form X_j, which is given as follows:

$$x(j) = \sum_{k=0}^{K-1} \rho_k x_k(j),\ j = 0, 1, \ldots, J-1 \tag{2}$$

where $\{\rho_k\}$ are constant weighting colour k. In this study, $\{\rho_k\}$ was obtained using the simulation-based power allocation method projected in Chen and Liang (2007).

The superimposed symbols are fed to a customary OFDM modulator unit that consists of an N-point inverse IDFT unit pursued by an N_g-point cyclic prefix (CP) converts linear convolution to circular convolution, applied as a guard band to avoid inter symbol interference (ISI). The IDFT output is obtained as vector $D_i = [D_i(0), D_i(1), \ldots, D_i(N-1)]^T$. The CP is appended to D_i such that

$$\begin{aligned} &S_i(n) = S_i(n+N) \text{ for } -N_g \le n \le N-1 \\ &S_i(k) = \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_i(k) e^{j2\pi nk/N} \right\}, \end{aligned} \tag{3}$$

where $X_i(k) = x(j)$ from equation (2) is the superimposed modulated data symbol assigned to the sub carriers and N_g represents the length of CP ($-N_g < n < N-1$).

2.2 Peak-to-average power ratio

The PAPR is important for enabling the high-power amplifier to operate in the linear region. The SCM-OFDM signal is oversampled '*L*' times to better approximate the PAPR. The oversampled SCM-OFDM signal in time domain is given as follows:

$$S_i(k) = \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} X_j(k)^{ej2\pi nk/LN} \right\}, 0 < n < LN \tag{4}$$

The PAPR is termed as the ratio of the maximum instantaneous power to the average power and is formulated mathematically as follows:

$$PAPR[S[n]] = \frac{\max\left[|s[K]|^2\right] 0 \le n \le LN-1}{E\left[|s[K]|^2\right]} \tag{5}$$

where $E[.]$ specifies the expectation operator.

2.3 Complementary cumulative distribution function

The CCDF refers to the probability that the distribution of the output power of the SCM-OFDM signal exceeds the predefined threshold value. It is done by determining the CCDF for the PAPR values. The CCDF is formulated and expressed as follows:

$$CCDF = \Pr\left(PAPR_{SCM-OFMD} > PAPR_o\right), \tag{6}$$

where $PAPR_{SCM-OFDM}$ and $PAPR_o$ are the output of the SCM-OFDM system and the threshold value respectively.

3 PAPR reduction techniques

3.1 Clipping technique for reducing the PAPR in the SCM-OFDM system

In the SCM-OFDM system, the PAPR is maximum for a given $\{\rho_k\}$ when all θ_k's are equal. To reduce the PAPR, clipping of X_j to form $\bar{X}_j$ is followed based on the rule presented below:

$$\bar{X}_j = \begin{cases} s_j, & |s_j| \le A \\ \dfrac{As_j}{|s_j|}, & |s_j| > A \end{cases} \tag{7}$$

where $A > 0$ is the clipping threshold.

From Liang (2015), we outline the clipping ratio as follows:

$$\gamma = A^2 \Big/ E\left(\left|X_j^2\right|\right) \tag{8}$$

The PAPR of the transmitted signal with clipping is given as follows

$$PAPR = A^2 \Big/ E\left(\left|\bar{X}_j^2\right|\right).$$

The received signal can be written as follows:

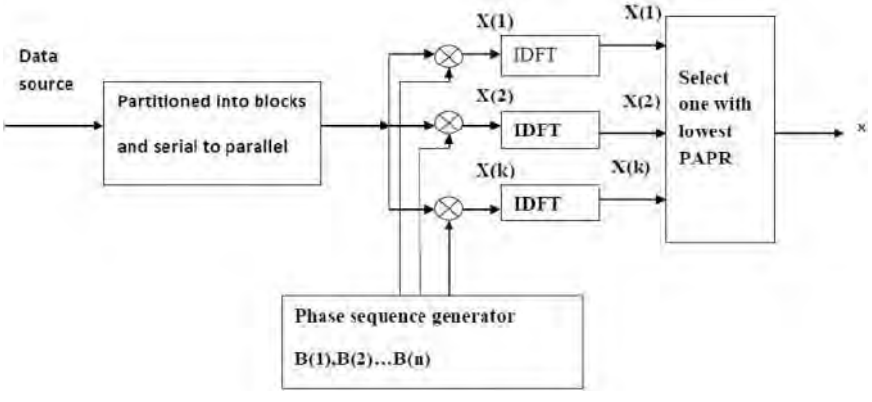
$$Y_j = X_j + w_j, \tag{9}$$

where w_j is the complex Gaussian noise with a zero mean and σ^2 variance.

3.2 Selective mapping

SM is a multiplicative PAPR reduction technique. The SCM-OFDM signal is copied and each copy is multiplied with different phase sequences, as a result different PAPR values are captured. The SM signal is a product of SCM-OFDM and phase sequences and represented by $X(1)$, $X(2)$, $X(3)$, …, $X(k)$, where k is the number of SCM-OFDM signals. Inverse DFT is performed on $X(k)$. $X(k)$ with lowest PAPR is selected and transmitted. With increase in number of k the performance of SM raises but as the k increases number of IDFTs also increases, increasing the complexity of the system. The block diagram of the SM is presented in Figure 2. This technique also requires side information. The side information is an extra bit that carries the phase sequences to receiver increasing the bandwidth. SM technique is not only spectral inefficient but also complex system.

Figure 2 Block diagram of SM



3.3 Partial transmit sequence

PTS is selection transmission technique with point-to-multipoint mapping, multiply single input data signals by a phase factor to map multiple candidate signals. Similar to SM technique, candidate signal with lowest PAPR is selected as the OFDM transmission signal. PTS also requires side information to be transmitted to receiver to retrieve which candidate signal was selected by the transmitter. Let X denote the random signal in frequency domain. Now X is partitioned into U disjoint subblocks represented by $\{X^{(u)}, u = 1, 2, \ldots, U\}$ where $X^{(u)}$ is given by

$$X^{(u)} = \left[X_1^{(u)}, X_2^{(u)}, \ldots, X_{M-1}^{(u)}\right] \tag{10}$$

$$X = \sum_{u=1}^{U} X^u \tag{11}$$

The phase rotation factors are given by

$$C_u = e^{j\theta u},\ u = 1, 2, \ldots, U \tag{12}$$

4 Proposed MMC-DHT PAPR reduction technique for SCM-OFDM system

The objective of this investigation is to realise PAPR reduction before OFDM modulation using modified companding and DHT precoding. Since the SCM symbols takes Gaussian distribution, it implies that there is no requirement of active shaping filters which is used to adapt the advanced modulation schemes such as QAM to Gaussian channel, thereby reducing the shaping filter burden on transmitter and receiver of the system. The cumulative distributed function (CDF) of the un-companded signal is expressed as

$$F_x(\cdot) = 1 - \exp\left(\frac{x_0^2}{\sigma_x^2}\right),\ x_o > 0 \tag{13}$$

where x_o is that the discrete envelope of $x(n)$ and σ_x^2 is the variance.

The μ companding was introduced based on Ghahremani and Shayesteh (2014)

$$F_{EC}\left(x(n)\right) = \text{sgn}(x) \cdot \sqrt{3\sigma_x^2} \cdot erf\left(\frac{|x|}{\sqrt{2\sigma_x^2}}\right),\ 0 \le x \le 1 \tag{14}$$

The above equation is the fundamental error function due to companding transform.

μ-law companding with constant value for μ applied to SCM-OFDM signal is given by

$$\begin{aligned} y_c(n) &= F\left(x_j(n)\right) \\ &= \text{sgn}\left(x_j(n)\right) \frac{A_s \times \log\left(1 + \mu\left|\frac{x_j(n)}{A_s}\right|\right)}{\log(1+\mu)} \end{aligned} \tag{15}$$

where $\text{sgn}(x(n)) = \dfrac{x(n)}{|x(n)|}$ is the phase and $A_s = \max(|x_j(n)|$.

The multi μ is simulated for SCM-OFDM system for better PAPR reduction. The fundamental problem in μ companding is, it enlarges the amplitudes of the lower amplitude signal keeping higher amplitude signals unchanged at the output of the compander. Thus increasing the average power and reduces the PAPR. However, this technique leads to unfair improvement of BER when compared to uncompanded signal. The companding profile known as peak ratio is presented in Ghahremani and Shayesteh (2014), it is articulated as the ratio of amplitude A of the signal specified in μ compander to the peak amplitude of the actual signal to be companded, i.e.,

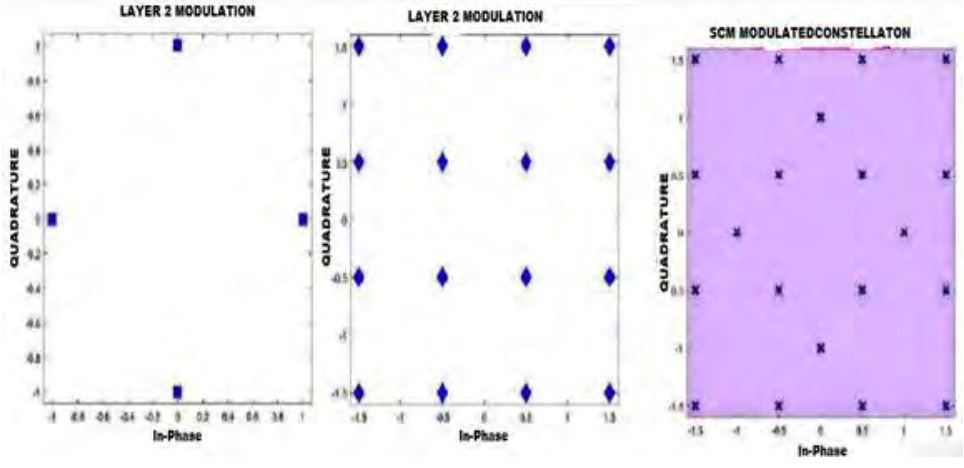
$$K = \frac{Peak\ amplitude\ of\ compressor}{Peak\ of\ actual\ signal} = \frac{A}{x_{jpeak}} \tag{16}$$

The transfer characteristics of the modified μ compander including new parameter peak ratio '*K*' is expressed by substituting the above equation in equation (15).

$$y = K \times x_{jpeak} \times \frac{\log\left(1+\mu\frac{|x_{jpeak}|}{K \times x_{jpeak}}\right)}{\log(1+\mu)}\text{sgn}(x) \tag{17}$$

The above companding profile permits the *μ-law* companding profiles to be modified such that all amplitudes as well as the peaks of the input signal can be amplified by changing the value of *K*. Higher the values of *K*, greater is the gain for the peaks and much higher the gain for the lower amplitude signals.

Figure 3 Simulation result of SCM of two signals QPSK and 16 QAM modulated signal (see online version for colours)



OFDM signal is precoded using DHT precoder before IFFT block. The Hartley transform of $S_i(n)$ is expressed as

$$H_s(\omega) = \int_{-\infty}^{\infty} f(t)Cas(\omega t)dt. \tag{18}$$

$$Cas(\omega t) = \cos(\omega t) + \sin(\omega t) \tag{19}$$

$$f(t) = \int_{-\infty}^{\infty} H_g(\omega)Cas(\omega t)d\omega \tag{20}$$

To reduce the number of functional computation $Cas(\omega t)$ can be expressed as:

$$\sqrt{2}\cos\left(\omega t - \frac{\pi}{4}\right) = \sqrt{2}\sin\left(\omega t + \frac{\pi}{4}\right) \tag{21}$$

For discrete signal Hartley transform is expressed as:

$$H_{s(k)} = \frac{1}{\sqrt{2}} \sum_{n=0}^{NL-1} y_i(n) Cas\left(\frac{2\pi nk}{N}\right) \tag{22}$$

For discrete signal Hartley transform is expressed as:

$$H_{s(k)} = \frac{\sqrt{2}}{\sqrt{N}} \sum_{n=0}^{NL-1} y_i(n) \cos\left(\frac{2\pi nk}{N} \frac{\pi}{-4}\right) \tag{23}$$

$k = 0, 1, \ldots, N-1$.

The DHT rotates the phase of input SCM symbols. While precoding, the signal is re-established to the single carrier. SCM-OFDM is analysed in AWGN channel. The received signal is as follows:

$$Y_j(n) = \left\{ \frac{1}{\sqrt{N}} \sum_{n=0}^{N-1} H_s(k) e^{\frac{j2\pi nk}{LN}} \right\} + w(n) \tag{24}$$

$0 < n < LN - 1$.

5 Simulation results

In this study, computer simulations were performed to evaluate the performance of the SCM-OFDM system. CP is added to the time domain signal to avoid the ISI. Besides PAPR drawback in OFDM system, OFDM is additionally sensitive to spectral null problem over a frequency selective fading channel. By using DHT over the entire bandwidth the nulls are spread to increase the probability of correctly receiving the transmitted symbols. Use of precoding, further improves the PAPR performance of an OFDM multicarrier system. SCM-OFDM system is simulated in MATLAB 2017 environment. Each layer in multicarrier is coded using convolution encoder with rate 1/2. The coded code word is interleaved to remove any burst error present in the data stream. Each layer is modulated using different modulation schemes. We adopted QPSK and 16 QAM for the layer 1 and layer 2 respectively. The simulation result of SCM scheme with two layers, layer 1 adopted QPSK and for layer 2 16 QAM and resultant superimposed signal is shown in Figure 3. For the higher modulation scheme when SCM is simulated with 64 QAM PAPR obtained is 6.4089 dB, OFDM with 64 QAM is 9.31 dB. Compared to OFDM PAPR in SCM improved by 31.2%. However, PAPR in the hybrid system SCM-OFDM system is 13.53 dB. PAPR in hybrid system increases due to nonlinearity in SCM, Since Gaussian like transmitted signal has a relatively high PAPR. SCM-OFDM with clipping resulted in reduction of PAPR by 2.87 dB. Even PAPR reduction using clipping and filtering technique PAPR is reduced to 2.87 dB, but BER performance of the same reduction scheme is poor as shown in Figure 4 compared to BER performance of the proposed method. PAPR reduction using DHT for SCM-OFDM signal is 5.8971 dB, thereby reducing PAPR by 4.7565 dB compared to clipping and without reduction technique by 7.6329 dB respectively, thereby improving performance of PAPR by 56.41%. With proposed technique the PAPR is further reduced to 1.7949 dB and PAPR performance improvement by 86.7%. The proposed system is also simulated with 256 QAM, with 256 QAM PAPR in OFDM is as high as 30 dB due to nonlinearity in the 256 QAM PAPR will increase up to a greater extent in OFDM system. Computer simulation

is done and results are shown in Figure 9. With the proposed hybrid scheme the PAPR is reduced significantly to 12.5 dB. Figure 3 shows the simulation of SCM symbols, considering $K = 2$ layers the number of symbols in SCM is 2. Figure 4 show that the PAPR reduction in SCM-OFDM system with DHT precoder and clipping as a reduction PAPR technique. DHT precoder outperforms clipping technique by 7.01 dB. BER performance is better compared to clipping technique for SCM-OFDM system. Figure 5 present that when modified μ-law compander with different value of μ, with $\mu = 30$ and above, the PAPR of SCM-OFDM system reduces to 3.8 dB. Increment in PR ratio to 2, increases the BER performance, compared to other μ-law companding technique as shown in Figure 6. The BER performance in DHT precoder is better than the BER of μ-law companding techniques. The simulation uses different μ-law companding levels and is denoted as U_1 and U_2 in the result. However, the PAPR is more in DHT precoding technique compared to μ-law companding. The proposed system considers the advantages of both the techniques. Figure 7 show different PAPR reduction technique for SCM-OFDM system, the proposed system outperforms the other reduction technique with PAPR of 1.79 dB. The BER performance of proposed system and its out performance with the other PAPR reduction techniques is shown in Figure 7. Table 2 shows the comparison of computational complexity in DHT and DFT scheme for PAPR reduction. Since DHT requires less multiplications compared to DFT, proposed system is low computationally complex system. Table 3 compares the spectral efficiency of the DHT with other alternative techniques. PTS and SLM techniques use side information which increases bandwidth requirement. The proposed method is bandwidth efficient, since no side information is required. Figure 8 provides BER comparative study of SCM-OFDM system with different PAPR reduction techniques. PAPR reduction for SCM-OFDM system with proposed reduction technique is 12.5 dB for 256 QAM modulations. Table 4 provides the PAPR results and comparisons with SCM only, OFDM only and hybrid combination of SCM-OFDM system with clipping, SLM, PTS and MMC-DHT precoder for 64 QAM and 256 QAM. It is evident that from Table 4 with MMC-DHT precoder improves PAPR performance of SCM-OFDM system.

Table 1 Simulation parameters

Bandwidth	20 MHz
Carrier frequency	2.5 GHz
Number of subcarriers	256
Subcarrier spacing	15 KHz
Number of cyclic prefix	64
Convolution encoder rate	1/2
Interleaver	Random interleaver
Modulation for layer 1	256 QAM
Modulation for layer 2	256 QAM

Table 2 Comparison of DHT and DFT for computational complexity

Schemes	*Real multiplications*	*Complex additions*
DHT	$2L^2$	$L(L-1)$
DFT	$4L^2$	$L(L-1)$

Notes: Where L is order of DHT matrix and DFT matrix.

Table 3 Comparison of DHT for spectral efficiency

Scheme	*Sideband information (bits/OFDM symbol)*
DHT	None
PTS	S $\log_2 \varnothing$
SLM	$\log_2 \varnothing$

Notes: Where S is the number of subblocks and ∅: is the number of phase factors.

Figure 4 BER performance of clipping and DHT PAPR reduction techniques in SCM-OFDM system (see online version for colours)

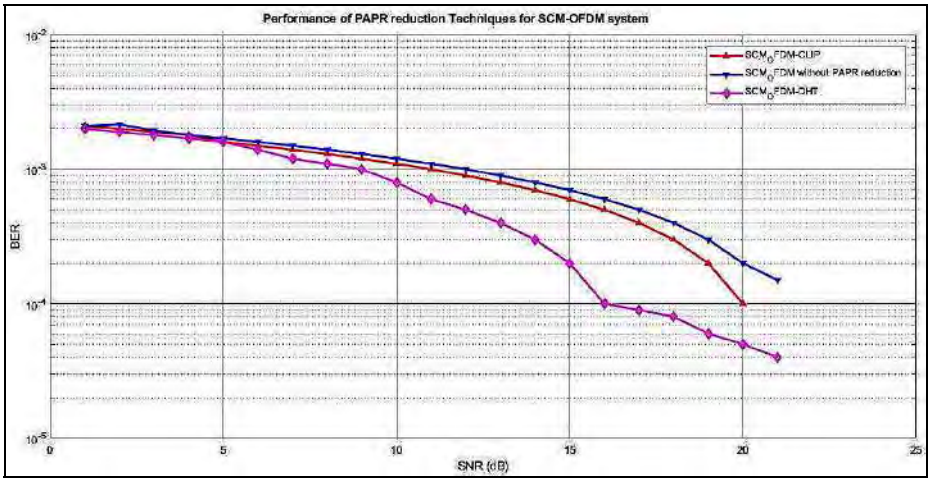


Figure 5 CCDF vs. PAPR for different μ companding profiles (u_1, u_2) and modified μ companding techniques (see online version for colours)

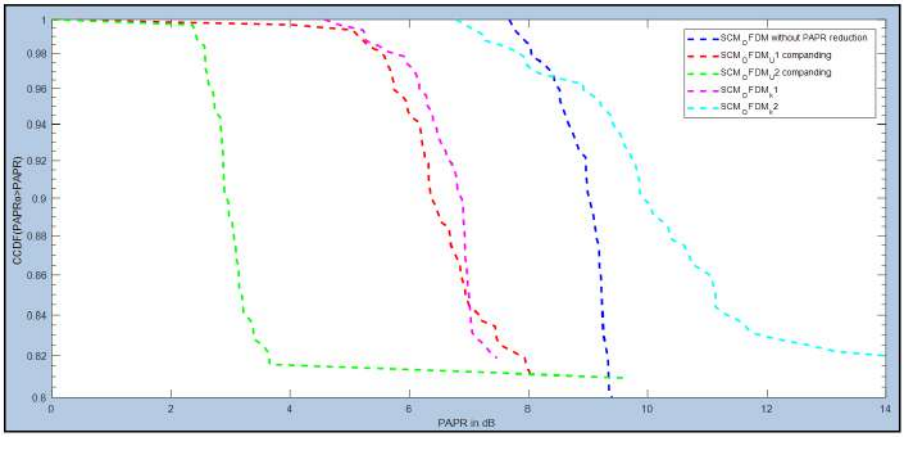


Table 4 Comparison of PAPR reduction techniques for SCM-OFDM system

PAPR reduction technique	*SCM*	*OFDM*	*SCM-OFDM without PAPR reduction*	*SCM-OFDM with clipping*
PAPR (in dB)	6.4089	9.31	13.53	10.6536
PAPR reduction technique	*SCM-OFDM with DHT*	*SCM-OFDM with μ companding*	*SCM-OFDM with proposed PAPR reduction for 64 QAM*	*SCM-OFDM with proposed PAPR reduction for 256 QAM*
PAPR (in dB)	5.8971	3.64	1.7949	12.5

Figure 6 BER performance of modified μ-law companding technique on SCM-OFDM system with different companding levels U_1, U_2 and U_3 (see online version for colours)

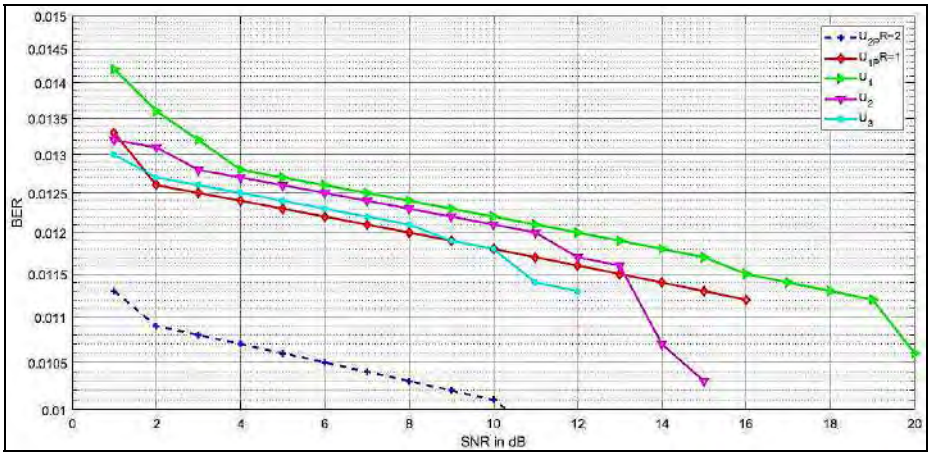


Figure 7 CCDF vs. PAPR for different PAPR reduction techniques on SCM-OFDM system (see online version for colours)

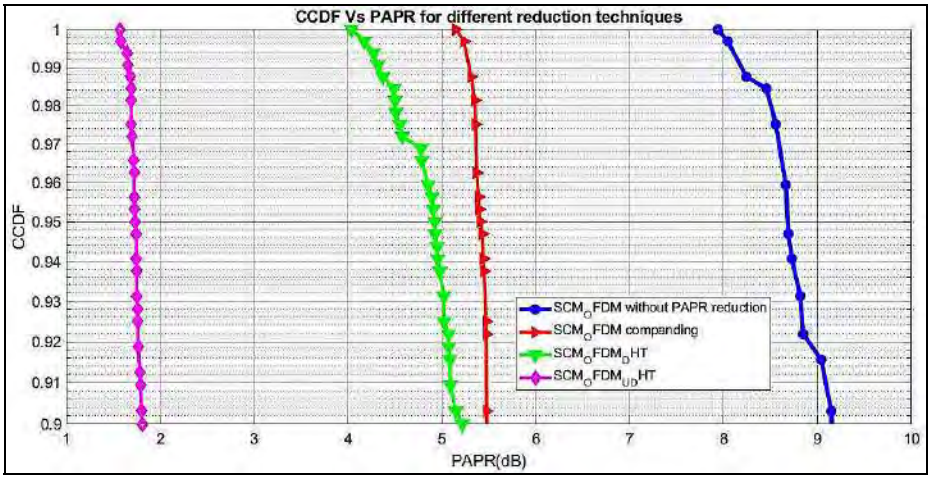


Figure 8 BER performance of different PAPR reduction technique on SCM-OFDM system (see online version for colours)

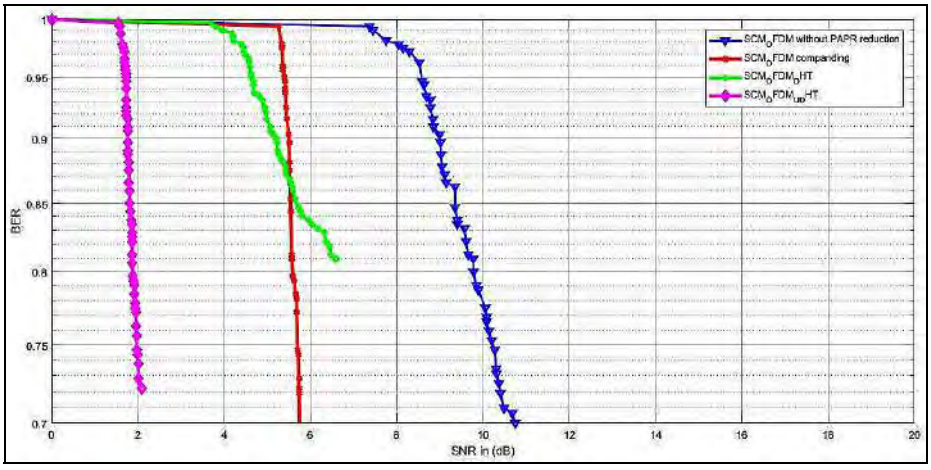
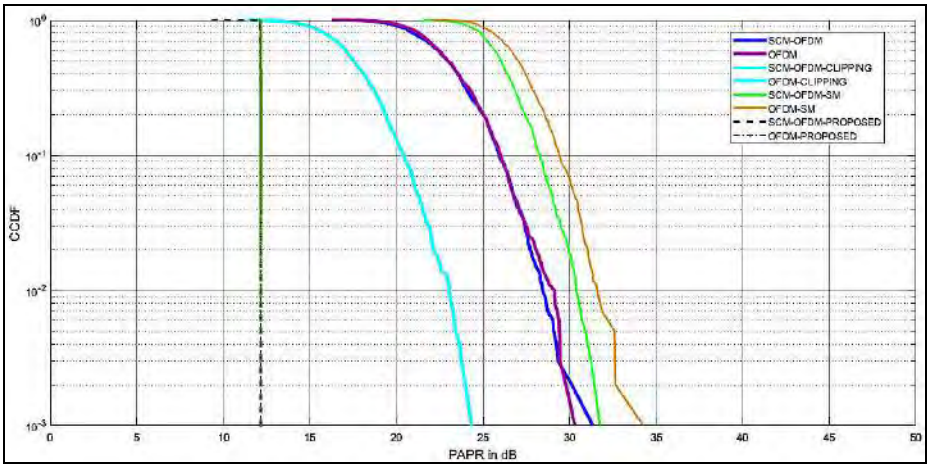


Figure 9 Comparative study of PAPR on SCM-OFDM systems for 256 QAM modulation with different reduction techniques (see online version for colours)



The SCM OFDM system is simulated for 64 QAM and 256 QAM respectively, PAPR for SCM signal and OFDM signal PAPR are 6.4089 dB and 9.31 dB respectively. With the combination of SCM-OFDM system PAPR increased to 13.53 dB due to Gaussian like signal of SCM-OFDM signal and with equal power allotted for two layers of SCM. Clipping technique reduces PAPR by 2.87 dB, where as DHT achieves PAPR reduction by 7.63 dB. PAPR is further reduced by 9.89 dB by μ companding technique. As observed from the simulation results DHT technique provides better BER performance, μ companding technique provides better PAPR performance. Combining the advantages of both the technique we achieved reduction in PAPR by 11.73 dB and improvement in PAPR by 86.7%. SCM-OFDM system is simulated with 256 QAM to demonstrate the proposed method works better compared to well known PAPR technique such as PTS and SM. Figure 9 shows that PAPR in SCM-OFDM without PAPR reduction is 33 dB,

with SM technique PAPR is reduced to 32 dB. Some improvement is observed when PTS is applied to SCM-OFDM system. Significant result is obtained with the proposed technique, PAPR is reduced to 12.5 dB, PAPR performance is improved by 62.1%.

6 Conclusions

A novel PAPR reduction technique for the OFDM system with SCM is presented and elaborated in this paper. An investigation on performance of different reduction techniques on SCM-OFDM system is achieved through computer simulation. The results indicated that the proposed technique outperformed the clipping technique by 83.1%, PTS by 58.3%, SM by 62% and achieved PAPR of 1.7949 dB with 64 QAM SCM-OFDM system, 12.5 dB with 256 QAM SCM-OFDM system. The proposed method is beneficial for high data rate systems, where many constellations are required to achieve a high data rate.

Future scope: For future research, time-varying channels can be studied; furthermore, different modulations and layers can be experimented with SCM-OFDM system.

References

Azim, A.W., Le Guennec, Y. and Maury, G. (2017) 'Decision-directed iterative methods for PAPR reduction in optical wireless OFDM systems', *Optics Communications*, Vol. 389, pp.318–330.

Chen, H. and Liang, H. (2007) 'PAPR reduction of OFDM signals using partial transmit sequences and Reed-Muller codes', *IEEE Communications Letters*, Vol. 11, No. 6, pp.528–530.

El-Hamed, M.A., Zekry, A., Elagooz, S.S. and El-Samie, F.E.A. (2018) 'Blind selective mapping for single-carrier frequency division multiple access system', *Digital Signal Processing*, Vol. 75, pp.25–37.

Ghahremani, R. and Shayesteh, M.G. (2014) 'BER performance improvement and PAPR reduction in OFDM systems based on combined DHT and μ-law companding', in *2014 22nd Iranian Conference on Electrical Engineering (ICEE)*, IEEE, May, pp.1483–1487.

Hassan, E. (2016) 'PAPR reduction using selective mapping scheme', in *Multi-Carrier Communication Systems with Examples in MATLAB®*, pp.82–109, CRC Press.

Lee, J.H. and Chung, H. (2018) 'Exact and approximate log-likelihood ratio of M-ary QAM with two-time dimensions', *ICT Express*, Vol. 5, No. 3, pp.173–177.

Liang, H.Y. (2015) 'Integrating CE and modified SLM to reduce the PAPR of OFDM systems', *Wireless Personal Communications*, Vol. 80, No. 2, pp.709–722.

Liang, H.Y., Chu, H.C., Lin, C.B. and Lin, K.H. (2014) 'A partial transmit sequence technique with error correction capability and low computation', *International Journal of Communication Systems*, Vol. 27, No. 12, pp.4014–4027.

Nadal, L., Moreolo, M.S., Fàbrega, J.M. and Junyent, G. (2014) 'Low complexity PAPR reduction techniques for clipping and quantization noise mitigation in direct-detection O-OFDM systems', *Optical Fiber Technology*, Vol. 20, No. 3, pp.208–216.

Nagaraju, C.H., Sharma, A.K. and Subramanyam, M.V. (2018) 'Reduction of PAPR in MIMO-OFDM using adaptive SLM and PTS technique', *International Journal of Pure and Applied Mathematics*, Vol. 118, No. 17, pp.355–373.

Rashmi, N. and Sarvagya, M. (2016) 'State space model based channel estimation using extended Kalman filter for superposition coded modulation OFDM system', *BVICA M's International Journal of Information Technology*, Vol. 8, No. 2, p.1017.

Satyavathi, K. and Rao, B.R. (2019) 'Modified phase sequence in hybrid PTS scheme for PAPR reduction in OFDM systems', in *Innovations in Electronics and Communication Engineering*, pp.327–333, Springer, Singapore.

Vittal, M.V.R. and Naidu, K.R. (2017) 'A novel reduced complexity optimized PTS technique for PAPR reduction in wireless OFDM systems', *Egyptian Informatics Journal*, Vol. 18, No. 2, pp.123–131.

Wang, H., Wang, X., Xu, L. and Du, W. (2016) 'Hybrid PAPR reduction scheme for FBMC/OQAM systems based on multi data block PTS and TR methods', *IEEE Access*, Vol. 4, pp.4761–4768.

Detection of Breast Cancer with Lightweight Deep Neural Networks for Histology Image Classification

H.S. Laxmisagar* & M.C. Hanumantharaju

Department of Electronics and Communication Engineering, BMS Institute of Technology Management, Bengaluru 560064, India

*Address all correspondence to: H.S. Laxmisagar, Department of Electronics and Communication Engineering, BMS Institute of Technology Management, Bengaluru 560064, India; Tel.: +08123008252, E-mail: sagar8.hs@bmsit.in

ABSTRACT: Many researchers have developed computer-assisted diagnostic (CAD) methods to diagnose breast cancer using histopathology microscopic images. These techniques help to improve the accuracy of biopsy diagnosis with hematoxylin and eosin-stained images. On the other hand, most CAD systems usually rely on inefficient and time-consuming manual feature extraction methods. Using a deep learning (DL) model with convolutional layers, we present a method to extract the most useful pictorial information for breast cancer classification. Breast biopsy images stained with hematoxylin and eosin can be categorized into four groups namely benign lesions, normal tissue, carcinoma *in situ*, and invasive carcinoma. To correctly classify different types of breast cancer, it is important to classify histopathological images accurately. The MobileNet architecture model is used to obtain high accuracy with less resource utilization. The proposed model is fast, inexpensive, and safe due to which it is suitable for the detection of breast cancer at an early stage. This lightweight deep neural network can be accelerated using field-programmable gate arrays for the detection of breast cancer. DL has been implemented to successfully classify breast cancer. The model uses categorical cross-entropy to learn to give the correct class a high probability and other classes a low probability. It is used in the classification stage of the convolutional neural network (CNN) after the clustering stage, thereby improving the performance of the proposed system. To measure training and validation accuracy, the model was trained on Google Colab for 280 epochs with a powerful GPU with 2496 CUDA cores, 12 GB GDDR5 VRAM, and 12.6 GB RAM. Our results demonstrate that deep CNN with a chi-square test has improved the accuracy of histopathological image classification of breast cancer by greater than 11% compared with other state-of-the-art methods.

I. INTRODUCTION

According to the American Cancer Society, cancer is the second most common cause of death for women. In 2018, an estimated 627,000 women died from cancer, accounting for roughly 15% of total cancer deaths in women.[1] The most essential techniques for preventing breast cancer fatalities are early identification[2] and receiving cutting-edge cancer therapy. X-ray radiography, three-dimensional ultrasonography, positron emission tomography, magnetic resonance imaging, tomography, and breast temperature monitoring are all approaches for screening for breast cancer.[3–7] Pathological diagnosis, on the other hand, is frequently viewed as a "golden standard." The extracted tissues are frequently stained for better observation and analysis, including hematoxylin and eosin (H&E) staining is the most common procedure. Hematoxylin paints the nuclei a dark purple color, while eosin dyes the cytoplasm, stroma, and other structures a pink color. However, even experienced pathologists find it difficult to analyze such images, resulting in inter and intra observer differences. Additionally, such analysis is time and resource consuming.[8] The different kinds of breast tissue scan stained with H&E[9] are shown in Fig. 1.

Normal, benign, carcinoma *in situ* (CIS), and invasive ductal carcinoma (IDC) are the four forms of breast tissue.[10] Benign tissue is a slight variation in the structure of the breast that is not cancerous and, in most situations, does not pose a health risk. CIS cancer is limited to the mammary duct lobule system and does not affect the rest of the body. *In situ* carcinoma can be treated if detected early enough. Invasive carcinoma, on the other hand, is a cancerous tumor that spreads to other organs. The most common[11] method for detecting many cancer kinds is histopathological imaging. A histopathologist examines the sections under an optical microscope to investigate the tissues' features and properties in histopathological research.[12] Traditionally, the histopathologist observes tissue sections with his or

0278-940X/22/$35.00

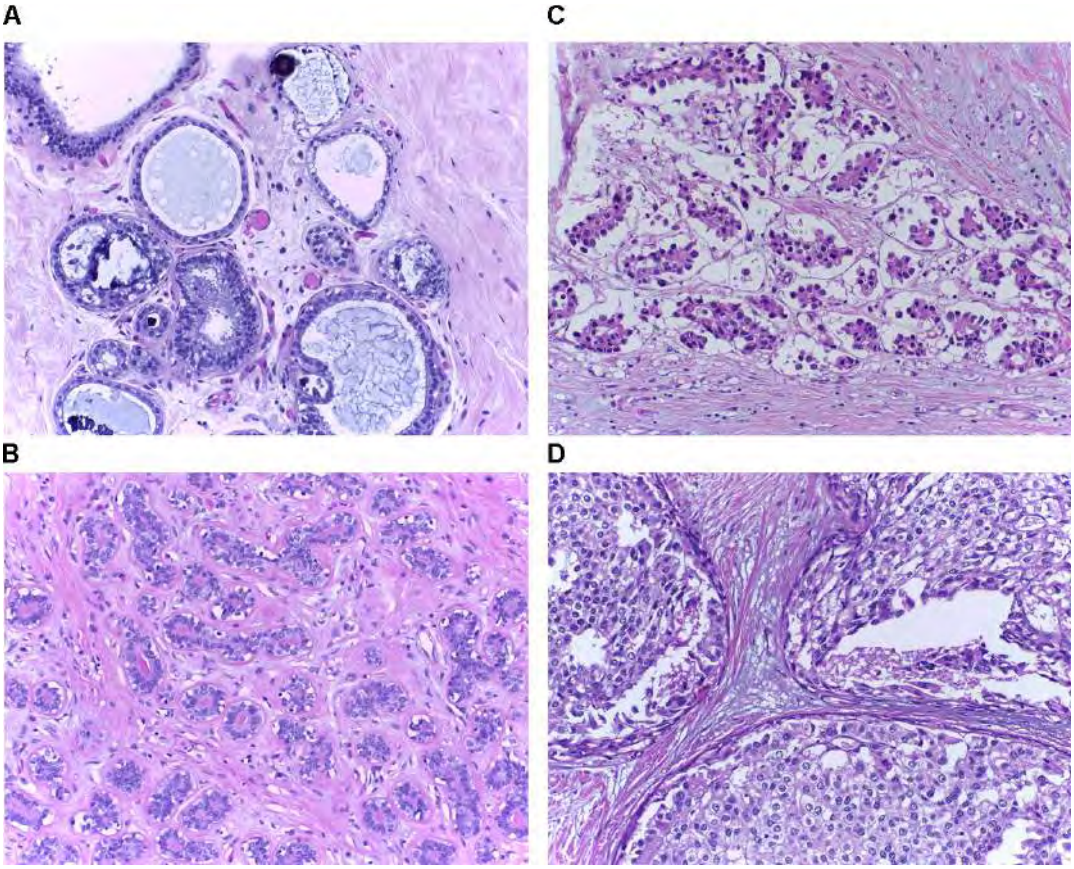


FIG. 1: Sample histopathology images obtained from ICIAR 2018 grand challenge (BACH). (A) Normal. (B) Benign. (C) Invasive. (D) *In situ*.

her naked eyes, and the visual data is manually examined based on existing medical knowledge. This manual analysis can take a long time due to the intricacy and diversity of histopathological pictures. The accuracy rate of diagnostician's manual classification of cancer kinds appears to be over 75%.[13] The CAD has been developed to distinguish and detect cancer kinds to tackle this challenge. Furthermore, CAD relies on machine learning (ML) algorithms to categorize and identify breast cancer, which is based on histopathological image discriminatory features.[14]

Artificial intelligence (AI) technology is advancing fast in recent years. It has made significant advances in computer vision, image processing, and analysis. In addition, AI has the potential to improve histopathology analysis. Fortunately, advances in computer vision and ML may provide more reliable classification approaches for hematoxylin and eosin-stained sections histological assessment. These approaches have a good classification rate for automatically classifying breast tissues into distinct categories. As a result, numerous researchers have created image examination procedures for breast cancer screening that are quick and exact. Their findings, however, fall far short of what is expected in clinical practice. As a result, the majority of investigators' efforts have gone into developing novel algorithms for histopathological image processing.[15–17] Pathologists' errors can be decreased to some extent, and their productivity can be increased. Deep learning (DL) with convolutional neural network (CNN) has appeared as one of the most powerful ML technologies in image classification in recent years.[18] According to new research, CNN-based algorithms outperform the best classical

ML approaches in terms of performance. In the field of histopathology, DL algorithms have been developed to improve diagnosis accuracy which is more accurate compared with trained pathologists for detection of breast cancer.[19]

Section II develops a comparative table of performances of neural network algorithms on image datasets such as BreakHis, ICIAR BACH, and IMEDIATREAT. Based on the inferences from the literature survey, section III proposes a model to improve the classification performance parameter. The architectural framework and other specifications of the model proposed are discussed in section IV. In section V, the summary of various CAD system development with percentage accuracy is presented in more detail. Finally, a summary and future implementation are presented in section VI.

II. RELATED WORK

The state-of-the-art algorithms for breast cancer identification and related disciplines, such as feature extraction utilizing CAD systems and DL for breast cancer detection are briefly reviewed and discussed in this part. In many hospitals, CAD has quickly grown in recent years to support clinicians in detecting cancer patients faster and more accurately. CAD is utilized to produce objective results and has aided in the diagnosis of medical images in general. The ability to discern between benign, normal, CIS, and malignant carcinoma is one of the most important CAD applications. ResNet50 and DenseNet121 deep CNN approaches for CAD tool development for breast cancer diagnosis were proposed by Yasin et al.[20] The total number of parameters in ResNet50 and DenseNet121 is insufficient to adequately cover the fine detail of 400 images. Only 82 patients were used to create the BreakHis dataset, resulting in a high level of data variability. Erfankhah et al.[22] proposed a local binary pattern (LBP) model with SVM classifier. IDC dataset's main limitations may be related to its small patch size (50 × 50) and binary classification.

Gour et al.[24] proposed a 152-layered CNN based on residual learning. In spite of the high performance of ResHist in detecting breast cancer, it must be validated on a larger dataset before it can be used in healthcare. Hu et al.[25] proposed the myResNet-34, a residual learning-based convolutional neural network. It is required to analyse the myResNet-34 with pre-training on ImageNet and investigate the outcomes of the proposed framework on various datasets, based on the effectiveness of transfer learning reported in recent papers. Yang et al.[29] suggested a multiscale network ensemble (EMS-Net) which is used to find and use the most discriminative areas in each microscope image. Table 1 describes the summary of different techniques proposed by various authors for the detection of breast cancer using BreakHis and IMEDIATREAT dataset.

Table 2 describes the summary of different techniques proposed by various authors for the identification of breast cancer using the BACH dataset. The uniqueness of this study consists in the development of a cost-effective CAD system based on feature extraction and classification utilizing DL approaches to differentiate between normal and abnormal breast cancer lesions.

The BACH challenge dataset was used to examine this CAD system. Multiple articles in the literature have used DCNNs to categorize breast cancer as part of their CAD systems,[20–30] but the accuracy of 83–96% was insufficient for a dependable and powerful CAD system. Furthermore, they did not look at ways to lower the CAD system's computing costs. To address these flaws, this research proposes a revolutionary CAD method for developing low-cost deep neural networks, which enhances the CAD system's accuracy. In medicine, images are vital[31] to improve the processes to treat patients in addition to helping doctors perform better diagnostics.

As per our knowledge, no paper is found related to the implementation of an efficient lightweight DCNN algorithm for the detection of breast cancer, and most existing development is related to other DCNNs such as ResNet, AlexNet, VggNet, and EMS-Net as discussed in the literature survey of state-of-the-art methods.

III. RESEARCH METHODOLOGY

The outline of the proposed work is shown in Fig. 2. In the data preprocessing steps such as data

TABLE 1: Summary of different techniques proposed by various authors for the detection of breast cancer using the BreakHis[34] and IMEDIATREAT[35] datasets

Author	Publisher, year	Dataset	Accuracy rate	Limitations
Yari et al.[20]	IEEE Transaction, 2020	Histopathology Image: BreakHis Dataset Method: DCNN: ResNet50 and DenseNet121	CAD tool development for breast cancer detection with 98 (100×) % accuracy	Another future scope could be to investigate the lower precision in the 400× image. The investigation of lower accuracy in the 400× image can be another future work.
Mansour et al.[21]	World Scientific Journal, 2020	Histopathology Image: BreakHis Dataset Method: CNN: AlexNet-DNN	CAD tool development for breast cancer detection with 96.20%	The proposed method can be examined with additional cutting-edge GPUs such as Nvidia and others. To achieve a more time-efficient method that produces the highest quality BC-CAD solution.
Erfankhah et al.[22]	IEEE Journal, 2019	Histopathology Image: BreakHis Dataset 400× Method: Local binary pattern (LBP) model with SVM classifier	CAD tool development for breast cancer detection with 83.40%	The fundamental restriction of this dataset is the nature of its classes. Two test portions of the same tissue type (for example, fat tissue) may come from two distinct scans.
Lichtblau, et al.[23]	IEEE, 2019	Histopathology Image: IMEDIATREAT data set Method: AlexNet model with various Classifier	CAD tool development for breast cancer detection with 87%	FTT+PCA makes a mistake by relying on the cumulative weight of the correct classifiers and assigning the proper FTT+PCA label where other classifiers may have performed better.
Gour et al.[24]	Wiley Publication, 2020	Histopathology Image: BreakHis Dataset Method: Residual learning-based CNN	CAD tool development for breast cancer detection with 84.34% accuracy	The suggested approach has numerous advantages over hand-crafted local feature descriptors; nonetheless, training the ResHist model requires enormous computational power and time when compared with feature descriptors.
Hu et al.[25]	Wiley Publication, 2021	BreakHis dataset: Method: Convolutional neural network named myResNet-34	CAD tool development for breast cancer detection with 94.03% accuracy	The classification accuracy may be improved by employing other DL-based classifier models and better stain normalization approaches.

augmentation and normalization techniques were used to increase the number of samples and model accuracy by removing image staining. The term "data augmentation" denotes the process of adding comparable data to an existing dataset to strengthen the model. Normalization is a pre-processing procedure for preparing biopsy imaging slides that employ a staining approach to eliminate staining**.** Finally, we used different MobileNet models to find out model accuracy.

TABLE 2: Summary of different techniques proposed by various authors for the identification of breast cancer using the BACH dataset

Author	Publisher, year	Dataset	Accuracy rate	Limitations
Aresta et al.[26]	Elsevier, 2019	Histopathology Image: ICIAR, 2018, BACH challenge dataset Method: Convolutional neuronal networks	CAD tool development for breast cancer detection with 87% accuracy	The BACH dataset remains freely accessible to support additional developments in digital pathology for CAD system development. The accuracy obtained is less due to the CNN model trained without a transfer learning approach.
Kassani et al.[27]	IEEE ICTC, 2019	Histopathology Image: ICIAR, 2018, BACH challenge dataset Method: Deep Convolutional Neural Network (DCNN)	CAD tool development for breast cancer detection with 92.50% accuracy	The classification accuracy may be improved by employing a DL-based classifier model and better stain normalization approaches.
Kausar et al.[28]	Elsevier, 2019	Histopathology Image: ICIAR, 2018 and BreakHis dataset Method: Deep convolution neural network (CNN)	ICIAR: CAD tool development for breast cancer detection with 91% accuracy BreakHis: 96.85% accuracy	The insufficient amount of accessible standard datasets lowers the refinement capacity of the deep CNN model.
Yang et al.[29]	Elsevier, 2019	Histopathology Image: ICIAR, 2018, BACH challenge dataset Method: Ensemble of MultiScale Networks (EMS-Net)	CAD tool development for detection of breast cancer with 90% accuracy	The network is time-consuming during training, it may be done offline. Introduce the new approach to locate and use the furthermost exclusionary areas in each microscope image.
Kumar et al.[30]	Cancer Management and Research, 2020	Histopathology Whole slide Images: ICIAR, 2018, BACH challenge dataset Method: k-means clustering algorithm	Normal Class: 85.5% accuracy	The outcome can be enhanced, if there is a significant number and variety of labeled whole slide image sample datasets.

FIG. 2: Overview of the proposed model

A. Dataset Description

A collection of 400 high-resolution histology images[33] is included in this archive. The resolution of each image ranges from 2040 × 1536 pixels. This image is made of histology microscopic images stained with H&E that were used in the bio-imaging breast histology categorization challenge 2018. There are four categories of image classes: benign, normal, *in situ*, and invasive. A diagnosis was produced from the contents of the image classes by two experienced pathologists. The bio-imaging challenge's purpose is to conduct data categorization amongst distinct classes. The collection contains 400 microscopic pictures, with 100 images each for benign, normal, invasive carcinoma, and carcinoma *in situ* conditions. The tiny pictures are 2048 × 1536 pixels and have an RGB color model with a .tiff format.

B. Data Pre-Processing

We pre-process the dataset which is collected in two steps such as normalization and data augmentation. Normalization is a pre-processing procedure for preparing biopsy imaging slides that employ a staining approach to eliminate staining as shown in Fig. 3. A logarithmic[13,14] transformation is used to convert the high-resolution pictures to an optical density. The optical density tuples are then subjected to the singular value decomposition, and the result in color space transform is utilized to produce the odd biopsy picture. The histogram picture[15] is then expanded to encompass the lower 90% of the dataset. The picture samples before and after normalizing are shown in the diagram below. Computers don't need to be stained; only pathologists need to dye histology pictures to distinguish between different regions of interest.

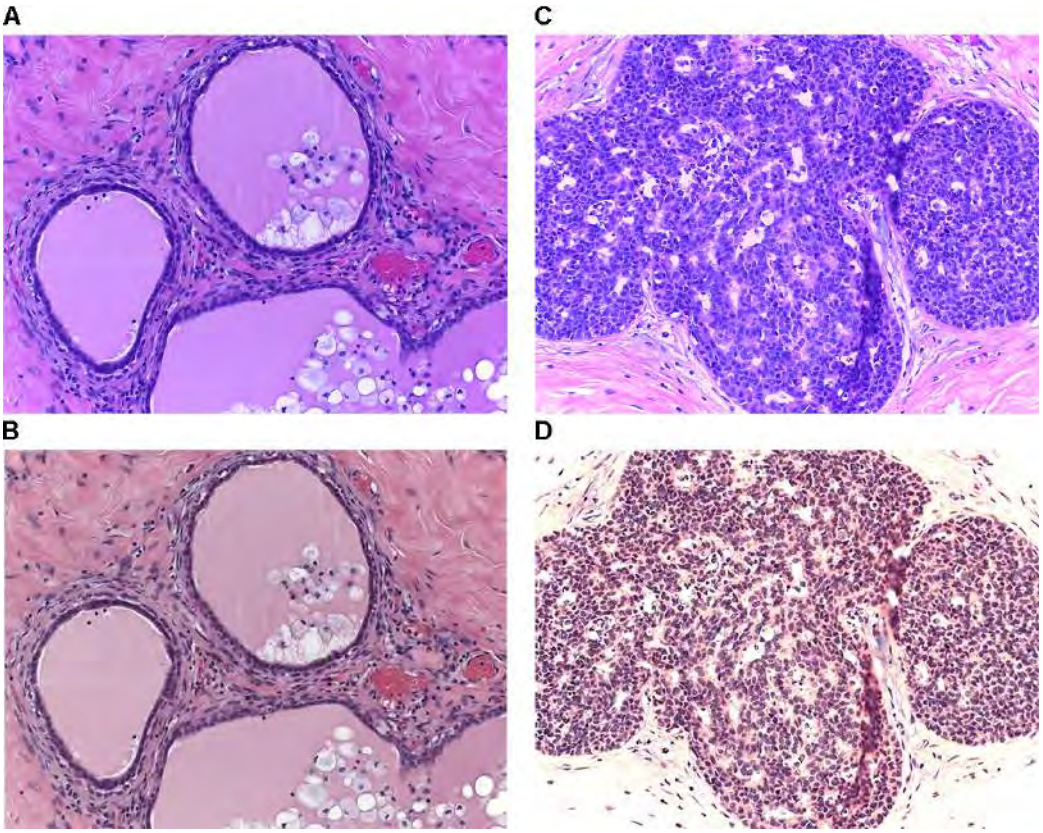


FIG. 3: Normalization of images (before vs. after): (A,C) before normalization of images; (B,D) after normalization of images

The term "data augmentation" refers to the process of adding comparable data to an existing dataset to strengthen the model. In addition, the normalized pictures are used to build a dataset that is used to rotate the image in the process. A data augmentation strategy based on picture rotation and flipping is used to enhance[15] the dataset without impacting diagnostic or image quality to avoid underfitting owing to a small number of samples. The data is supplied into the CNN model once it has been normalized and enhanced. Once the model has learned the underlying attributes of these four classes, it will be able to predict images without labels. Image Data Generator is a class in Keras that allows handlers to perform image augmentation in a very basic method.

For our deep neural network to train the model, the dataset, whatever it is utilized for, is extremely little. To avoid this, researchers employed several data augmentations on pre-processed histology pictures, such as rotating the x- and y-axis, flipping and cropping the image, to construct a robust model and avoid overfitting issues. As illustrated in Fig. 4, data augmentation of the original image with respect to the axis has been performed.

To perform high-precision diagnostics, convolutional neural networks require a large dataset size. One of the challenges in CNN is insufficient medical

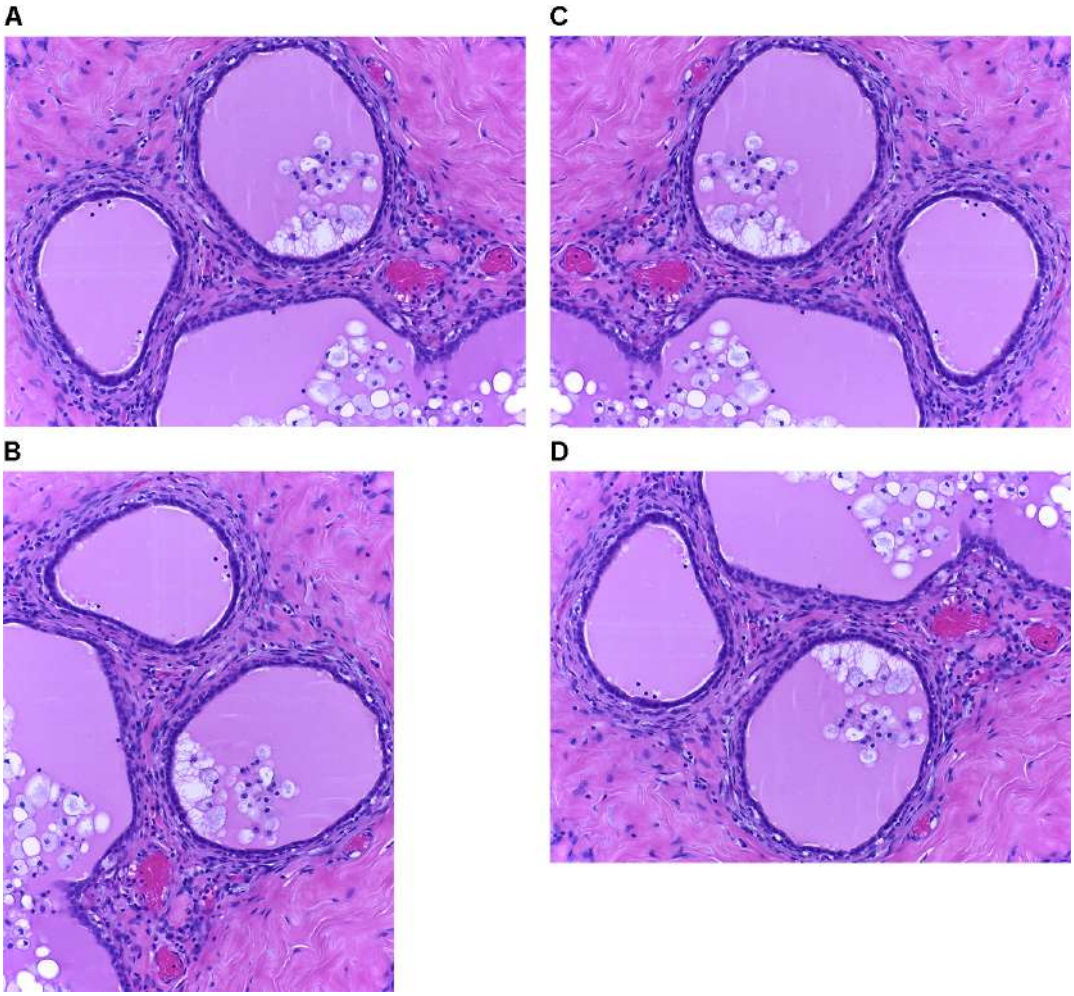


FIG. 4: Image augmentation (rotation about the axis). (A) Original image. (B) Data augmentation 1. (C) Data augmentation 2. (D) Data augmentation 3.

data set size, and the over-fitting problem which means CNN performed in terms of accuracy is high for the training dataset, but it performs less with the test dataset. Therefore, in the proposed system, the present study has used the data augmentation technique to solve the problems of data set size and over-fitting. In the data augmentation technique, the number of each microscopic image was increased by geometric transformations such as translation, scaling, rotation, and flipping.

IV. MODEL ARCHITECTURE

Convolutional Neural Networks are a family of deep, feed-forward artificial neural networks, most typically employed to evaluate visual data. CNNs employ a form of multilayer perceptron's designed to need little preprocessing. In a convolutional neural network, each neuron is only linked with a few local neurons in the previous layer, and the weight is distributed for every neuron in that layer. Convolutional neural networks are useful for image classification tasks because the convolution procedure generates information on spatially associated elements of the image.

With MobileNet, you are primarily going to use depth-wise discrete convolutions, which multiply a regular convolution by a depth-wise convolution as well as a 1 × 1 convolution can be called a pointwise convolution.[8] As shown in Fig. 5, the convolutional neural network technique MobileNet 2.10ex is employed to implement the operation, which includes two layers, one of which performs depth-wise convolutions and another of which performs pointwise convolutions. The ReLu activation function is employed at the output of each neuron to prevent neuron loss. We can use depth-wise convolutional layers and a convolutional layer of 1 × 1 for combining.[19] After the first phase is completed, the next phase involves the extraction of a feature matrix that contains labeled data. Four hidden layers of the deep neural network analyze the feature data and classify the data into four different categories. Models with low latency such as MobileNet are ideal for systems and applications that require models to be small. The underlying architecture of MobileNet is tiny and has low latency, making it ideal for the needs of embedded vision and mobile applications. In order to improve the classification performance on a small dataset, it is required to incorporate a trained weight transfer model using ImageNet database as shown in Fig. 6.

The proposed convolutional neural network, MobileNet model, was used to classify four different classes of data for the detection of breast cancer. This method has included four steps such

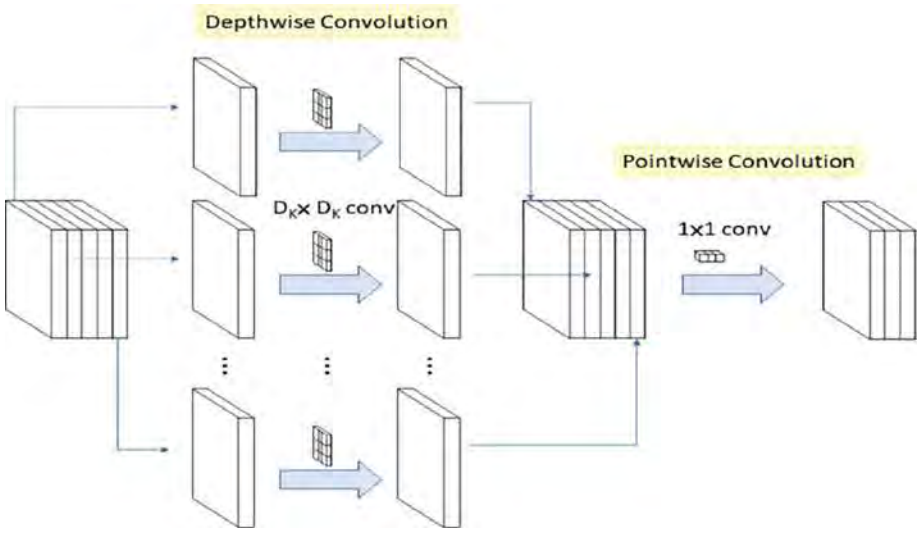


FIG. 5: MobileNet architecture[8]

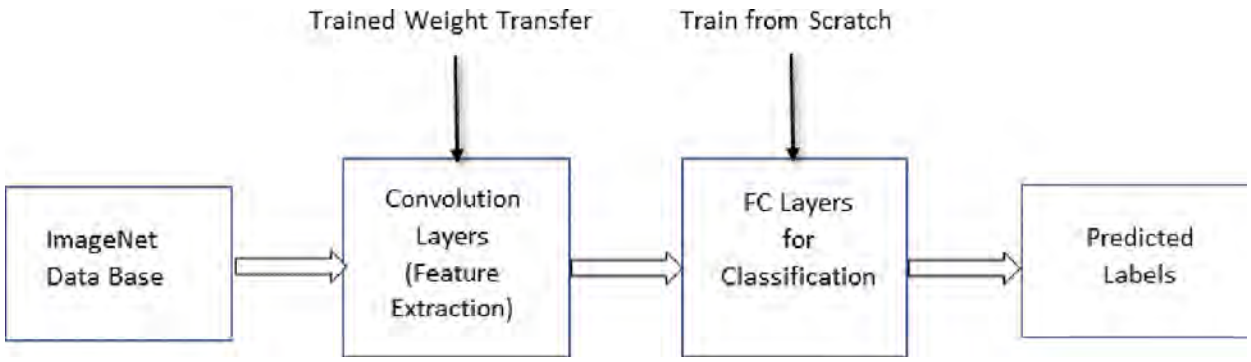


FIG. 6: Trained weight transfer model using the ImageNet database

as data normalization, data augmentation, feature extraction, and classification as shown in Fig. 7. First, pre-treatment is necessary to remove noise from histology images, and normalization is done as a pre-processing procedure to eliminate staining from microscopic images. This step improves diagnostic performance. Second, the data augmentation technique is used to increase the amount of training data effectively. A data augmentation strategy based on picture rotation and flipping is used to enhance the dataset without impacting diagnostic or image quality to avoid underfitting owing to a small number of samples. Third, the augmented dataset is used to train the network on images of various multi-scales,

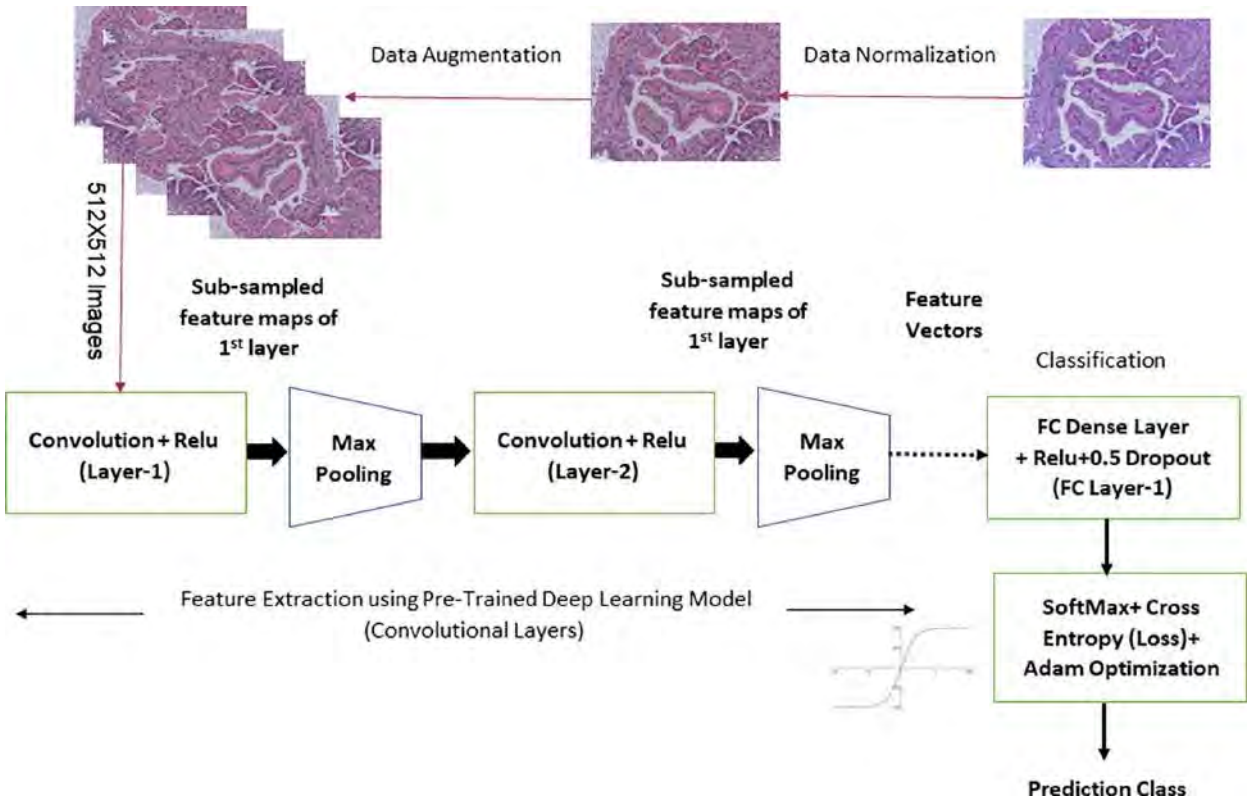


FIG. 7: Framework for histology image classification

and then the deep features of the breast histology are extracted. In the feature extraction downsampling is performed using max-pooling at various stages to get feature vectors. Once the model has learned the underlying attributes of these four classes, it will be able to predict images without labels. In the proposed system, the MobileNet model was used to extract deep features. Fourth, the deep features are fed into the fully connected layer, which classifies the microscopic images into four different classes.

All layers of MobileNet follow a batch-normalization layer and include rectified linear unit (ReLu) nonlinearity except for the last fully connected layer, on which there is no nonlinearity. Additionally, there is the output dense layer for prediction with four different classes of data. The equation of rectified linear unit is given below,

$$R(Z) = \max(0, Z) \tag{1}$$

where R(Z) is the activation function and Z denotes the input to a neuron.

This function value is set to 0 if it receives a negative input, and if it receives a positive value, the function will retain the same positive value. The softmax function is very essential for multiple classes which are more than two.

It converts real numbers to normalized probability distribution by taking the exponents of each output. Once that has been done, regularize each number by the sum of each exponent so that the entire output vector adds up to one. At the last stage of an image classifier, a softmax function is used, such as those used in ImageNet competitions, such as CNN (VggNet).

$$\sum_{i=1}^{C} P_i = 1 \tag{2}$$

The sum of all probabilities at the output of softmax equals to 1 and the number of output probabilities varies from i = 1 to C, where C denotes total number of output probabilities of softmax function. Using the following equation, the softmax classification result may be predicted with more confidence as one of several outcomes,

$$\text{Confidence} = max_i P_i \tag{3}$$

where $max_i P_i$ is the maximum probability value at the output of softmax function.

The softmax function output is selected based on probability value which is higher compared with all others. The higher probability value is selected with more confidence to classify correct class of data in case of multiple classes to be classified.

$$P_i \geq 0,\ P_i \leq 1, \text{ where, } i = 1,2,3,\ldots\ldots,C$$

The output probability value of softmax function lies between the range 0 to 1 ($P_i \geq 0, P_i \leq 1$) as mentioned above. Softmax function converts all score values into normalized probability distribution,

$$S(y_i) = \frac{e^{y_i}}{\sum_j e^{y_j}} \tag{4}$$

where $S(y_i)$ is the softmax function for input real number y_i, $\sum_j e^{y_j}$ is the sum of all the exponents of input data y_i and it is a single number, e^{y_i} denotes exponent of input data y_i.

Categorical cross-entropy loss is a loss function used in multiclass classification problems. A probability density function model is designed to measure how different two probability densities differ when an example relates only to one of several possible categories. Its formal purpose is to identify the difference between two probability densities. The categorical cross-entropy is used by the model to learn to assign a high probability to the right class and a low probability to the other categories. The error is calculated using cross-entropy and the equation below

$$L = -\sum t_i \log y_i \quad (5)$$

where L denotes categorical cross-entropy loss, t_i is the truth label and y_i is the softmax probability for i*th* class. The aim is to minimize the loss, i.e., the smaller the loss the better the model. As an alternative to stochastic gradient descent, Adam can be used to iteratively update the weights of networks based on training data, instead of the traditional stochastic gradient descent process. The Adam technique is a stochastic gradient descent replacement method for training DL models. Table 3 shows pseudo-code for data normalization and augmentation, mobileNet model creation, training and validation of images, model evolution, and inferring the class of an image.

The notion of transfer learning is utilized throughout the model training process. Transfer learning in computer vision may be depicted using pre-trained models. Transfer learning models outperform typical ML models in terms of performance. This is because models that use information (features, weights, etc.) from previously trained models already comprehend the features. It is speedier than training neural networks from the ground truth images. Thus, it is best to use models such as VggNet, Inception, and MobileNet to reduce the need for multiplication when training such models.

In this approach, a transfer learning method is used for classifying breast cancer histology images. The network has learned the deep features of the MobileNet architecture, which is pre-trained on ImageNet.

The ICIAR BACH dataset has been classified into benign, normal, *in situ*, and invasive. ImageNet pre-trained models were used for transfer learning to get better accuracy. One of the main challenges of employing DL models in the field of medicine is a lack of training data due to difficulty in collecting and labeling data, which needs to be performed by experts. To overcome this drawback, transfer learning (TL) has been utilized to solve several medical imaging tasks using pre-trained state-of-the-art models from the ImageNet dataset as shown in Fig. 8.

Figure 9 depicts steps involved to perform histopathology image classification for detection of breast cancer. Breast cancer histopathology image datasets used in the present work have been obtained from https://iciar2018-challenge.grand-challenge.org/Dataset/. The resolution of each image ranges from 2040 × 1536 pixels. This image is made of histology microscopic images stained with H&E that were used in the bio-imaging breast histology categorization challenge 2018. There are four categories of image classes: benign, normal, *in situ*, and invasive. After collecting the dataset, it is required to pre-process the dataset in two steps, such as normalization and data augmentation, to remove staining from biopsy images and to increase the size of the dataset. To perform high-precision diagnostics, convolutional neural networks require a large dataset size. In the data augmentation technique, the number of each microscopic image was increased by geometric transformations such as translation, scaling, rotation, and flipping. Image Data Generator is a class in Keras that allows handlers to perform image augmentation in a very basic method. In the proposed system, the MobileNet model was used to extract deep features. In the feature extraction downsampling is done using max-pooling at various stages to get feature vectors. Finally, the deep features are fed into the fully connected layer, which classifies the microscopic images into four different classes. Four hidden layers of the deep

TABLE 3: Overview of the results obtained through the process of model validation

Models	Benign	*In Situ*	Invasive	Normal	Model accuracy
Model 2.00	10	10	4	12	75.0%
Model 2.10ex	10	11	7	10	79.2%
Model 1.40	10	11	10	11	87.5%
Model 1.40ex	11	11	7	6	72.9%

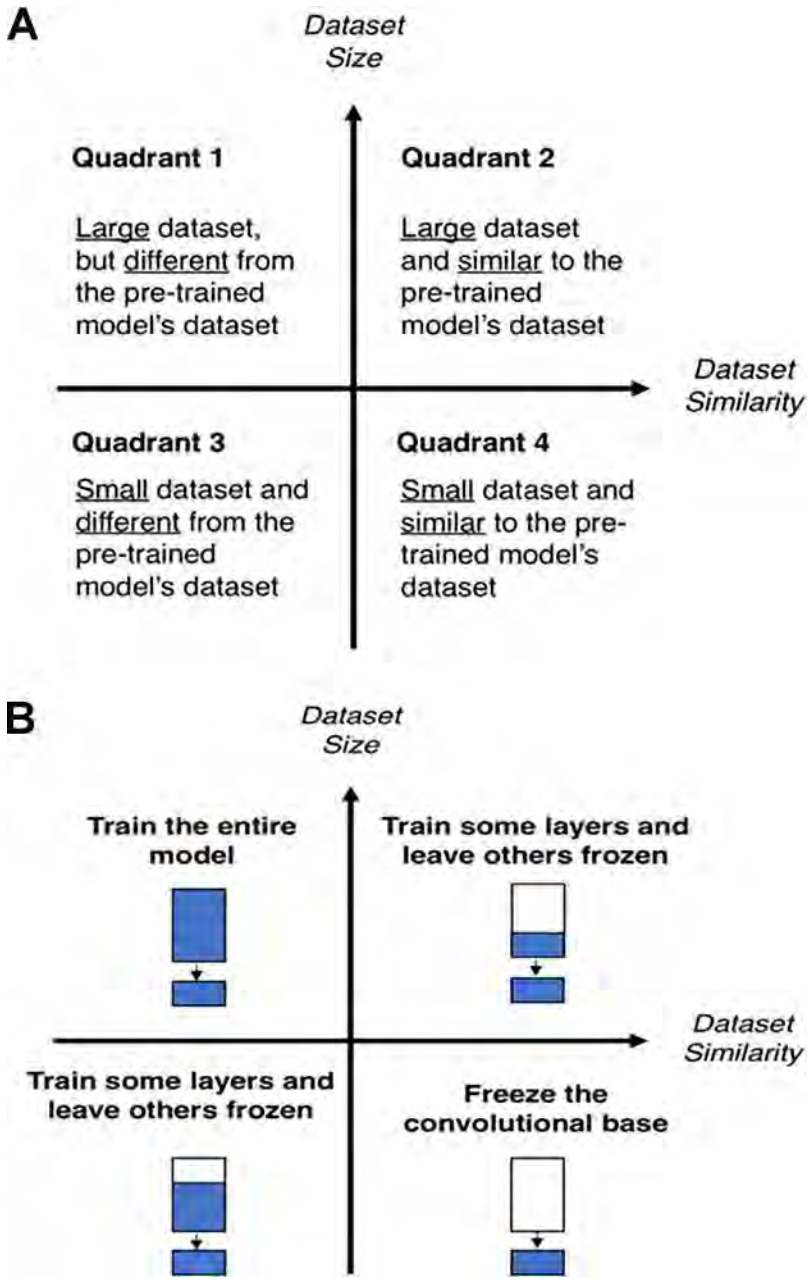


FIG. 8: Transfer learning strategies.[32] (a) Dataset size-similarity matrix; (b) Fine-tuning pre-trained models.

neural network analyze the feature data and classify the data into four categories.

V. RESULTS AND ANALYSIS

The model was trained up to 280 epochs in total to achieve the best results feasible given the technological restrictions. An overview of the results obtained through the process of model validation with respect to different MobileNet models is depicted in Table 3.

For our tests, we utilized the Tensorflow backend and the Keras frameworks for neural - network development as well as dataset preparation. We

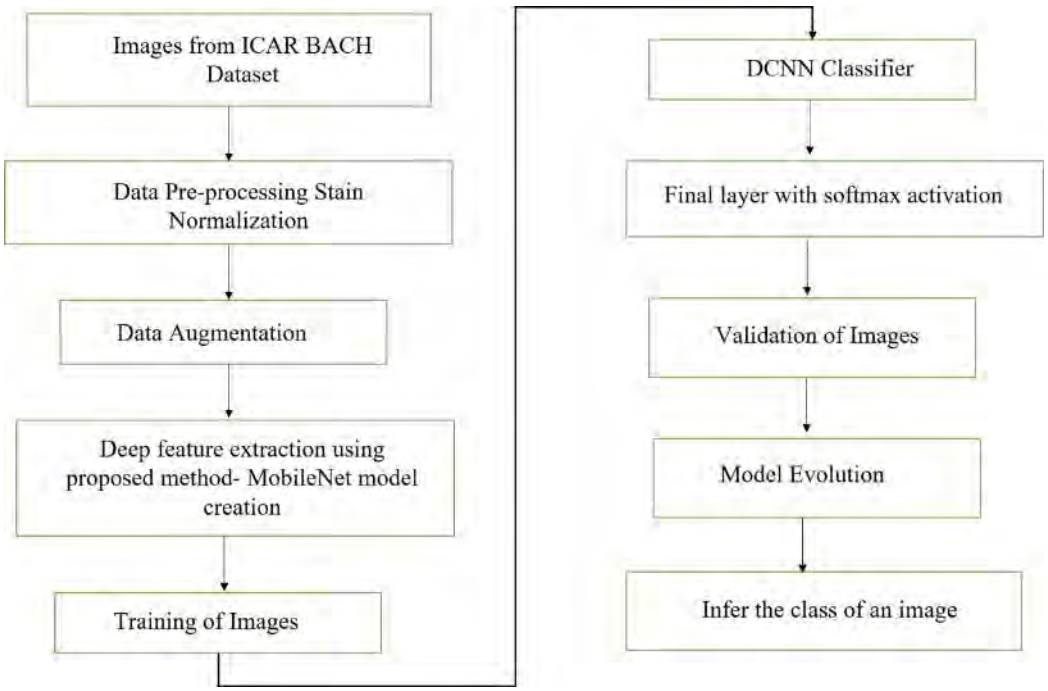


FIG. 9: Flowchart depicting steps involved to perform histopathology image classification for detection of breast cancer

divided the data into three groups, with 88% of the instances used for training and 12% utilized for testing. The training period was set at 280 epochs with 25 training steps each epoch. The overall graphical representation of model training and validation accuracy is shown in Fig. 10.

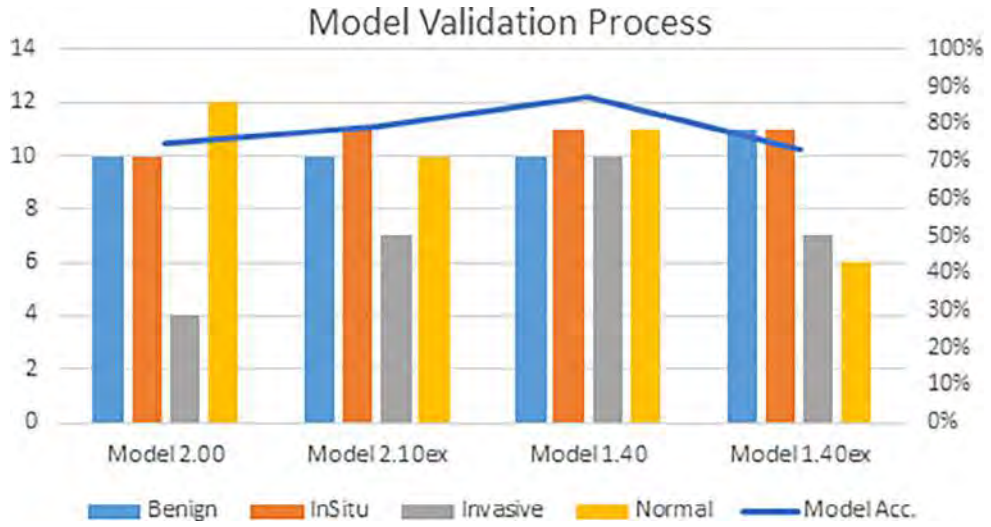


FIG. 10: Graphical overview of the results obtained over the process of the model validation. The x-axis denotes different MobileNet models and the y-axis represents model accuracy (RHS) and the number of correct classifications (LHS).

Figure 11 shows the learning approach for the four classifiers in the BACH dataset with 400 tissue pictures. The model's accuracy improves with time as the epochs continue. In the meanwhile, we can identify some unexpected declines inaccuracy. This is mostly due to the learning rate, which is set at 0.0001. According to our data and the model's properties, this learning rate seems to be rather high. Google Colab trained the model for 280 epochs using its powerful GPU with 2496 compute units, 12.6 GB RAM, and 12 GB GDDR5 memory. The training and validation accuracy of the model can be plotted by the number of DL epochs applied once it has been trained. This depicts that they are moving extremely near to each other, with the distance between the two being as small as possible.

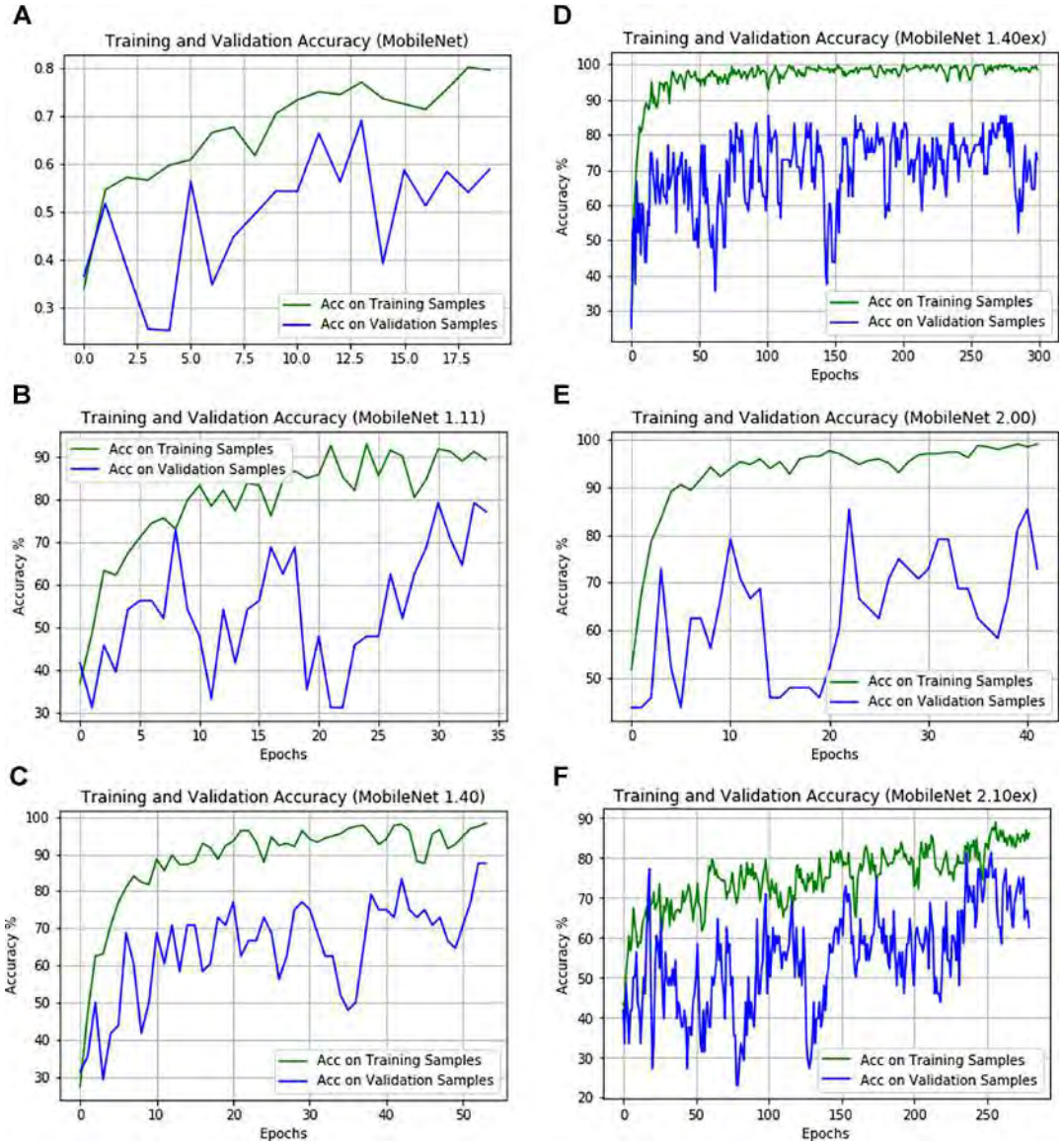


FIG. 11: Training and validation accuracy of different MobileNet models. The x-axis denotes the number of epochs corresponding to each model and the y-axis represents model accuracy in terms of percentage accuracy for MobleNet models. (A) Model 1.10; (B) model 1.11; (C) model 1.40; (D) model 1.40ex; (E) model 2.00; (F) model 2.10ex.

Figure 12 shows the confusion matrix for the classification of histopathology images into normal, benign, *in situ*, and invasive by the number of models trained. The feature score can be computed using the chi-square test with degrees of freedom, k = 10 for selecting the best 30 features out of 1024 features extracted using deep CNN as shown in Table 4. Figure 13 shows the learning approach for the four classifiers in the BACH dataset with 400 tissue pictures.

The scikit train, test and split technique is used to further separate features data (x data) and target data (y data) into train and test. The first 80% of the data was allocated to training, while the remaining 20% was allocated to testing. Setting the test size option to 0.2 controls the train, test, and split fraction. The data that is prepared can be fed to the ML model such as linear support vector classifier kernel to compute percentage accuracy. Similarly, the same steps are carried out to select the 39 best features out of 1024 features.

The graphical representation of the chi-square test score computed for the 30 best features to compute percentage accuracy is shown in Fig. 14. The summary of CAD system development with percentage accuracy is presented in Table 5.

VI. CONCLUSIONS

The main contribution of this article is the categorization of four distinct classes of data using microscopic images as benign, normal, invasive, and *in situ* carcinoma. The ICIAR-BACH competition dataset was enhanced to make it more adaptable for convolutional neural networks. MobileNet's deep CNN architecture is employed in the feature extraction step. This MobileNet has been trained to identify microscopic medical images, and its parameters have been modified to differentiate between four types of input. With multiple architectural versions developed and tested, we created a high-performance design that has an estimated accuracy of 87.5%.

As a result of its excellent accuracy, as well as its resource-efficient architecture, this version of MobileNet design has high potential. This low-cost model presents an intriguing prospect for industry-related studies including breast cancer screening. In the future, we can use modern FPGA boards to build this lightweight deep neural network model. Other deep neural network architectures, such as the residual network (ResNet) and visual geometry group (VggNet) will be proposed in the future to analyze performance parameters.

It is critical to correctly categorize histopathological pictures to correctly determine the kind of breast cancer. DL was used to effectively classify breast cancer. The categorical cross-entropy is used by the model to learn to assign a high probability to the right class and a low probability to the other classes. It is employed in the classification step of

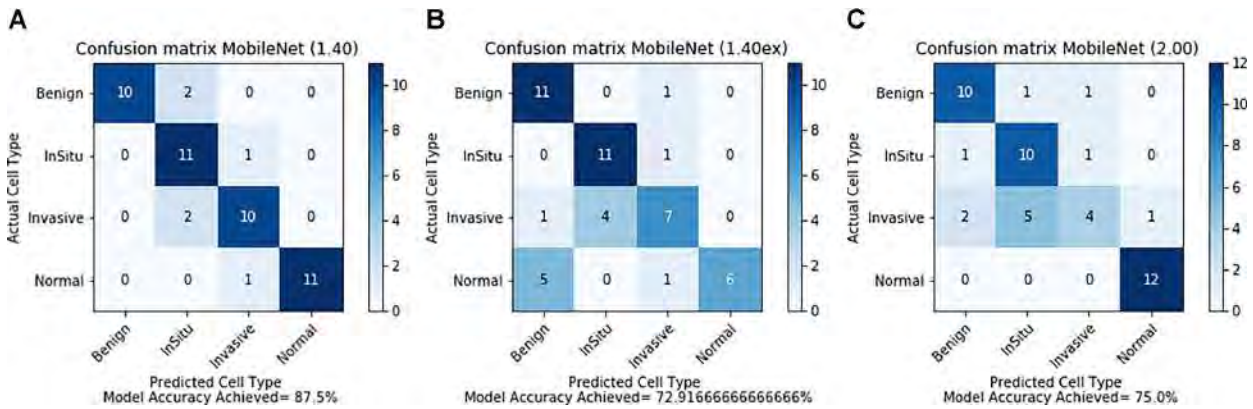


FIG. 12: Confusion matrix for the four classes normal, benign, invasive carcinoma, and *in situ* carcinoma. The x-axis denotes the predicted cell type corresponding to each class of data and the y-axis represents the actual cell type. (A) Model 1.40; (B) model 1.40 ex; (C) model 2.00.

TABLE 4: Summary of the chi-square test score for the 30 best features to compute accuracy

Best feature	Specs	Score
630	630	719.05
798	798	712.83
371	371	697.30
533	533	665.06
183	183	645.84
175	175	612.44
669	669	594.89
830	830	590.01
48	48	589.36
20	20	574.84
210	210	558.34
777	777	551.21
836	836	541.32
949	949	540.86
530	530	522.40
758	758	515.41
540	540	513.50
121	121	511.61
712	712	499.24
914	914	494.21
198	198	493.98
330	330	493.71
806	806	485.31
574	574	481.13
682	682	472.58
282	282	469.01
802	802	467.02
888	888	466.09
877	877	453.98
335	335	442.52

the CNN following the clustering stage, which improves the performance of the suggested system. The suggested method made use of the Python programming language, which enhanced classification accuracy while shortening the model's processing time. To measure training and validation accuracy, the model was trained on Google Colab for 280 epochs with a powerful GPU with 2496 CUDA cores, 12 GB GDDR5 VRAM, and 12.6 GB RAM. The feature score is computed using the chi-square test with degrees of freedom, k = 10 for selecting the best 30 features out of 1024 features extracted using deep CNN. The data that is prepared using the train_test_split technique is fed to the ML model such as linear support vector classifier kernel to compute percentage accuracy. Therefore, deep CNNs with chi-square tests have improved the histopathological image classification of breast cancer

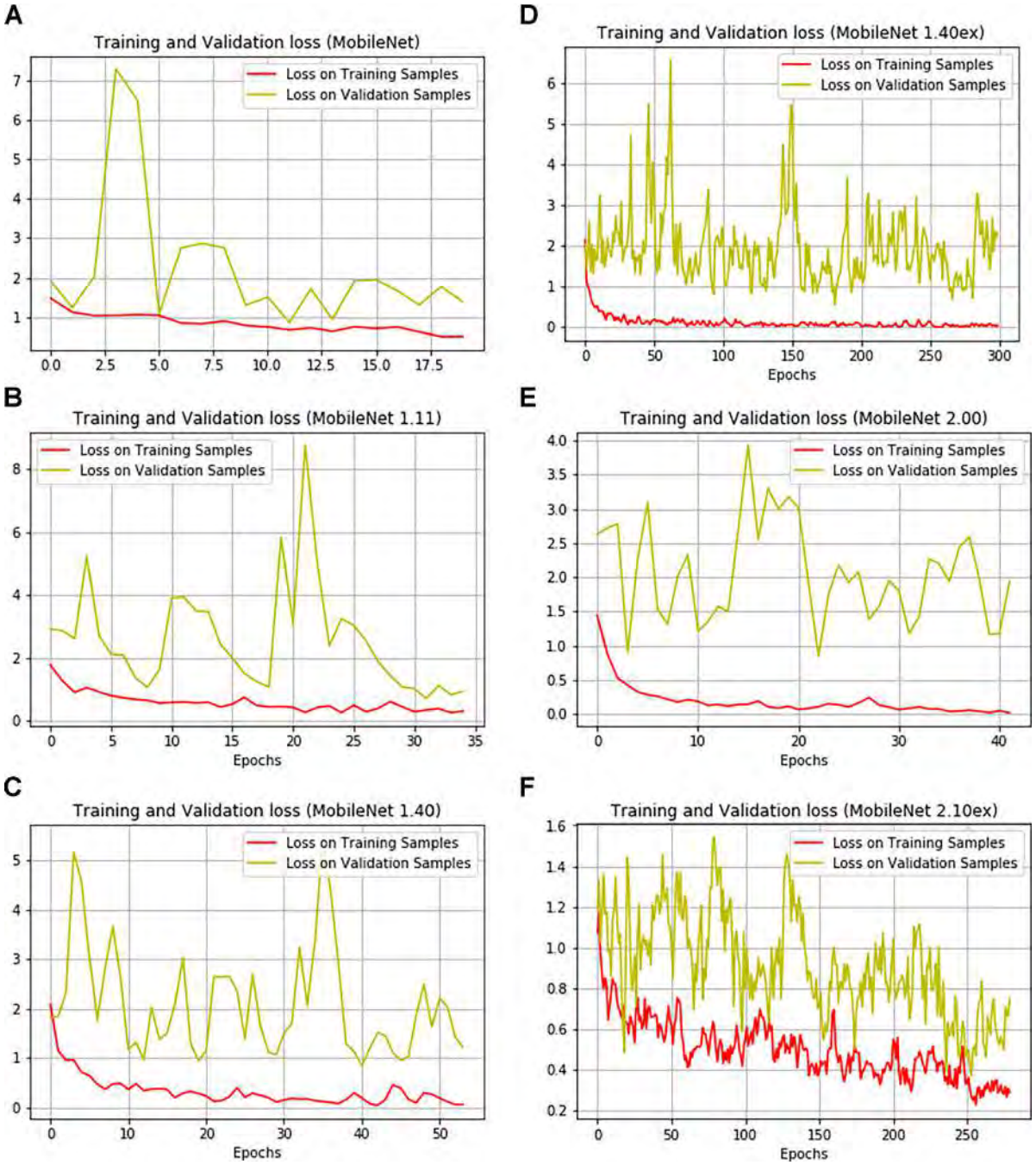


FIG. 13: Training and validation loss (red or bottom lines and yellow or top lines, respectively). The x-axis denotes the number of epochs and the y-axis represents model loss for MobileNet models. (A) Model 1.10; (B) model 1.11; (C) model 1.40; (D) model 1.40ex; (E) model 2.00; (F) model 2.10ex.

with an increase of 11.25% accuracy. The proposed system might be improved further by future studies utilizing a huge dataset of histopathology images in the CNN training process, and even by using huge feature extractions in CNN to build models that can further improve the proposed system. Furthermore,

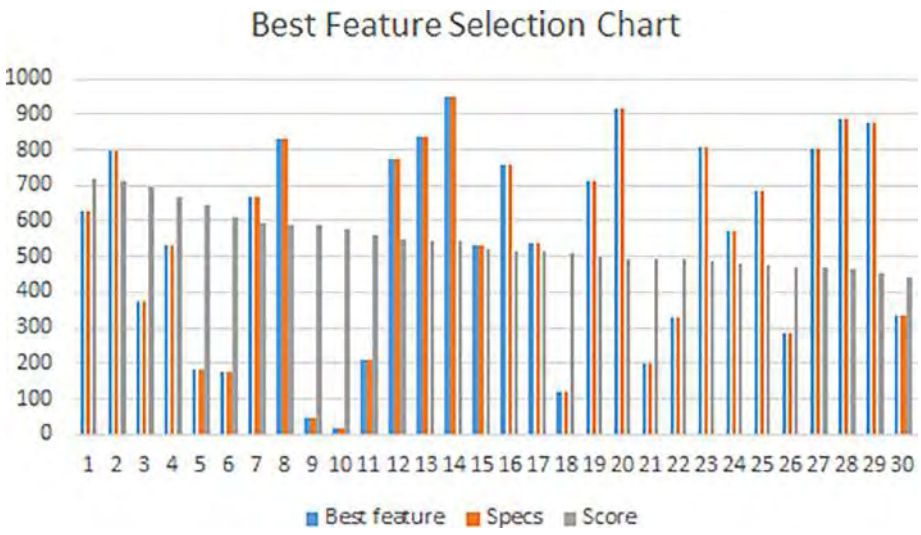


FIG. 14: Graphical chart of chi-square test score computed for 30 best features. The x-axis denotes the 30 best features and the y-axis represents the best score value corresponding to each feature.

TABLE 5: Summary of CAD system development with percentage accuracy

Method	Dataset	Features	Percentage accuracy	Number of classes	Model
Software	Histopathology	1024	87.50	Four	DCNN-Mobilenet 1.40
Software	WBCD	30	97.20	Two	SVM
Software	Histopathology	30	98.75	Four	SVM
Software	Histopathology	39	94.00	Four	SVM

more research is needed on the multi-level categorization of breast cancer.

REFERENCES

1. World Health Organization. Preventing cancer. Available from: http://www.who.int/cancer/prevention/diagnosis-screening/breast-cancer/en/.
2. American Cancer Society. Breast cancer early detection and diagnosis. Available from: https://www.cancer.org/cancer/breast-cancer/screening-tests-and-early-detection.html.
3. Moghbel M, Ooi CY, Ismail N, Hau YW, Memari N. A review of breast boundary and pectoral muscle segmentation methods in computer-aided detection/diagnosis of breast mammography. Artif Intell Rev. 2020;53(3):1873–918.
4. Kozegar E, Soryani M, Behnam H, Salamati M, Tan T. Computer aided detection in automated 3-D breast ultrasound images: A survey. Artif Intell Rev. 2020;53(3):1919–41.
5. Domingues I, Pereira G, Martins P, Duarte H, Santos J, Abreu PH. Using deep learning techniques in medical imaging: A systematic review of applications on CT and PET. Artif Intell Rev. 2020;53(6):4093–160.
6. Murtaza G, Shuib L, Abdul Wahab AW, Mujtaba G, Nweke HF, Al-Garadi MA, Zulfiqar F, Raza G, Azmi NA. Deep learning-based breast cancer classification through medical imaging modalities: State of the art and research challenges. Artif Intell Rev. 2020;53(3):1655–720.
7. Moghbel M, Mashohor S. A review of computer assisted detection/diagnosis (CAD) in breast thermography for breast cancer detection. Artif Intell Rev. 2013;39(4):305–13.
8. De Matos J, Britto Jr AD, Oliveira LE, Koerich AL. Histopathologic image processing: A review. arXiv preprint arXiv:1904.07900. 2019 Apr 16.
9. ICIAR Grand Challenge 2018 on Breast Cancer Histology Images. Available from: https://zenodo.org/record/3632035#.YqgaL6FBxPY.
10. Zhou X, Li C, Rahaman MM, Yao Y, Ai S, Sun C, Wang Q, Zhang Y, Li M, Li X, Jiang T. A comprehensive review for breast histopathology image analysis using classical and deep neural networks. IEEE Access. 2020;8:90931–56.
11. Vo DM, Nguyen NQ, Lee SW. Classification of breast

cancer histology images using incremental boosting convolution networks. Inf Sci. 2019;482:123–38.
12. Li C, Chen H, Li X, Xu N, Hu Z, Xue D, Qi S, Ma H, Zhang L, Sun H. A review for cervical histopathology image analysis using machine vision approaches. Artif Intell Rev. 2020;53(7):4821–62.
13. Jadoon MM, Zhang Q, Haq IU, Butt S, Jadoon A. Three-class mammogram classification based on descriptive CNN features. BioMed Res Int. 2017;2017:3640901.
14. Loukas C, Kostopoulos S, Tanoglidi A, Glotsos D, Sfikas C, Cavouras D. Breast cancer characterization based on image classification of tissue sections visualized under low magnification. Comput Math Methods Med. 2013;2013:829461.
15. Gurcan MN, Boucheron L, Can A, Madabhushi A, Rajpoot N, Yener B. Histopathological image analysis: A review. IEEE Rev Biomed Eng. 2009;2:147–71.
16. Kowal M, Filipczuk P, Obuchowicz A, Korbicz J, Monczak R. Computer-aided diagnosis of breast cancer based on fine needle biopsy microscopic images. Comput Biol Med. 2013;43(10):1563–72.
17. Tang J, Rangayyan RM, Xu J, El Naqa I, Yang Y. Computer-aided detection and diagnosis of breast cancer with mammography: Recent advances. IEEE Trans Inf Technol Biomed. 2009;13(2):236–51.
18. Bengio Y, Courville A, Vincent P. Representation learning: A review and new perspectives. IEEE Trans Pattern Anal Mach Intell. 2013;35(8):1798–828.
19. Litjens G, Sánchez CI, Timofeeva N, Hermsen M, Nagtegaal I, Kovacs I, Hulsbergen-Van De Kaa C, Bult P, Van Ginneken B, Van Der Laak J. Deep learning as a tool for increased accuracy and efficiency of histopathological diagnosis. Sci Rep. 2016;6:26286.
20. Yari Y, Nguyen TV, Nguyen HT. Deep learning applied for histological diagnosis of breast cancer. IEEE Access. 2020;8:162432–48.
21. Mansour RF. A robust deep neural network based breast cancer detection and classification. Int J Comput Intell Appl. 2020;19(1):2050007.
22. Erfankhah H, Yazdi M, Babaie M, Tizhoosh HR. Heterogeneity-aware local binary patterns for retrieval of histopathology images. IEEE Access. 2019;7:18354–67.
23. Lichtblau D, Stoean C. Cancer diagnosis through a tandem of classifiers for digitized histopathological slides. PLoS One. 2019;14(1):e0209274.
24. Gour M, Jain S, Sunil Kumar T. Residual learning based CNN for breast cancer histopathological image classification. Int J Imaging Syst Technol. 2020;30(3):621–35.
25. Hu C, Sun X, Yuan Z, Wu Y. Classification of breast cancer histopathological image with deep residual learning. Int J Imaging Syst Technol. 2021;31(3):1583–94.
26. Aresta G, Araújo T, Kwok S, Chennamsetty SS, Safwan M, Alex V, Marami B, Prastawa M, Chan M, Donovan M, Fernandez G. Bach: Grand challenge on breast cancer histology images. Med Image Anal. 2019;56:122–39.
27. Kassani SH, Kassani PH, Wesolowski MJ, Schneider KA, Deters R. Classification of histopathological biopsy images using ensemble of deep learning networks. arXiv preprint arXiv:1909.11870. 2019 Sep 26. Available from: https://arxiv.org/abs/1909.11870
28. Kausar T, Wang M, Idrees M, Lu Y. HWDCNN: Multi-class recognition in breast histopathology with Haar wavelet decomposed image based convolution neural network. Biocybern Biomed Eng. 2019;39(4):967–82.
29. Yang Z, Ran L, Zhang S, Xia Y, Zhang Y. EMS-Net: Ensemble of multiscale convolutional neural networks for classification of breast cancer histology images. Neurocomputing. 2019;366:46–53.
30. Kumar A, Prateek M. Localization of nuclei in breast cancer using whole slide imaging system supported by morphological features and shape formulas. Cancer Manag Res. 2020;12:4573.
31. Chorianopoulos AM, Daramouskas I, Perikos I, Grivokostopoulou F, Hatzilygeroudis I. Deep learning methods in medical imaging for the recognition of breast cancer. In 2020 11th International Conference on Information, Intelligence, Systems and Applications (IISA). IEEE Computer Society; 2020. p. 1–8.
32. Marcelino P. Transfer learning from pre-trained models: How to solve any image classification problem quickly and easily. Available from: https://towardsdatascience.com/transfer-learning-from-pre-trained-models-f2393f124751.
33. Aresta G, Araújo T, Kwok S, Chennamsetty SS, Safwan M, Alex V, Marami B, Prastawa M, Chan M, Donovan M, Fernandez G. Bach: Grand challenge on breast cancer histology images. Med Image Anal. 2019;56:122–39.
34. Breast Cancer Histopathological Database (BreakHis). Available from: https://www.kaggle.com/datasets/ambarish/breakhis.
35. IMEDIATREAT Dataset. Available from: https://sites.google.com/site/imediatreat/data-sets.

Received: January 16, 2023. Revised: April 13, 2023. 509

International Journal of
Intelligent Engineering & Systems

http://www.inass.org/

Multi Objective Energy Based Improved Jellyfish Swarm Optimization for Effective Cluster Head Discovery in UWSN

Seema Swamy Gowda[1]* **Ambika Ramalingappa[2]**

[1]*Department of Electronics and Communication Engineering, Sir M Visvesvaraya Institute of Technology, Bengaluru, India*
[2] *Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bengaluru, India*
* Corresponding author's Email: seema_ec@sirmvit.edu

Abstract: Underwater wireless sensor networks (UWSNs) have a huge amount of sensors located underwater to collect data from the underwater scenario. UWSN is considered a promising method for monitoring and exploring an underwater scenario. Energy-efficient and reliable data broadcasting are considered challenging tasks, because of the limited energy source of sensors. To address this issue, an energy-efficient cluster head (CH) selection and multi-hop routing are developed in UWSN. The multi objective energy based improved jellyfish swarm optimization (MOEIJSO) is proposed to select optimal CHs from normal sensors. The input parameters of MOEIJSO are residual energy, neighbor node distance, sink distance, and CH balancing factor. Next, multi-hop routing is developed by using ant colony optimization (ACO) for delivering the data packets. Therefore, the proposed MOEIJSO-ACO method is used to improve energy efficiency while increasing data delivery. The proposed MOEIJSO-ACO method is evaluated by using the alive and dead nodes, total energy consumption, data received in base station (BS), and life expectancy. The existing researches such as metaheuristics-based clustering with a routing (MCR) for UWSN, modified low energy adaptive clustering hierarchy (MLEACH) and cooperative energy-efficient routing (CEER) approach are used to compare the MOEIJSO-ACO method. The alive nodes of the MOEIJSO-ACO are 399 for 800 rounds, which is high when compared to the MCR-UWSN.

Keywords: Ant colony optimization, Energy efficiency, Life expectancy, Multi-hop routing, Multi-objective energy based improved jellyfish swarm optimization, Underwater wireless sensor networks.

1. Introduction

Earth is occupied by a 3/4 ratio of water using seas, lakes, canals, rivers and streams. Since a huge amount of unexplored and hidden resources exist underwater which require to be discovered whereas underwater environments are highly sophisticated for humans to discover. Accordingly, an exploration of the underwater environment is possible with the use of wireless technology [1, 2]. UWSN is developing technology that is utilized for observing and discovering the changes in an aqueous environment [3]. UWSN is a wireless network that comprises a group of sensors and autonomous underwater vehicles for collecting and sensing data. These sensors accomplish direct or indirect transmission of information to the surface sensor (i.e., sink). Next, the data from the sink is transmitted to an offshore monitoring center to analyze and study the gathered data [4]. The sensors of UWSN are located underwater for evaluating the monitoring features such as density, temperature, pressure and so on [5]. UWSN is extensively used in diverse fields such as military target tracking, oil/gas spill monitoring, submarine detection, offshore exploration, disaster prevention and so on [6]. The UWSN faces various issues such as strong network dynamics, expensive deployment, less available bandwidth and restricted battery energy [7].

The battery energy of underwater sensors is restricted whereas these batteries cannot be simply

replaced or charged in the ocean. The energy consumption of the UWSN is high because of the higher communication power, restricted bandwidth capacity, higher and variable time delay and higher bit error rate [8-10]. The limited battery source of the node creates the failure nodes over the UWSN [11]. These aforementioned features of UWSN create an effect on sensor data gathering and long-distance communication. Hence, the routing is required to be developed as simple and energy efficient for UWSNs. This indicates that effective and reliable data broadcasting to the sink node is a highly challenging task as well as it is one of the essential concerns in UWSN [12, 13]. A cluster-based data collection is developed for improving the energy efficacy of the WSN [14]. The main objective of the clustering approach is used to separate the network into small divisions namely clusters. An entire cluster is represented by using the CH that gathers the observed information from sensors and it is broadcasted to the sink by utilizing the CH as the next hop. The developed clustering approach is used to minimize the overall routing distance and overhead of sensors [15].

The contributions are concise as follows:

- The MOEIJSO based CH discovery is developed for enhancing the energy efficiency of UWSN. The IJSO is chosen for this research because of its enhanced local search capacity which is used to obtain in best CH solutions.
- Further, the ACO-based route discovery is used for identifying the route from transmitter CH to BS. Therefore, the MOEIJSO-ACO is used to improve the life expectancy of UWSN while enhancing the data delivery.

The paper organization is sorted as follows: Section 2 delivers the related works of the energy-efficient data transmission of UWSN. The MOEIJSO-ACO based data broadcasting is detailed in section 3 whereas the results are presented in section 4. Finally, the conclusion is given in section 5.

2. Related work

This section provides the existing energy efficient data broadcasting-related research in UWSN.

Chenthil and Jesu Jayarin, [16] presented the multilayer clustering-based butterfly optimization routing (MCBOR) for broadcasting the data to the receiver without loss. The list of parameters such as energy, node coverage, processing power and communication cost were utilized to choose the candidate CH. Further, butterfly optimization was used for transmitting the data to the receiver. The developed MCBOR was used to improve the PDR and minimize the loss over the UWSN. Moreover, the developed MCBOR was required to be analyzed in a large-scale network.

Bhattacharjya, [17] developed an energy-efficient UWSN for reducing energy expenditure and improving performance in the underwater environment. The developed cluster-based UWSN (CUWSN) utilized the advantages of CH and multi-hop data transmission. Here, the CH was chosen based on the residual energy of the sensors over the network. The developed CUWSN was used to improve life expectancy based on multi-hop transmission. The CH discovery was considered only the residual energy, however, a data transmission distance was required to be considered in CUWSN.

Faheem, [18] developed the dynamic firefly mating optimization-based routing protocol namely FFRP for the UWSN scenario. A stable and reliable route was discovered by using the developed FFRP. A balancing of data traffic load using FFRP was used to reduce energy usage and latency. Accordingly, the data transmission over the stable link was used to improve the data delivery of the network.

Subramani, [19] presented MCR for UWSN to enhance energy efficiency. In MCR-UWSN, the cultural emperor penguin optimizer-based clustering (CEPOC) was developed for generating the clusters. The developed CEPOC considered the node's motion and average distance for discovering the CH. Further, the routing was developed by using grasshopper optimization according to the list of nodes, energy, and distance. For an effective CH selection, cluster balancing among the nodes was essential to enhance the performance.

Rizvi, [20] presented an energy efficient approach i.e., MLEACH for UWSN. The CHs were randomly chosen in this MLEACH for preserving the energy in each round. The non‐persistent carrier sense multiple access was used to establish the data transmission among CH and its respective node. The direct data transmission from the CH to BS was resulted in higher energy depletion of nodes.

Ahmad [21] developed the cooperative energy-efficient routing (CEER) approach for generating the reliable network and increasing the network lifetime. The problem of hotspot was eliminated by minimizing the energy utilization based on sink mobility approach. The network reliability was obtained by using the cooperative technique in UWSN. However, the clustering over the network

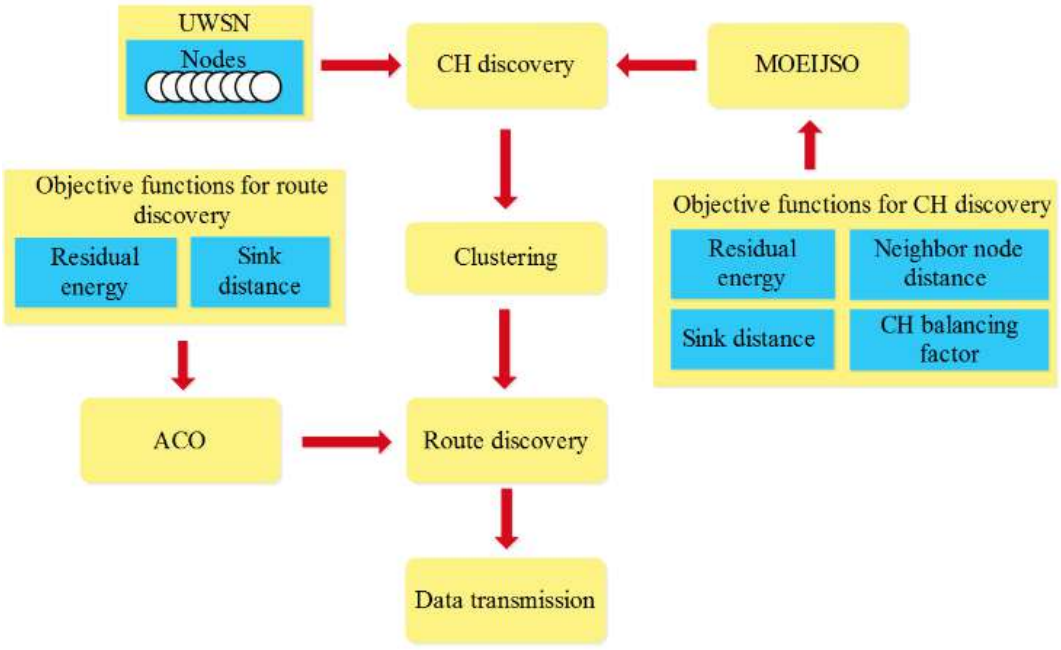


Figure. 1 Block diagram for the MOEIJSO-ACO

was required to be developed for an additional improvement of lifetime.

The limitations found from the related works are inadequate fitness function, higher energy usage and routing without clustering. In this research, an appropriate objective measures are used with MOEIJSO to develop an effective clustering and routing for minimizing the energy usage of overall network.

3. MOEIJSO-ACO method

In this research, effective data transmission is obtained by using a MOEIJSO based optimal CH selection and ACO based routing in UWSN. The energy usage of sensors is minimized by developing a clustering along with a multi-hop transmission using MOEIJSO-ACO. The reduction in energy utilization helps to improve the life expectancy of the sensors which results in higher data delivery. The block diagram of the MOEIJSO-ACO is presented in Fig. 1.

3.1 Sensor initialization

The nodes are positioned randomly in the WSN followed by the CHs chosen using MOEIJSO. The clusters are formed in the UWSN, once the CHs are chosen from the network. Further, the route from the transmitter CH to receiver BS is discovered using ACO. The optimal CH and route discovery are explained in the following sections.

3.2 Discovering optimal CH using MOEIJSO

The MOEIJSO is used for discovering the optimal CHs from the sensors of UWSN. The conventional Jellyfish Swarm Optimization (JSO) is motivated based on jellyfish's search-feeding activities and movement patterns through the ocean. The main difference between JSO and IJSO [22] is that the incorporation of the sinusoidal factor for improving the local search capacity resulted in the best optimal solutions. Further, the IJSO is developed as MOEIJSO for selecting optimal CHs from the sensors.

3.2.1. Representation and initialization

An initial solution of MOEIJSO i.e., jellyfish has a set of nominee nodes to be chosen as CH. Here, each jellyfish is initialized with the ID of a random node between 1 and N, where the total amount of nodes in the underwater scenario is denoted as N. The MOEIJSO's i th jellyfish is denoted as $X_i = (X_{i,1}, X_{i,2}, \ldots, X_{i,dim})$, where dim denotes each jellyfish's dimension as amount of CHs.

3.2.2. Iterative process

The IJSO has three different rules that are mentioned as follows: 1) two types of motion exist in jellyfish such as following the ocean current motion and moving in the jellyfish population. Here, the time-controlled mechanism is used to switch among these motions; 2) Jellyfish moves in the ocean for searching food, where the jellyfishes are attracted to

the position with more food, and 3) Food position and objective value defines the amount of discovered food.

3.2.2.1. Following the movement of ocean current

The movement way of current is denoted as $\overrightarrow{trend}$ that defines the average of the overall vector sum from every individual current location of jellyfish to the current finest location. Eq. (1) shows the current $\overrightarrow{trend}$ for the direction of motion.

$$\overrightarrow{trend} = X^* - DF \tag{1}$$

Where the optimal location is denoted as X^* and the difference between the current jellyfish placed in an optimal location and the average location of all jellyfish is denoted as DF is expressed in Eq. (2).

$$DF = e_c \mu \tag{2}$$

Where, the parameter which defines the food attractiveness to jellyfish is denoted as e_c which is expressed in Eq. (3) and the average location of all jellyfish is denoted as μ.

$$e_c = \beta \times rand(0,1) \tag{3}$$

Where, the distribution coefficient is denoted as β. Eq. (4) shows the jellyfish's location update for following ocean current motion.

$$X_i(t+1) = X_i(t) + rand(0,1) \times \overrightarrow{trend} = X_i(t) + rand(0,1) \times (X^* - \beta \times rand(0,1) \times \mu) \tag{4}$$

Where, $X_i(t)$ is jellyfish's current location and $X_i(t+1)$ is location obtained after updating in search space.

3.2.2.2. Group movements

The jellyfish's group motion is categorized into passive (i.e., Class A) and active motion (i.e., Class B). The movement of Class A occurs, when the swarm of jellyfish is generated whereas the movement of Class B is carried out after the Class A movement. The class A motion is accomplished as the jellyfish traveling around their current location. In IJSO, the sinusoidal adaptive factor (S) shown in Eq. (5) is used for improving the local search capability.

$$S = 1 + \sin\frac{\pi(2T+t)}{2T} \tag{5}$$

The location update of class A motion with sinusoidal adaptive factor is expressed in Eq. (6).

$$X_i(t+1) = X_i(t) + S \times \gamma \times rand(0,1) \times (u_b - l_b) \tag{6}$$

Where lower and upper bound are denoted as l_b and u_b respectively; jellyfish's movement coefficient is denoted as γ where it is taken as 0.1. Accordingly, the range of motion is mainly based on l_b, u_b and γ.

The jellyfish accomplishes the motion of class B and approaches the discrete jellyfish with huge amount of food at its position and collects with food as its target. The location update of class B motion is expressed in Eq. (7).

$$X_i(t+1) = X_i(t) + \overrightarrow{step} \tag{7}$$

Where, i th jellyfish's step length with the direction of motion is denoted as $\overrightarrow{step}$ which is expressed in Eq. (8).

$$\overrightarrow{step} = rand(0,1) \times \vec{D} \tag{8}$$

Where, ith jellyfish's motion direction is denoted as $\vec{D}$ which is shown in Eq. (9).

$$\vec{D} = \begin{cases} X_j(t) - X_i(t) & , if\ f(X_i(t)) \geq f\left(X_j(t)\right) \\ X_i(t) - X_j(t) & , if\ f\left(X_j(t)\right) \geq f(X_i(t)) \end{cases} \tag{9}$$

Where, i th and j th jellyfish's location are represented as $X_i(t)$ and $X_j(t)$ respectively; the objective function according to the X is denoted as f.

3.2.2.3. Time control mechanism

The time control mechanism is used for performing the switching among three modes of movement of jellyfish where the time control function $c(t)$ is expressed in Eq. (10).

$$c(t) = \left|\left(1 - \frac{t}{T}\right) \times (2 \times rand(0,1) - 1)\right| \tag{10}$$

Where, the current iteration is denoted as t; maximum iterations are denoted as T and the random number is generated in the range of $[0, 1]$. If $c(t) \geq 0.5$, then the jellyfish movement is controlled in ocean currents; otherwise, the jellyfish movement is controlled by the intra-group motion of jellyfish.

3.2.3. Derivation of multiple objective functions for MOEIJSO

Multiple objectives used in this MOEIJSO for selecting the optimal CHs are residual energy (f_1), neighbor node distance (f_2), sink distance (f_3) and CH balancing factor (f_4). Eq. (11) shows the multiple objective functions used in MOEIJSO for improving the searching process during the CH discovery.

$$f = \psi_1 \times f_1 + \psi_2 \times f_2 + \psi_3 \times f_3 + \psi_4 \times f_4 \quad (11)$$

Where, $\psi_1 - \psi_4$ is the weight value assigned to each objective value. The multiple objectives are defined as follows:

- The energy usage of CH is essential in UWSN, due to the different processes such as information collection, aggregation, and distribution over the network. For an effective transmission, the node with huge remaining energy is desired as a CH and Eq. (12) shows the computation of remaining energy.

$$f_1 = \sum_{i=1}^{dim} \frac{1}{E_{CH_i}} \quad (12)$$

Where, E_{CH_i} denotes the remaining energy of the ith CH

- The distance between the sensors shown in Eq. (13) is the neighbor node distance and the distance between the CH and BS shown in Eq. (14) is the sink distance. The energy depletion of the sensor is proportional to the distance among the sensors. Therefore, the CH with a lesser broadcast distance is preferred as CH.

$$f_2 = \sum_{j=1}^{dim} \left(\sum_{i=1}^{CM_j} dis(N_i, CH_j)/CM_j \right) \quad (13)$$

$$f_3 = \sum_{i=1}^{dim} dis(CH_i, BS) \quad (14)$$

Where, the number of cluster members for jth cluster is denoted as CM_j; distance from node i to CH j is denoted as $dis(N_i, CH_j)$ and distance from the CH i to receiver BS is denoted as $dis(CH_i, BS)$.

- In UWSN, there is a possibility that some big clusters are formed along with the small clusters. Therefore, the CH balancing factor shown in Eq. (15) is used for balancing the clusters which result in improved energy balancing over the UWSN.

$$f_4 = \sum_{i=1}^{dim} \frac{AN}{dim} - CM_j \quad (15)$$

Where the amount of alive nodes is denoted as AN.

The above-mentioned objective values are utilized for selecting adequate CHs from normal sensors. The remaining energy of the sensors is used to compute the possibility of a failure node during the data broadcasting because the failure node needs to be avoided for increasing the data delivery. On the other hand, the distance factors and CH balancing factor is used for improving the energy efficiency of the UWSN.

3.3 Cluster formation

After selecting the CHs using MOEIJSO, the cluster members are assigned to their respective CHs. The distance and residual energy are considered in a potential function shown in Eq. (16) for creating the clusters.

$$Potential\ function\ (N_i) = \frac{E_{CH}}{dis(N_i, CH)} \quad (16)$$

3.4 Route discovery using ACO

The route among the transmitter CH and receiver BS is discovered by ACO, once the clustering is done in the network. The initial weight value of each path is calculated according to the distance between the nodes. The node transition rule is the possibility of selecting m as relay CH for the lth CH computed by ant n is expressed in Eq. (17).

$$P_{lm}^{n} = \begin{cases} \frac{[\tau_{lm}(t)]^{\alpha}[\eta_{lm}(t)]^{\beta}}{\sum_{o \epsilon \mathcal{N}_n}[\tau_{lo}(t)]^{\alpha}[\eta_{lo}(t)]^{\beta}} & if\ m \epsilon \mathcal{N}_n \\ 0 & otherwise \end{cases} \quad (17)$$

Where, P_{lm}^{n} denotes the probability of node selection; τ_{lm} and η_{lm} denotes the pheromone intensity and heuristic value respectively; α and β are used to control the relative importance of τ_{lm} and η_{lm}, and the set of CHs n that doesn't visit the ant is denoted as $\mathcal{N}_n$.

The real ant's foraging process is replicated by artificial ants. The rule of node transition is used to select the next relay CH when the transmitter CH is required for transmitting the data. The path is retraced to the transmitter CH, when the relay CH i.e., ant reached the BS. Accordingly, the pheromone value of the path is updated according to the pheromone

update rule expressed in Eq. (18) which includes pheromone evaporation and reinforcement. The pheromone evaporation and reinforcement minimize or maximize the path's pheromone respectively.

$$\tau_{lm}^{new} = (1-\rho)\tau_{lm}^{old} + \sum_{n=1}^{A} \Delta\tau_{lm}^{n} \quad (18)$$

Where, the amount of ants in ACO is denoted as A and the pheromone decay coefficient is ρ. Eq. (19) shows the computation of pheromone quantity.

$$\Delta\tau_{lm}^{n} = \begin{cases} \frac{Q}{a_n} & if\ the\ ant\ n\ travelled\ route\ (l,m) \\ 0 & otherwise \end{cases} \quad (19)$$

Where, the constant value is denoted as Q and the objective function of the route is a_n that is shown in Eq. (20). The objective function of ACO uses the residual energy and sink distance

$$a_n = \varepsilon_1 \times \sum_{i=1}^{dim} \frac{1}{E_{CH_i}} + \varepsilon_2 \times \sum_{i=1}^{dim} dis(CH_i, BS) \quad (20)$$

Where, ε_1 and ε_2 are the weight parameters assigned to objective values considered in route discovery. The failure node is avoided by considering the remaining energy whereas the distance is considered for minimizing the energy by generating the route with less transmission distance.

4. Results and discussion

The outcomes of the MOEIJSO-ACO are detailed in this section. The implementation and simulation of the MOEIJSO-ACO are developed in MATLAB R2020b where the system uses the i5 processor and 6GB of RAM. The underwater scenario is designed with 50 sensors deployed in the area of $100 \times 100 \times 50m^3$. The base station for this UWSN scenario is located in the coordinates of $50, 50, and\ 50$ whereas the sensors are equipped with 5J of energy. Table 1 shows the simulation parameters considered for this MOEIJSO-ACO method.

4.1 Performance analysis

The MOEIJSO-ACO method is analyzed by using alive and dead nodes, total energy consumption, data received in BS and life expectancy. Classical approaches such as Low Energy Adaptive Clustering Hierarchy (LEACH) and Distributed Energy Efficient Clustering (DEEC) are used to evaluate the MOEIJSO-ACO method.

Table 1. Simulation parameters

Parameter	Value
Number of sensors	50
Location of BS	$50, 50, 50$
Network size	$100 \times 100 \times 50m^3$
The initial energy of sensors	5J
Size of packet	4000s

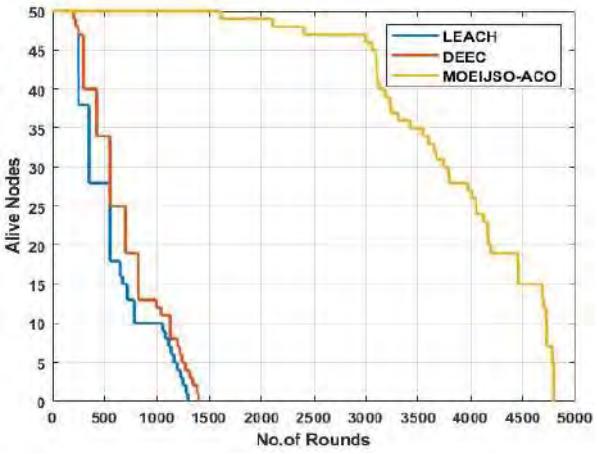


Figure. 2 Alive nodes vs rounds

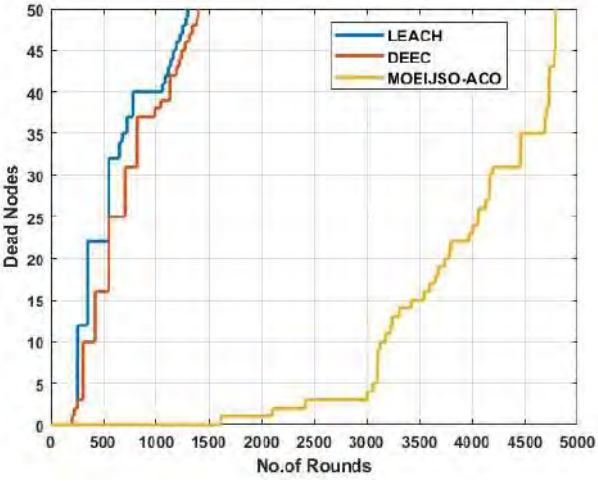


Figure. 3 Dead nodes vs rounds

4.1.1. Alive & dead nodes

The amount of nodes with sufficient energy to perform the communication is alive nodes whereas dead nodes don't have sufficient energy to accomplish the data transfer over the network. The alive nodes and dead node comparison for MOEIJSO-ACO with LEACH and DEEC are shown in Fig. 2 and Fig. 3 respectively. The alive nodes of MOEIJSO-ACO are increased than the LEACH and DEEC by minimizing the energy usage of the nodes. The energy usage of nodes for MOEIJSO-ACO is reduced based on effective balancing between clusters and shortest path discovery.

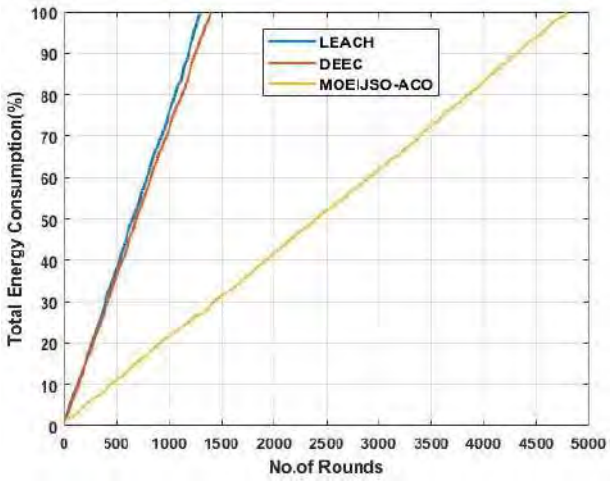


Figure. 4 Total energy consumption vs rounds

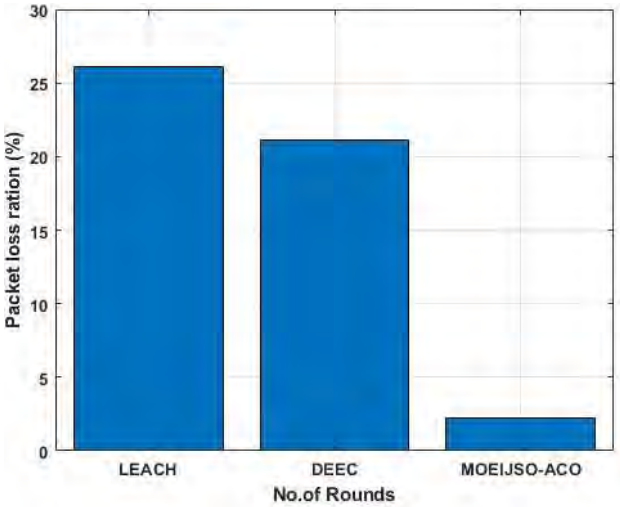


Figure. 7 Packet loss ratio

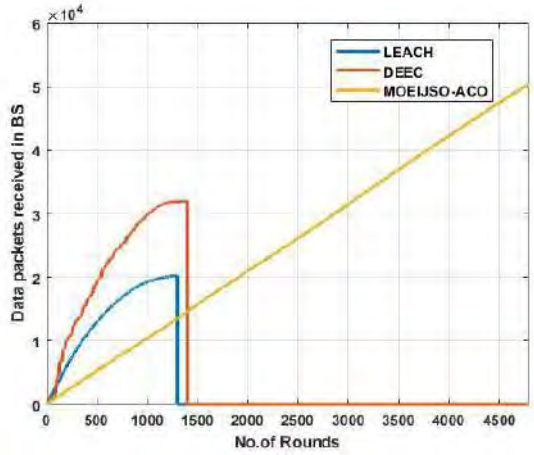


Figure. 5 Data packets received in BS vs rounds

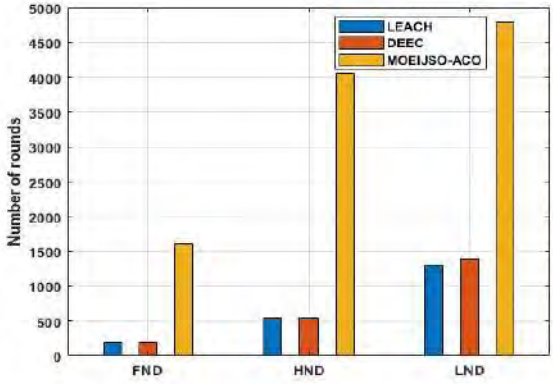


Figure. 8 Life expectancy

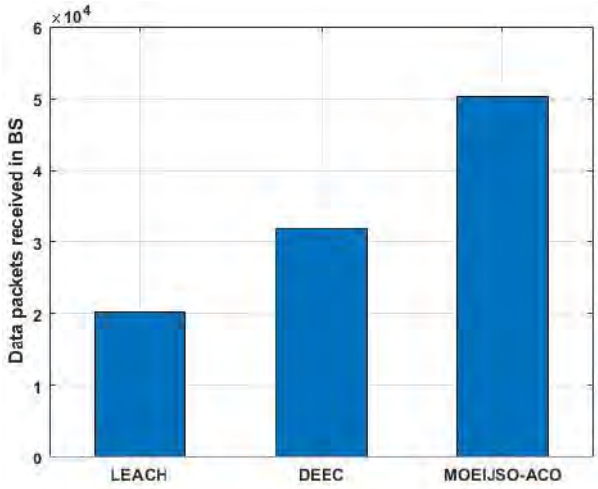


Figure. 6 Maximum data packets received in BS

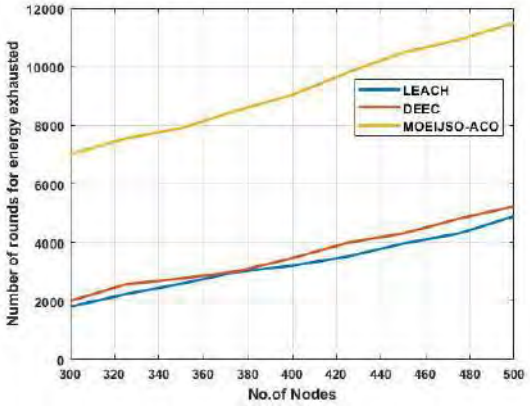


Figure. 9 Number of rounds for energy exhausted vs nodes

4.1.2. Total energy consumption

The energy depleted while transmitting and receiving the data is total energy consumption. Fig. 4 displays the energy expenditure analysis for MOEIJSO-ACO with LEACH and DEEC. This analysis shows that the MOEIJSO-ACO achieves lesser energy consumption than the LEACH and DEEC. An optimal CH selection using MOEIJSO along with the multi hop routing developed by ACO is used to reduce energy usage. The ACO-based shortest route discovery is also used to minimize the node's energy usage.

4.1.3. Data packets received in BS and PLR

The amount of data packets collected by the BS and its maximum value are illustrated in Figs. 5 and 6 where the MOEIJSO-ACO was compared with LEACH and DEEC. Further, the graph for PLR is shown in Fig. 7. This analysis shows that the MOEIJSO-ACO has high data delivery than the LEACH and DEEC. The mitigation of failure nodes in MOEIJSO based CH selection and reducing the energy consumption of nodes through ACO based multi hop data broadcasting helps to increase the amount of data packets gathered in BS.

4.1.4. Life expectancy

Life expectancy represents the time at which the nodes active during the transmission. There are three different parameters such as first node die (FND), half node die (HND) and last node die (LND) are used while analyzing the life expectancy.

The life expectancy comparison for MOEIJSO-ACO with LEACH and DEEC are displayed in Fig. 8. Further, the number of rounds for energy exhausted is shown in Fig. 9 in terms of nodes. This analysis depicts that the MOEIJSO-ACO has a better lifetime than the LEACH and DEEC.

4.2 Comparative analysis

The existing research of MCR-UWSN [19], MLEACH [20] and CEER [21] are used for evaluating the MOEIJSO-ACO method. Here, the comparison is made for 400 nodes, 50 nodes and 225 nodes variations. The MCR-UWSN [19] is taken for 400 nodes comparison, MLEACH [20] is taken for 50 nodes comparison and CEER [21] is taken for 225 nodes comparison. Tables 2, 3 and 4 show the comparison of MOEIJSO-ACO with MCR-UWSN [19], MLEACH [20] and CEER [21] respectively. The life expectancy comparison of MOEIJSO-ACO is shown in Table 5. The graph for life expectancy comparison of MOEIJSO-ACO with CEER [21] is shown in Fig. 10. This comparison shows the MOEIJSO-ACO has improved performance than the MCR-UWSN [19], MLEACH [20] and CEER [21]. The usage of energy is minimized in MOEIJSO-ACO based on balancing between clusters and shortest path discovery. The multi hop routing developed by ACO is used to achieve the reliable data transmission while reducing the energy. Thus, the life expectancy of the MOEIJSO-ACO is improved than the MCR-UWSN [19], MLEACH [20] and CEER [21].

Table 2. Comparison of MOEIJSO-ACO with MCR-UWSN

Performances	**Methods**	**Number of rounds**				
		200	400	600	800	1200
Alive nodes	MCR -UWSN [19]	400	400	400	400	0
	MOEIJSO-ACO	400	399	399	399	394
Dead nodes	MCR -UWSN [19]	0	0	0	0	400
	MOEIJSO-ACO	0	1	1	1	6
Total Energy Consumption (%)	MCR-UWSN [19]	3	12	22	42	100
	MOEIJSO-ACO	2.12	4.2	7.15	9.53	11.54

Table 3. Comparison of MOEIJSO-ACO with MLEACH

Performances	**Methods**	**Number of rounds**				
		500	1000	1500	2000	2500
Alive nodes	MLEACH [20]	29	9	3	1	0
	MOEIJSO-ACO	50	50	50	49	47
Dead nodes	MLEACH [20]	21	41	47	49	50
	MOEIJSO-ACO	0	0	0	1	3
Total Energy Consumption (%)	MLEACH [20]	42	81	96	99	100
	MOEIJSO-ACO	11.63	22.38	33.39	44.01	52.49

Table 4. Comparison of MOEIJSO-ACO with CEER

Performances	**Methods**	**Number of rounds**				
		2000	4000	6000	8000	10000
Alive nodes	CEER [21]	223	204	182	145	134
	MOEIJSO-ACO	225	225	225	188	153
Dead nodes	CEER [21]	2	21	43	80	91
	MOEIJSO-ACO	0	0	0	37	72
Total Energy Consumption (%)	CEER [21]	23	38	49	63	71
	MOEIJSO-ACO	17.33	29.07	36.12	58.19	65.14

Table 5. Life expectancy comparison of MOEIJSO-ACO

Life expectancy	400 nodes		50 nodes		225 nodes	
	MCR-UWSN [19]	MOEIJSO-ACO	MLEACH [20]	MOEIJSO-ACO	CEER [21]	MOEIJSO-ACO
FND	852	182	110	1890	1700	6100
HND	1121	4185	600	4027	10200	11054
LND	1187	9089	1500	4851	12310	13008

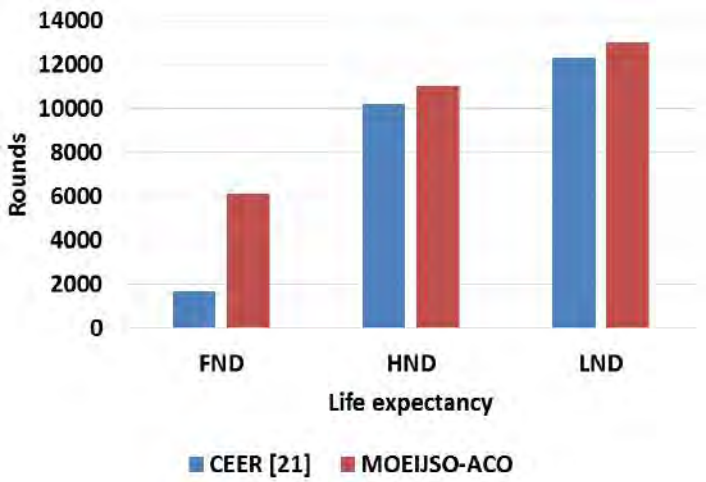


Figure. 10 Life expectancy comparison of MOEIJSO-ACO with CEER

5. Conclusion

In this research, energy-efficient CH and route discovery are performed for improving the energy efficiency of UWSN. The MOEIJSO is used for selecting the optimal CHs from the normal sensors by optimizing it with residual energy, neighbor node distance, sink distance and CH balancing factor. Next, ACO-based routing is developed for discovering the optimal route from the transmitter CH to BS. The sensor's energy usage is decreased by balancing the clusters and identifying the route with a lesser transmission distance. The reduction in energy usage is used to increase the UWSN life expectancy. On the other hand, data delivery is enhanced by improving energy efficiency and avoiding node failure. The MOEIJSO-ACO achieved better performances than the MCR-UWSN, MLEACH and CEER. The alive nodes of the MOEIJSO-ACO are 399 for 800 rounds, which is high when compared to the MCR-UWSN. In the future, modern optimization can be used to enhance UWSN performances.

Conflicts of interest

The authors declare no conflict of interest.

Author contributions

For this research work all authors' have equally contributed in Conceptualization, methodology, validation, resources, writing—original draft preparation, writing—review and editing.

References

[1] S. Karim, F. K. Shaikh, B. S. Chowdhry, Z. Mehmood, U. Tariq, R. A. Naqvi, and A. Ahmed, "GCORP: Geographic and cooperative opportunistic routing protocol for underwater sensor networks", *IEEE Access*, Vol. 9, pp. 27650-27667, 2021.

[2] M. Ismail, M. Islam, I. Ahmad, F. A. Khan, A. B. Qazi, Z. H. Khan, Z. Wadud, and M. A. Rakhami, "Reliable path selection and opportunistic routing protocol for underwater wireless sensor networks", *IEEE Access*, Vol. 8, pp. 100346-100364, 2020.

[3] S. Chinnasamy, J. Naveen, P. J. A. Alphonse, C. Dhasarathan, and G. Sambasivam, "Energy-Aware Multilevel Clustering Scheme for Underwater Wireless Sensor Networks", *IEEE Access*, Vol. 10, pp. 55868-55875, 2022.

[4] L. Alsalman and E. Alotaibi, "A Balanced Routing Protocol Based on Machine Learning for Underwater Sensor Networks", *IEEE Access*, Vol. 9, pp. 152082-152097, 2021.

[5] T. R. Chenthil and P. J. Jayarin, "An energy-efficient distributed node clustering routing protocol with mobility pattern support for underwater wireless sensor networks", *Wireless Networks*, Vol. 28, No. 8, pp. 3367-3390, 2022.

[6] Z. Liu, X. Jin, Y. Yang, K. Ma, and X. Guan, "Energy-Efficient Guiding-Network-Based Routing for Underwater Wireless Sensor Networks", *IEEE Internet of Things Journal*, Vol. 9, No. 21, pp. 21702-21711, 2022.

[7] K. Hao, Y. Ding, C. Li, B. Wang, Y. Liu, X. Du, and C. Q. Wang, "An energy-efficient routing void repair method based on an autonomous underwater vehicle for UWSNs", *IEEE Sensors Journal*, Vol. 21, No. 4, pp. 5502-5511, 2020.

[8] J. Wu, X. Sun, J. Wu, and G. Han, "Routing strategy of reducing energy consumption for underwater data collection", *Intelligent and Converged Networks*, Vol. 2, No. 3, pp. 163-176, 2021.

[9] T. Hussain, Z. U. Rehman, A. Iqbal, K. Saeed, and I. Ali, "Two hop verification for avoiding void hole in underwater wireless sensor network using SM-AHH-VBF and AVH-AHH-VBF routing protocols", *Transactions on Emerging Telecommunications Technologies*, Vol. 31, No. 8, p. e3992, 2020.

[10] H. Alasarpanahi, V. Ayatollahitafti, and A. Gandomi, "Energy-efficient void avoidance geographic routing protocol for underwater sensor networks", *International Journal of Communication Systems*, Vol. 33, No. 6, p. e4218, 2020.

[11] K. Latif, N. Javaid, I. Ullah, Z. Kaleem, Z. A. Malik, and L. D. Nguyen, "DIEER: Delay-Intolerant Energy-Efficient Routing with Sink Mobility in Underwater Wireless Sensor Networks", *Sensors*, Vol. 20, No. 12, p. 3467, 2020.

[12] M. Zhang and W. Cai, "Energy-Efficient Depth Based Probabilistic Routing Within 2-Hop Neighborhood for Underwater Sensor Networks", *IEEE Sensors Letters*, Vol. 4, No. 6, pp. 1-4, 2020.

[13] B. Priyalakshmi and S. Murugaveni, "Emperor Penguin Optimized Q Learning Method for Energy Efficient Opportunistic Routing in Underwater WSN", *Wireless Personal Communications*, 2022.

[14] A. Chaaf, M. S. A. Muthanna, A. Muthanna, S. Alhelaly, I. A. Elgendy, A. M. Iliyasu, and A. A. A. E. Latif, "Energy-efficient relay-based void hole prevention and repair in clustered multi-AUV underwater wireless sensor network", *Security and Communication Networks*, Vol. 2021, p. 9969605, 2021.

[15] S. Karim, F. K. Shaikh, K. Aurangzeb, B. S. Chowdhry, and M. Alhussein, "Anchor Nodes Assisted Cluster-Based Routing Protocol for Reliable Data Transfer in Underwater Wireless Sensor Networks", *IEEE Access*, Vol. 9, pp. 36730-36747, 2021.

[16] T. R. Chenthil and P. J. Jayarin, "An Energy-Aware Multilayer Clustering-Based Butterfly Optimization Routing for Underwater Wireless Sensor Networks", *Wireless Personal Communications*, Vol. 122, No. 4, pp. 3105-3125, 2022.

[17] K. Bhattacharjya, S. Alam, and D. De, "CUWSN: energy efficient routing protocol selection for cluster based underwater wireless sensor network", *Microsystem Technologies*, Vol. 28, No. 2, pp. 543-559, 2022.

[18] M. Faheem, R. A. Butt, B. Raza, H. Alquhayz, M. W. Ashraf, S. Raza, and M. A. B. Ngadi, "FFRP: dynamic firefly mating optimization inspired energy efficient routing protocol for internet of underwater wireless sensor networks", *IEEE Access*, Vol. 8, pp. 39587-39604, 2020.

[19] N. Subramani, P. Mohan, Y. Alotaibi, S. Alghamdi, and O. I. Khalaf, "An Efficient Metaheuristic-Based Clustering with Routing Protocol for Underwater Wireless Sensor Networks", *Sensors*, Vol. 22, No. 2, p. 415, 2022.

[20] H. H. Rizvi, S. A. Khan, and R. N. Enam, "Clustering Base Energy Efficient Mechanism for an Underwater Wireless Sensor Network", *Wireless Personal Communications*, Vol. 124, No. 4, pp. 3725-3741, 2022.

[21] I. Ahmad, I. T. Rahman, A. Zeb, I. Khan, M. T. B. Othman, and H. Hamam, "Cooperative Energy-Efficient Routing Protocol for Underwater Wireless Sensor Networks", *Sensors*, Vol. 22, No. 18, p. 6945, 2022.

[22] D. D. Yuan, M. Li, H. Y. Li, C. J. Lin, and B. X. Ji, "Wind power prediction method: Support vector regression optimized by improved jellyfish search algorithm", *Energies*, Vol. 15, No. 17, p. 6404, 2022.

SSRG International Journal of Electrical and Electronics Engineering
ISSN: 2348 – 8379 / https://doi.org/10.14445/23488379/IJEEE-V9I9P106

Volume 9 Issue 9, 47-53, September 2022

Original Article

Evaluating the Statistical Stability of POSIT Arithmetic and IEEE 754 Float to Accelerate Data for Detection of Breast Cancer

H. S. Laxmisagar[1], M. C. Hanumantharaju[2]

[1,2]*Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bengaluru, India.*

[1]*Corresponding Author : sagar8.hs@bmsit.in*

Received: 25 July 2022 Revised: 12 September 2022 Accepted: 19 September 2022 Published: 30 September 2022

Abstract *- In recent years, the researcher has focused on the hardware implementation of Floating Point Units (FPUs), which have a huge area and energy footprint. Due to their greater accuracy, speed, and simpler hardware design, Posit Arithmetic Units (PAUs) are proposed to replace IEEE 754-2008 compliant FPUs. It is important to improve existing floating-point IP cores for field-programmable gate array (FPGA) based applications. In comparison with other floating-point formats, posits number representation offers greater dynamic range and numerical accuracy. Various researchers attempted to implement a support vector machine (SVM) using hardware implemented on FPGA platforms to achieve high performance at lower power consumption and cost. As a result, the algorithm is unsuitable for embedded real-time applications. Therefore, SVM linear classifier is implemented on hardware, decreasing the latency and executing the task in real-time. This paper proposes an SVM linear classifier with pipeline architecture for fast processing in Verilog HDL using a single-precision IEEE standard 754 number format with 32-bit representation. The POSIT algorithm design is done with 24-bit representation using a software approach to determine the accuracy of prediction for the detection of Breast Cancer. The accuracy rate is computed both using software and hardware for performance evaluation. The pipelined SVM architecture is designed using Verilog HDL and synthesized using the Vivado simulation tool. The design is configured to the Xilinx KC705 Kintex-7 evaluation board for implementation.*

Keywords *- Breast Cancer, FPGA, IEEE 754 format, SVM, Vivado tool, Verilog HDL.*

1. Introduction

Breast cancer seriously risks women's lives and health [1]. The illness is treatable in the early stages, but it is not recognized until later, which is the primary reason for the death of so many women worldwide. Breast cancer detection at an early stage is critical for successful treatment and lowering the mortality rate. In the United States, 12.15 percent of women [2] will develop this cancer during their lifetime. For the last two decades, machine-learning researchers have been improving classifier effectiveness. As a result, machine-learning research has resulted in a new generation of cutting-edge supervised machine learning [29] classification algorithms, such as SVMs.

SVMs are considered a very good supervised learning algorithm that offers cutting-edge accuracy at the cost of high computational complications [4]. On the other hand, the standard versions of the SVM algorithm are very time-consuming and computationally intensive, presenting a challenge to investigate other hardware architectures than the CPU. It motivates the SVM to be realized on hardware platforms to attain high-performance computation at low-power consumption and cost.

Posit has various benefits over floating point numbers, including greater accuracy, a wide dynamic range, the capacity to provide results that are bitwise equal across platforms, simpler hardware, and the capacity to manage exceptions more effectively [5]. Compared to IEEE float floating point units, posit processing units require less circuitry. It is more efficient to employ low-precision posits than inexact computing methods that attempt to compensate for decreased quality of answers [6]. A wide range of applications is accelerated and performed with a substantial performance by FPGAs, proving that they are superior to other comparable platforms, such as graphics processing units (GPUs) and general-purpose processors [25]. As a result, FPGAs have become increasingly popular for realizing and accelerating SVM on hardware for low-power embedded system applications.

Section 2 describes the literature review related to the identification of Breast cancer, SVM, IEEE 754 float, POSIT arithmetic, and FPGAs, which are necessary for understanding the suggested method for the identification of cervical cancer. Section 3 discusses the system architecture implementation with floating point multipliers and adder trees in more detail. Section 4 presents the proposed design simulation, synthesis, and implementation

in more detail. Finally, a brief conclusion and future implementation are presented in section 5.

2. Related Work

In real-time scenarios, Deep Neural Networks (DNNs) are increasingly constrained by real-time constraints, so information representation must be re-evaluated. Encoding information in a way that can be processed rapidly and represented in a hardware-friendly way is a very challenging endeavor. In recent studies, posit formats have been shown to satisfy real-time constraints better than IEEE 754 standards for floating point representations of real numbers. According to Cococcioni et al. [8], Posit-based DNNs can be activated in the same way as 32-bit floats with 16-bit down to 10 bits. Despite their lower performance, 8-bit Posits could be a good substitute for 32-bit floats due to their high speed and low storage properties.

According to Van Dam et al. [9], there are approximately 160 MPOPS for multiplication, 250 MPOPS for addition, and 180 MPOPS for accumulations. As a result, intermediate results can be rounded off without affecting decimal accuracy. For the same number of bits, posit arithmetic gives higher decimal accuracy than the IEEE floating point format. Selvathi et al. [10] proposed integrating a computerized diagnostic system for breast cancer identification onto FPGA using Artificial Neural Network (ANN). The features of the WBCD are used to train and test the Multilayer Perceptron Neural Network (MLPNN). Baez et al. [11] proposed using Xilinx SDSoCTM to optimize the HLS methodology (Software-Defined SoC (System on chip)). The multiclass SVM classifier algorithm is designed using C programming and implemented on ZC7020 (ZedBoard) and ZC7045 (ZC706) devices.

Mohammadi et al. [12] proposed hardware implementation of SVM with parallel processing using Stochastic Gradient Descent for implementation onto FPGA. Fiolhais et al. [26] proposed a fused dot product processor with exact operand arithmetic that can output an incomplete result per cycle. The processor employs a long interval accumulator with full fixed-point precision to achieve full accuracy. Bassoli et al. [14] investigated FPGA implementations of IEEE 754 floating-point adder and multiplier for fast data processing using a conventional approach and customizations to save the chip area. The binary parallel multiplier and adder, which is a digit-serial multiplier, were proposed by Louca et al. [15]. One of the most significant limitations of existing embedded FPGAs is their major hardware development effort and time-to-market.

Table 1. Summary of different methods for breast cancer detection with limitations for the WBCD dataset

Author	Method	Application	Hardware	Acc. Rate	Area, Timing, and Power	Limitations
T Chandrasekaran et al., 2021 [16]	ANN based machine learning classifier	Breast Cancer Detection	Hardware-Software Co-design (CMOS-IC)	ANN=96.9%	-------	The limited dataset has been used for training. The ANN requires a huge volume of data for proper training.
Selvathi, D et al. 2018 [17]	ANN implementation using FPGA	Breast Cancer Classification	FPGA Virtex 5 board using XILINX ISE tool	90.83%	-------	limitation is optimization is not done, and no on-board Implementation
Batista et al. 2020 [18]	Pseudo-logarithmic number representation	Data classification	FPGA	-------	--------	Accuracy rate not computed Timing optimization not done Expensive multipliers with simpler adders
González-Díaz_Conti et al. 2022 [19]	ANN	Breast Cancer Detection	Xilinx Vivado 2016.1 FPGA Device: Zynq-7 ZC702	-------	LUTs:2267 Power:1.752 W Timing: ------	Accuracy rate not computed Timing optimization is not done

Numerous researchers [20]–[23] have contributed their work to implement the SVM machine learning algorithm, which is becoming increasingly popular in research due to the promising opportunities it provides in this field. Kim et al. [28] presented an IEEE 754 single precision multiplication method to cut down on the number of logic blocks. The suggested work estimates only one of the operands and iteratively compensates for the calculation error to attain limited errors in respective applications.

Most authors proposed CAD system development for the detection of Breast cancer, and very few papers report FPGA implementation for the detection of various types of cancer. As per the literature review, no paper is found on implementing an efficient SVM classification system for the identification of Breast cancer onto FPGA. Most existing implementations other than Breast cancers are not realized on modern FPGAs board, as evident from the literature review.

3. Proposed Architecture for Faster Data Classification

3.1. System Architecture

The SVM linear classifier's proposed FPGA design is shown in Figure 3. The architecture's adder tree constructs the classification function f(X) with the bias 'b' and polynomials such as patient data and weight value. Once testing data is passed into the FPGA board, the trained model uses it to perform a classifying procedure. The results are compared using the decision function to obtain either benign or malignant.

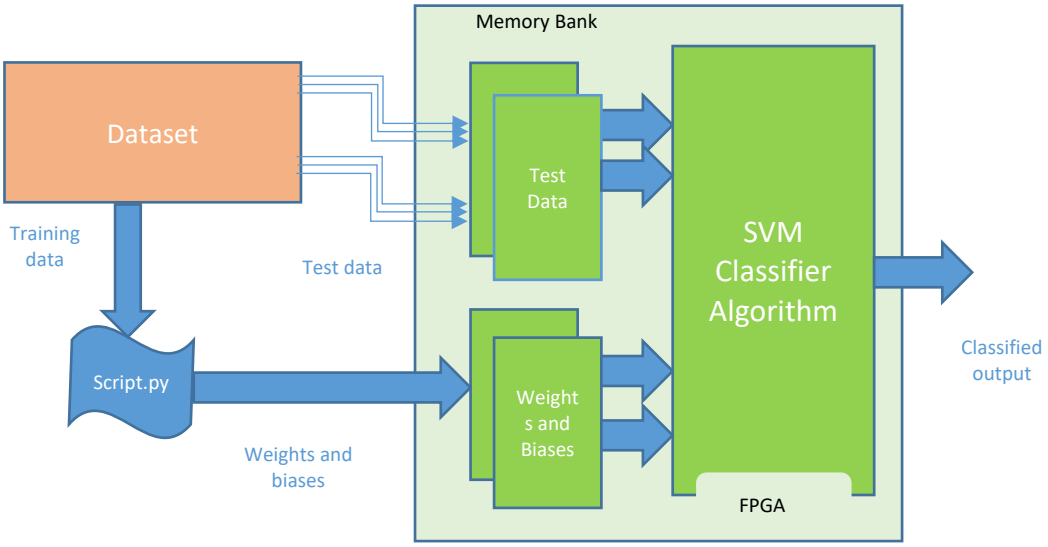


Fig. 1 System Architecture for Classification

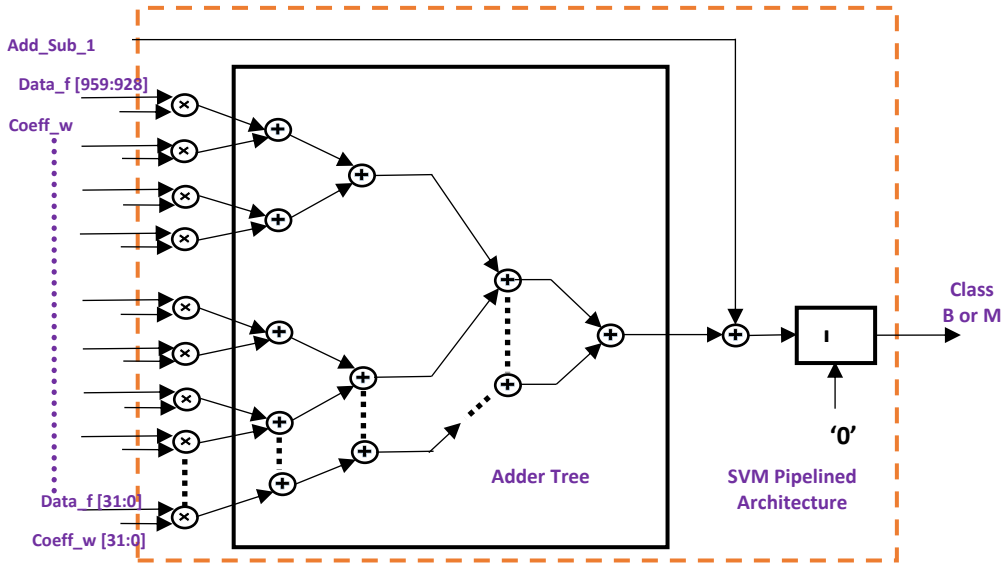


Fig. 2 SVM Linear Classifier Architecture

3.2. Design of BRAM Controller

The block RAM controller's design is essential to load breast cancer patient data and computed weight values using python script corresponding to each feature in Wisconsin Breast Cancer Dataset. The loaded data can be used further as input by designing a linear SVM classifier to create a training model to classify two data classes.

3.3. Computing for Linear SVM Classifier

32-bit handles the system's computations signed floating-point multipliers and adders with a single sign bit, a 23-bit mantissa, and an 8-bit exponent. The floating-point math packages are based on the VHDL 1164 numeric_std package and use the package's signed and unsigned arithmetic. Because the numeric_std package is well supported by simulation and synthesis tools, they are extremely efficient.

4. Experimental Results and Discussions

This study seeks to determine the most useful features in identifying malignant or normal cancer. The main objective is to discover whether breast cancer belongs to malignant or benign. Here, the standard WBCD dataset is analyzed for 569 patients with 30 best features to categorize between two data classes for the detection of Breast Cancer.

4.1. Simulation Output Waveform Analysis of Top Module of SVM Linear Classifier

The pipelined SVM architecture is designed using Verilog HDL, then synthesized by Vivado (ver. 2017.4) simulation tool. The design is implemented through translating, mapping, placing and routing, and generating a bit stream. Ultimately, it is configured to the Xilinx KC705 Kintex-7 evaluation board. Below are the simulation results for the single precision IEEE standard 754-based SVM linear classifier algorithm.

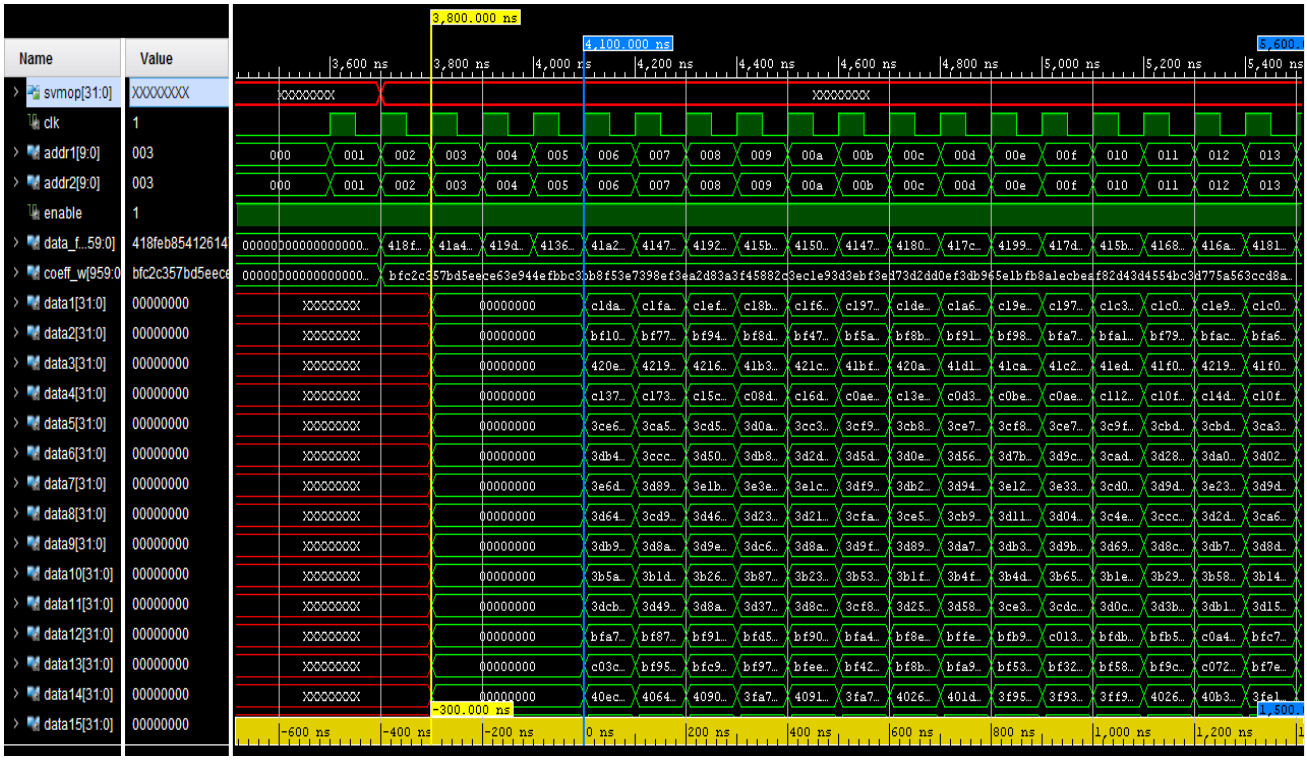


Fig. 3 The Output Waveform for Top Module of SVM Classifier: Depicts Data Loaded into BRAM at 3700 ns

The Simulation result of the SVM linear classifier depicts data loaded into BRAM for training to the positive edge of the clock, as shown in Figure 3. There are 30 features in the WBCD dataset, and each feature is represented with a 32-bit floating point number. So total bandwidth required to accommodate each patient's data is 960 bits wide. The simulation window shown in Figure 4 depicts pipelined adder tree operation with different stages of operation.

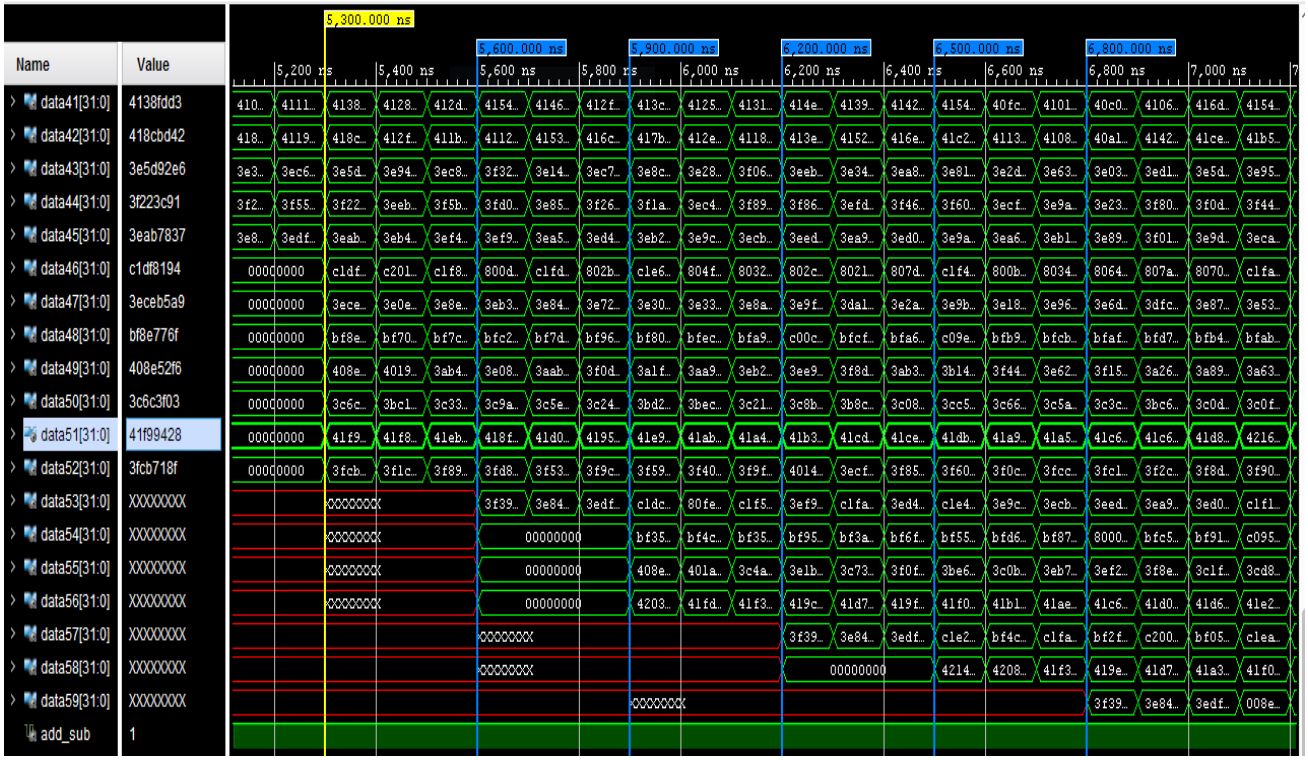


Fig. 4 Simulation Waveform of Adder Tree Pipelined Operation in SVM Linear Classifier at Different Stages

4.2. Logic Synthesis of top module SVM classifier system

After synthesis and implementation, it becomes very easy to check and analyse the different metrics such as area, power, and timing. The implementation result is more accurate in evaluating the performance analysis of the system. The schematic view of the top module SVM classifier after synthesis is shown in Fig. 5.

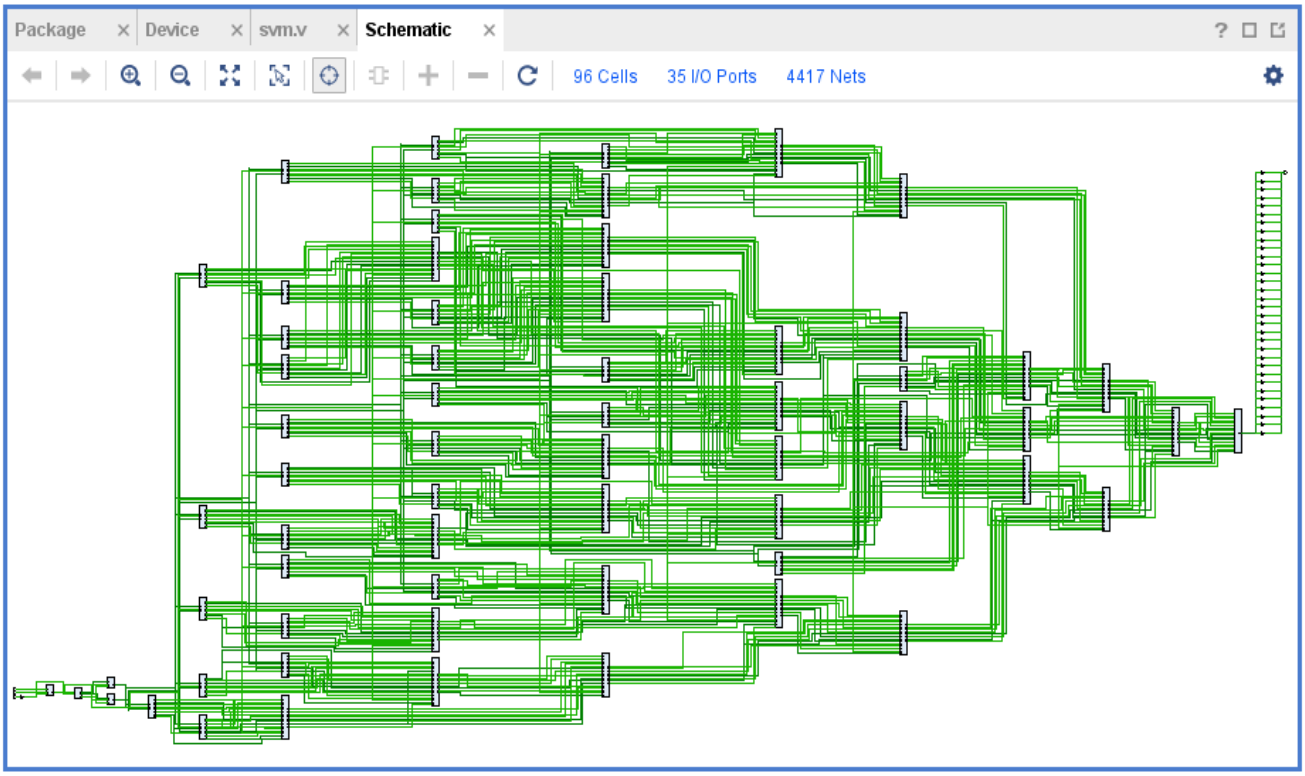


Fig. 5 Schematic View of Top Module SVM Classifier after Synthesis

4.3. Analysis of SVM output

The obtained hexadecimal number needs to be converted into a real number for comparison to find out the classified output, whether it belongs to 'Benign' or 'Malignant.' If the SVM output is positive or equal to '0', then the decision is taken w.r.t 'Malignant.' Similarly, if the SVM output is negative or less than '0', then a decision is taken w.r.t 'Benign.' It results in more inaccuracy than only positive, negative, and '0' for classification. To overcome this problem, a confusion matrix needs to be incorporated to increase the percentage accuracy by considering true positive or negative and false positive or negative. The accurate classification rate of the proposed system is 91%. The classification accuracy can further be increased by using an exact dot product accumulator with additional hardware cost.

There is a slight improvement in accuracy percentage obtained using Posit 24-bit format to IEEE 754 floating point algorithm implementation for detecting breast cancer using the WBCD dataset. The accuracy of precision for detection of breast cancer is obtained after running software code in Java for 24-bit posit format. The accuracy obtained is about 0.9122 using the Posit 24-bit format for IEEE 754 single precision 32-bit representation. The obtained result indicates that using the Posit 24-bit format, it is possible to obtain a better accuracy of IEEE 754 single precision 32-bit representation.

Table 2. Summary of POSIT arithmetic and IEEE 754 float

Format	Representation	Accuracy
IEEE 754 float	32 bits	91%
POSIT arithmetic	24 bits	91.22%

The summary of performance evaluation of POSIT arithmetic and IEEE 754 float is presented in table 2.

5. Conclusion

The proposed parallel implementation of the SVM algorithm for detecting breast cancer using the Vivado simulation tool as a training system is presented. The major objective of this solution is to reach a high data processing rate to satisfy the needs of computationally demanding applications. As a result, all possible multiplications and additions were parallelized. Finally, based on interpretations using the synthesis results, it can be concluded that implementing this technique in hardware would result in significant performance enhancements over the existing methods. The simulation and synthesis results exhibit that the SVM linear classification system is more effective in fast data classification. According to experimental results with a maximum frequency of 100 MHz in linear classification, the classification accuracy of about 91% in linear one has been achieved. The accuracy obtained is about 0.9122 using the Posit 24-bit format with respect to IEEE 754 single precision 32-bit representation. The obtained result indicates that using the Posit 24-bit format, it is possible to obtain a better accuracy of IEEE 754 single precision 32-bit representation.

The FPGA implementation of the SVM linear classifier with an exact dot product accumulator improves the percentage of accuracy related to state-of-the-art. It provides more flexibility to alter the structure in the future.

Conflicts of Interest

"Laxmisagar H.S and Hanumantharaju M.C declare that there is no conflict of interest regarding the publication of this paper."

References

[1] J. Ferlay et al., "Cancer Statistics for the Year 2020: An Overview," *Int. J. Cancer*, vol. 149, no. 4, pp. 778–789, 2021. Doi: https://doi.org/10.1002/ijc.33588.

[2] P. Mathur et al., "Cancer Statistics, 2020: Report from National Cancer Registry Programme, India," *JCO Glob. Oncol.*, no. 6, pp. 1063–1075, 2020. Doi: 10.1200/GO.20.00122.

[3] Dr.A.Nitesh and J.Deveesh, "Analysis and Optimization of a Floating Point Representation of a Complex Numbers using FPGA," *SSRG International Journal of VLSI & Signal Processing*, vol. 3, no. 2, pp. 17-21, 2016. *Crossref*, https://doi.org/10.14445/23942584/IJVSP-V3I2P104

[4] M. Papadonikolakis and C.S. Bouganis, "A Novel FPGA-Based SVM Classifier," in *International Conference on Field-Programmable Technology*, pp. 283–286, 2010. Doi: 10.1109/FPT.2010.5681485.

[5] J. Gustafson, "Posit Arithmetic," *Math. Noteb. Descr. Posit Number Syst.*, vol. 30, 2017.

[6] J. L. Gustafson and I. T. Yonemoto, "Beating Floating Point at its Own Game: Posit Arithmetic," *Supercomput. Front. Innov.*, vol. 4, no. 2, pp. 71–86, 2017.

[7] Chiranjeevi G N, Dr Subhash Kulkarni, "Enhanced MAC Controller Design for 2D Convolution Image Processing on FPGA," *SSRG International Journal of Engineering Trends and Technology*, vol. 69, no. 9, pp. 51-55, 2021. *Crossref*, https://doi.org/10.14445/22315381/IJETT-V69I9P207

[8] M. Cococcioni, F. Rossi, E. Ruffaldi, and S. Saponara, "Fast Approximations of Activation Functions in Deep Neural Networks when using Posit Arithmetic," *Sensors*, vol. 20, no. 5, 2020. Doi: 10.3390/s20051515.

[9] L. Van Dam, J. Peltenburg, Z. Al-Ars, and H. P. Hofstee, "An Accelerator for Posit Arithmetic Targeting Posit Level 1 Blas Routines and Pair-Hmm," in *Proceedings of the Conference for Next Generation Arithmetic*, vol. 2019, pp. 1–10, 2019.

[10] D. Selvathi and R. D. Nayagam, "FPGA Implementation of on-Chip ANN for Breast Cancer Diagnosis," *Intell. Decis. Technol.* vol. 10, no. 4, pp. 341–352, 2016.

[11] A. Baez et al., "High-Level Synthesis of Multiclass SVM using Code Refactoring to Classify Brain Cancer from Hyperspectral Images," *Electronics*, vol. 8, no. 12, 2019. Doi: 10.3390/electronics8121494.

[12] M. Mohammadi, A. Krishna, N. S., and S. K. Nandy, "A Hardware Architecture for Radial Basis Function Neural Network Classifier," *IEEE Trans. Parallel Distrib. Syst.*, vol. 29, no. 3, pp. 481–495, 2018. Doi: 10.1109/TPDS.2017.2768366.

[13] Ramya R, Madhura R, "FPGA Implementation of Optimized BIST Architecture for Testing of Logic Circuits," *SSRG International Journal of VLSI & Signal Processing*, vol. 7, no. 2, pp. 36-42, 2020. *Crossref*, https://doi.org/10.14445/23942584/IJVSP-V7I2P106

[14] M. Bassoli, V. Bianchi, and I. De Munari, "A Simulink Model-Based Design of a Floating-Point Pipelined Accumulator with HDL Coder Compatibility for FPGA Implementation," in *Applications in Electronics Pervading Industry, Environment and Society*, pp. 163–171, 2020.

[15] Louca, Cook, and Johnson, "Implementation of IEEE Single Precision Floating Point Addition and Multiplication on FPGAs," in *Proceedings IEEE Symposium on FPGAs for Custom Computing Machines*, pp. 107–116, 1996. Doi: 10.1109/FPGA.1996.564761.

[16] S. T. Chandrasekaran, R. Hua, I. Banerjee, and A. Sanyal, "A Fully-Integrated Analog Machine Learning Classifier for Breast Cancer Classification," *Electronics*, vol. 9, no. 3, 2020. Doi: 10.3390/electronics9030515.

[17] D. Selvathi and A. Aarthy Poornila, "Deep Learning Techniques for Breast Cancer Detection Using Medical Image Analysis," in *Biologically Rationalized Computing Techniques for Image Processing Applications*, J. Hemanth and B. V. Emilia, Eds. Cham: Springer International Publishing, pp. 159–186, 2018. Doi: 10.1007/978-3-319-61316-1_8.

[18] G. C. Batista, D. L. Oliveira, O. Saotome, and W. L. S. Silva, "A Low-Power Asynchronous Hardware Implementation of a Novel SVM Classifier, with an Application in a Speech Recognition System," *Microelectronics J.*, vol. 105, p. 104907, 2020. Doi: https://doi.org/10.1016/j.mejo.2020.104907.

[19] G. González-Díaz_Conti et al., "Hardware-Based Activation Function-Core for Neural Network Implementations," *Electronics*, vol. 11, no. 1, 2022. Doi: 10.3390/electronics11010014.

[20] H. M. Hussain, K. Benkrid, and H. Seker, "The Role of FPGAs as High Performance Computing Solution to Bioinformatics and Computational Biology Data," *AIHLS*, vol. 102, 2013.

[21] A. Eklund, P. Dufort, D. Forsberg, and S. M. LaConte, "Medical Image Processing on the GPU – Past, Present and Future," *Med. Image Anal.*, vol. 17, no. 8, pp. 1073–1094, 2013. Doi: https://doi.org/10.1016/j.media.2013.05.008.

[22] A. HajiRassouliha, A. J. Taberner, M. P. Nash, and P. M. F. Nielsen, "Suitability of Recent Hardware Accelerators DSPS, FPGAS, and GPUS for Computer Vision and Image Processing Algorithms," *Signal Process. Image Commun.*, vol. 68, pp. 101–119, 2018. Doi: https://doi.org/10.1016/j.image.2018.07.007.

[23] S. Asano, T. Maruyama, and Y. Yamaguchi, "Performance comparison of FPGA, GPU and CPU in Image Processing," in *International Conference on Field Programmable Logic and Applications*, pp. 126–131, 2009. Doi: 10.1109/FPL.2009.5272532.

[24] Mehboob Hasan Ahmed, Rutuja Jagtap, Roopal Pantode and Prof. S. S. Phule, "An FPGA Chip Identification Generator using Configurable Ring Oscillator," *SSRG International Journal of Electronics and Communication Engineering*, vol. 3, no. 4, pp. 10-14, 2016. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V3I4P103

[25] S. Afifi, H. GholamHosseini, and R. Sinha, "A System on Chip for Melanoma Detection using FPGA-Based SVM Classifier," *Microprocess. Microsyst.*, vol. 65, pp. 57–68, 2019. Doi: https://doi.org/10.1016/j.micpro.2018.12.005.

[26] L. Fiolhais and H. Neto, "An Efficient Exact Fused Dot Product Processor in FPGA," in *28th International Conference on Field Programmable Logic and Applications FPL*, pp. 327–3273, 2018. Doi: 10.1109/FPL.2018.00062.

[27] E.Ganesan and V.Sakthivel, "A Novel FPGA Design with Hybrid LUT / MUX Architecture," *SSRG International Journal of Electronics and Communication Engineering*, vol. 3, no. 11, pp. 6-8, 2016. *Crossref*, https://doi.org/10.14445/23488549/IJECE-V3I11P112

[28] S. Kim and R. A. Rutenbar, "An Area-Efficient Iterative Single-Precision Floating-Point Multiplier Architecture for FPGA," in *Proceedings of the on Great Lakes Symposium on VLSI*, pp. 87–92, 2019. Doi: 10.1145/3299874.3318002.

[29] W. Chen, H. R. Pourghasemi, A. Kornejady, and N. Zhang, "Landslide Spatial Modeling: Introducing New Ensembles of ANN, Maxent, and SVM Machine Learning Techniques," *Geoderma*, vol. 305, pp. 314–327, 2017. Doi: https://doi.org/10.1016/j.geoderma.2017.06.020.

AI and IoT-based Intelligent Health Care & Sanitation

Back Book Details

IoT Based Website for Identification of Acute Lymphoblastic Leukemia using DL

Author(s): R. Ambika*, S. Thejaswini, N. Ramesh Babu, Tariq Hussain Sheikh, Nagaraj Bhat and Zafaryab Rasool

Pp: 1-15 (15)

DOI: 10.2174/9789815136531123010003

Abstract

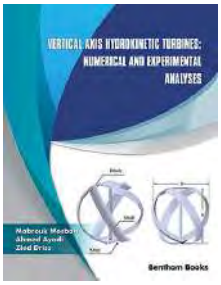
A form of cancer known as leukemia, attacks the body's blood cells and bone marrow. This happens when cancer cells multiply rapidly in the bone marrow. The uploaded image is analyzed by the website, and if leukemia is present, the user is notified-a collection of pictures depicting leukemia as well as healthy bones and blood. Once collected from Kaggle, the data is preprocessed using methods like image scaling and enhancement. To create a Deep Learning (DL) model, we use the VGG-16 model. The processed data is used to “train” the model until optimal results are achieved. A Hypertext Markup Language (HTML) based website is built to showcase the model. Using a DL model, this website returns a response indicating whether or not the user's uploaded photograph shows signs of leukemia. The primary aim of this site is to lessen the likelihood that cancer cells may multiply while the patient waits for test results or is otherwise unaware of their condition. Waiting for results after a leukemia test can cause further stress and even other health problems, even if the person is found to be leukemia-free. This problem can be fixed if this website is used as a screening tool for leukemia.

Keywords: Deep Learning, Image Augmentation, Image Processing, Leukemia, VGG-16, Web Development.

Recent Patents on Space Technology

View More »

Related Books



Recent Advances in Renewable Energy

View More »

Chapter 9
An Exhaustive Analysis of Energy Harvesting Absorbers and Battery Charging Systems for the Internet of Things

C. Padmavathy
Sri Ramakrishna Engineering College, India

Dankan Gowda V.
https://orcid.org/0000-0003-0724-0333
B.M.S. Institute of Technology and Management, India

Vaishali Narendra Agme
Bharati Vidyapeeth College of Engineering, India

Algubelly Yashwanth Reddy
https://orcid.org/0000-0001-6422-1943
Sree Dattha Group of Institutions, India

D. Palanikkumar
Dr. N.G.P. Institute of Technology, India

ABSTRACT

Nearly all application fields are paying increased attention to the internet of things (IoT). Nearly 20 billion devices are now linked to the internet. With several applications ranging from smart buildings and smart cities to smart devices, IoT has progressed over the last few decades. As a result, the quantity of sensors, sensor nodes, and gateways has increased, making these battery-powered devices power-hungry. It will be a laborious operation to change the battery in remote monitoring applications for these smart sensors or nodes. By gathering RF energy from the environment and converting it to DC power, RF energy harvesting is a cost-effective method of extending the lifespan of wireless sensor networks (WSNs). A brand-new, IoT-based smart universal charger is suggested in this chapter for charging multichemistry

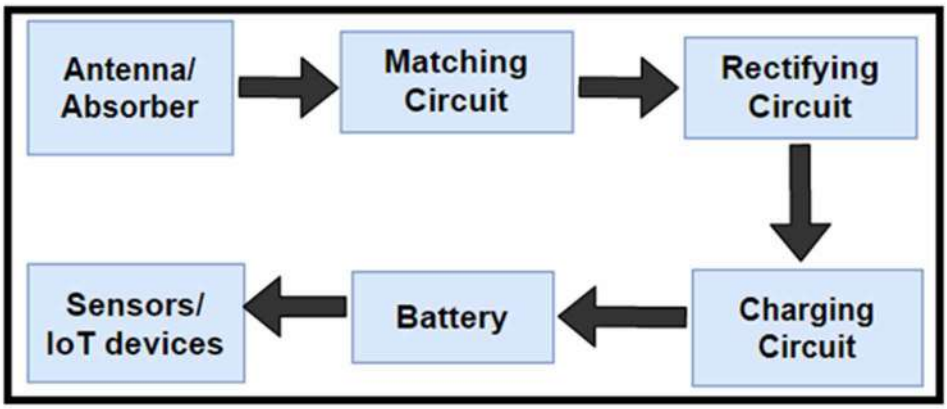
DOI: 10.4018/978-1-6684-4974-5.ch009

batteries. The suggested charger has an advantage over traditional chargers since it can charge both already installed batteries and any future batteries.

INTRODUCTION

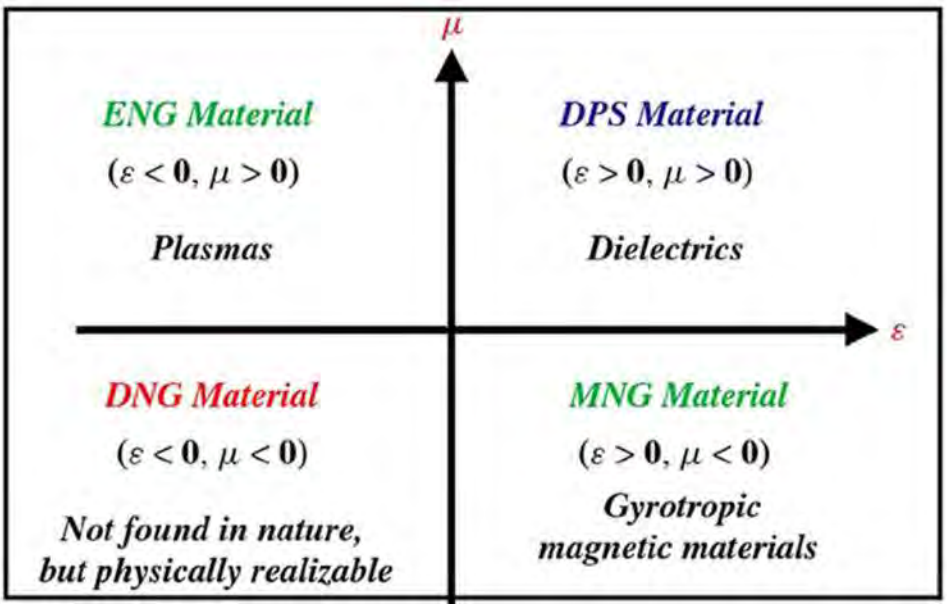
In the last few years, the demand for the power has increased enormously. Researchers are looking for alternate energy sources such as solar, wind, thermal, vibration and radio frequency (RF) for energy harvesting. Inexhaustible RF energy sources with zero harmful emissions can pave way for harvesting energy unlimitedly for powering low power sensors and microcontrollers. Internet of Things (IoT) enables to connect enormous sensor nodes for remote monitoring applications to the internet. RF Energy harvesting reduces the complexity of IoT nodes by avoiding power

Figure 1. RF energy harvesting System (RFEH)



circuits, reduces the overall cost and improves the overall efficiency of the system. RF Energy harvesting or Green Energy harvesting is an excellent technique which allows size reduction in comparison with other harvesting systems such as photonic cells or wind turbines, where miniaturization really matters a lot in portable devices which uses battery (Almoneef, T. S. 2014). RF energy harvesting plays a key role in providing a sustainable energy source towards the future of wearable electronics too. Further, RF energy harvesting is reliable, portable environmental friendly and also cost effective. These advantages have attracted the researchers to unveil novel research in this field. A typical EHS consists of Antenna/absorber, matching circuit,

Figure 2. Material classification

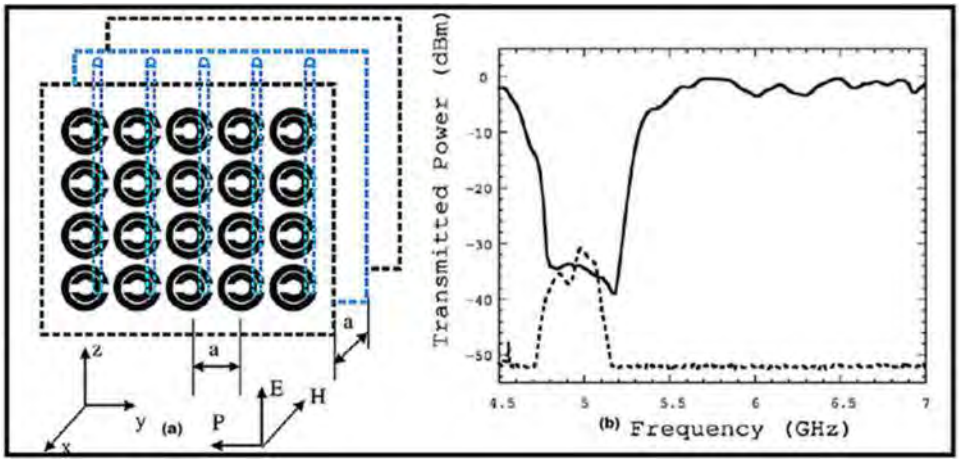


rectifier circuit, charging circuit, battery and low power electronic devices such as sensors, microcontrollers etc. as shown in Figure.1.

Generally materials are classified on the basis of permittivity (Ɛ) and permeability (μ) in the four quadrants as shown in Figure.2.

Double positive materials (DPS) with (Ɛ > 0; μ > 0) i.e. naturally existing dielectric materials lie in quadrant-I. In such materials, the direction of poynting vector and wave number is same. These are known as right handed materials (Almutairi, A. F. 2019). The refractive index is positive in these materials. Epsilon negative materials (ENG) with (Ɛ < 0; μ > 0) i.e. plasmas and metals at optical frequencies belong to quadrant-II. The wave travels in evanescent mode in these materials. Double negative materials (DNG) with (Ɛ < 0; μ < 0) are referred as metamaterials and lie in quadrant-III. The refractive index is negative in these materials and are also known as Left handed materials (LHM). The effective (Ɛ & μ) of a material can be tailored to design an artificial structure for any specific application. The real part of the permittivity Ɛ′(ω) defines the dielectric constant of the material and the imaginary part Ɛ″(ω) defines the attenuation provided by the material (Aminov, P. 2018). The real part of permeability μ′(ω) defines the energy stored in the magnetic field and the imaginary part μ″(ω) gives the measure of energy dissipation in the material for the incident magnetic field. by combining the structural elements of electric and magnetic response together in a unit cell a material with effective

Figure 3. Negative Index metamaterial (a) Design of an SRR array (b) Transmitted power vs Frequency



metamaterial properties can be designed. The electric and magnetic responses can be generated by coupling between metamaterial elements. Moreover, the flexibility of engineering 'μ' and 'Ɛ' values match the impedance of a designed metamaterial absorber with free space to reduce the reflection (Amiri, M. 2018). Subsequently, it leads to the perfect absorption at the desired frequency. To synthesize negative permeability was demonstrated as shown in Figure.3(a). The transmission response for SRR is represented by bold line and the response for the combined structure (thin wire and SRR) is shown by dashed curve in Figure.3(b). For the bulk structure, if SRR is alone present, there is no transmission around 5 GHz. But when SRR is combined with thin wire, energy is transmitted. The effect of the combined negative permeability and negative permittivity derives the energy transmission.

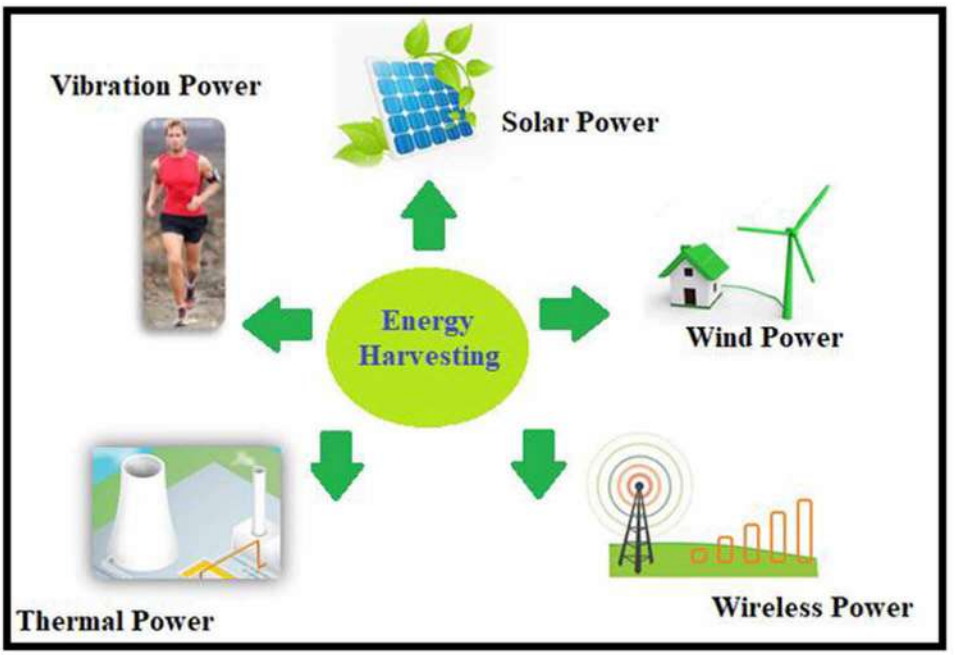
To understand the wave absorption theory, the interaction of the incident wave with the boundary surface is studied. The ways through which an incident wave can interact with a boundary surface includes: reflection, transmission, absorption, scattering and the excitation of surface waves. The surface electromagnetic waves (SEWs) (Ritchie, 1957; Watts, Liu and Padilla, 2012) are generated due to the incident wave (Ang, K. H., Chong, G. 2015). The effect of these SEWs can be studied based upon their propagation length (Lp). Though the propagation length (Lp) gives insufficient information about an incident wave coupled with the boundary surface, it defines the distance over which the intensity of SEWs decays exponentially (Arrawatia, M. 2016). In order to describe the coupling of the incident wave with the boundary surface, an equation for Figure of merit is described by Gil et al. (2007). The scattering of the waves at the interface is considered to be negligible if the

operating wavelength is much larger than the surface roughness ($\lambda > Rs$) causing the surface waves to die out before the occurrence of re-scattering. Further, based on the Fresnel equation and transmission matrix approach, the basic absorption theory for a magneto dielectric medium backed by a conductive ground metal was discussed by Watts, Liu and Padilla (2012). In this, all possible ways an incident wave can interact with boundary surface is considered. The metamaterial absorber consists of a dielectric medium which is supported by an opaque metallic surface as the ground plane, therefore there is zero transmission (S21) and the maximum absorption solely depends upon the minimum reflection. Most of the researchers focussed on metamaterial absorber design with single or dual band absorption characteristics (Bagmanci, M. 2018). Very few works have been published on absorbers with wide band absorption characteristics. Hence, the present research work was envisioned to explore the metamaterial based RF energy harvesting for ultra-wide frequency band (UWB) with maximum absorptivity. The existing multichemistry batteries such as lead-acid, Li-ion, etc., requires different chargers for charging (Bakır, M., Karaaslan. 2017). However, universal chargers can afford solution to some extent. Moreover, the existing chargers will not support, when a new type of battery is invented. Hence, these sort of issues necessitate an alternative way to address these problems. As an alternative, a novel IoT based smart universal charger has been proposed in this chapter for charging multichemistry batteries.

LITERATURE SURVEY

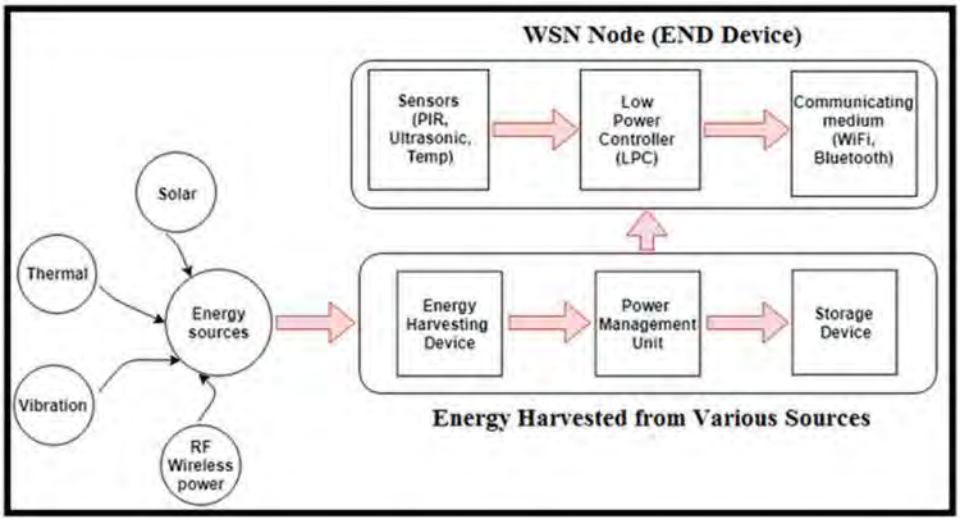
WSNs are one of the prominent areas of research which provides a background framework for Internet of Things (IoT). The key components of WSNs are sensors, actuators, embedded controllers and connectivity devices through which data or information is communicated. One of the major challenge lies in the incessant power supply required for these enormous smart connected devices which are mainly battery operated. RF energy harvesting provides an encouraging and economically friendly solution to provide low power DC supply from the harvested RF energy sources such as Wi-Fi, GSM, Radio/TV broadcasting stations, radar, etc. Numerous research works are carried out in the development of RF Energy harvesting systems (RFEHs) with research challenges in the antenna design, bandwidth and DC conversion efficiency (Baqir, M. A., Ghasemi, M. 2019). An overview of the different RFEH systems along with the challenges in the antenna design, rectifier design, matching circuit design are reviewed to afford a general framework for designing RF energy harvesting systems. In addition, an IoT based universal battery charging system has been developed to replace the conventional batteries. Advancement in the technologies has brought about a demand for instant access to everything. Engineers are bound

Figure 4. Various Energy Harvesting sources in the wireless networks



to revamp the existing technologies such as IoT (Benayad, A., & Tellache. 2020), 5G technologies, artificial intelligence (AI), (Bently, W. F., & Heacock. 1996) and wearable electronics (Brown, W. C. 1980). These technologies involve huge measure of sensors which are deployed in the remote places, but these sensors require batteries to satisfy their power requirement. But replacing these batteries periodically from these ubiquitous sensors is quite a herculean task, which makes the environment more and more hazardous. Energy Harvesting (EH) or Green Energy (GE) could be an alternate or a feasible solution to this environmental challenge which involves extracting energy from the readily available renewable energy sources. Figure.4. illustrates the various energy harvesting sources for WSNs. The ambient energy from the various sources such as solar power, wind power, thermal power, vibration power and dedicated RF energy sources such as Wi-Fi, WLAN, Bluetooth, AM/FM etc. are converted into electrical energy. The right choice of energy harvesting helps in reducing the complexity of end devices (WSN/IoT nodes). This is done by avoiding the power circuits which reduces the overall cost and improves the overall efficiency of the devices/system. Figure.5. represents the architecture of a wireless sensor node (WSN). A typical WSN consists of energy harvesting device, power management unit and storage device. Energy harvesting device converts the incident ambient energy from various energy sources such as solar, wind, thermal,

Figure 5. Architecture of a WSN node



vibration, and RF wireless energy into DC voltage. The power management unit (PMU) is responsible for either storing the harvested energy or regulating the power to the WSN Node (End Device). The storage device conserves the harvested power for future use.

In the past, several reviews were reported by the researchers on different energy harvesting techniques. (Chen, T., Li, S.J. 2019) demonstrated energy harvesting from body heat (thermal energy) to power WSN node. (Chen, W. C. 2018) demonstrated piezoelectric techniques for energy harvesting. (Chen, X., Tomasz. 2017) focussed on small scale wind mills for powering low power WSN nodes. Though, battery has been the main source of power, replacing and discarding them remains a challenge for uninterrupted data and also hazardous to the environment (Dawar, P., Raghava. 2019). Therefore, GE or environment friendly technology is required to avoid the toxic chemicals and metals which pollute the environment. Energy harvesting was achieved through near field meth (Derbal, M. C., & MouradNedil. 2020). In the chapter, Far field method is being focussed more for RF energy harvesting.

NEED FOR IoT BASED UNIVERSAL CHARGERS

The rapid evolution of technology has yielded advances in introducing portable devices and gadgets with profound impacts on society. Promisingly, in every domain, new portable devices keep updating the market. Rechargeable batteries are one of

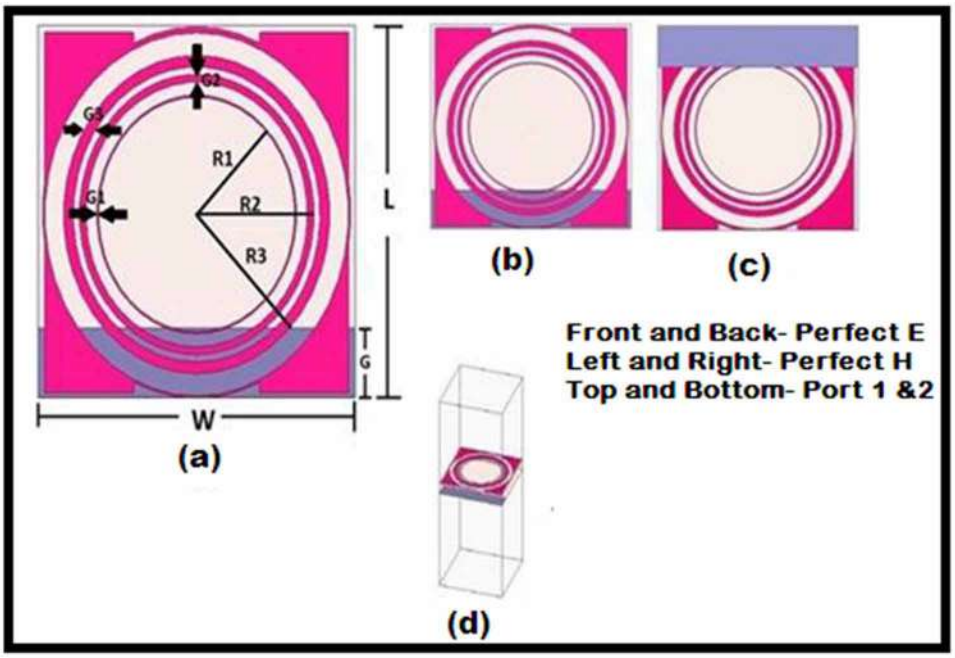
the major sources to operate these portable devices. Hence, development of new rechargeable batteries has become a major research area in low power electronics. Since its invention, researchers have concentrated mainly on increasing charge density, decreasing charging time, reducing size and minimizing the cost of the rechargeable battery (Devi, K. K. A., Hau, N. C. 2018). This has resulted in the development of new rechargeable batteries with different chemical composition. Currently, development of such batteries are in a greater pace, as it replaces the conventional batteries. Although different kinds of batteries utilize different charging algorithm, the major problem was every battery should follows its charging algorithm to get maximum life time, reduced charging time and no hazards. Secondly, different batteries have to use different chargers (Devi, K. K. A., Sadasivam. 2019). The need of universal chargers are much important which supports charging many kinds of batteries instead of buying separate chargers, reduces the overall size and cost. IoT refers to the interconnection of embedded devices through internet. IoT enables the devices to share and interact each other without manual intervention. The application of IoT has expanded to every field to facilitate real time monitoring by accessing remote database. Moreover, incorporation of IoT platform with the universal charger seems to be a novel alternative to add new battery profile from the server to the client, which in turn paves the way to charge existing battery profile as well as upcoming batteries in the near future. number of cells for NiCd, NiMH, Li-ion and Li-Polymer batteries (Ding, F., Dai, J. 2018). It had the advantage of charging Constant Current (CC), Constant Voltage (CV) with a wide range of currents and voltages. But the CMOS processor used was more expensive and also adds more complexity to the design. The design experiment was carried out using DU2004S3 a Nickel/Li-ion development system. But the limitations were manual selection of batteries and fixed battery profiles. The PWM signal generated by the microcontroller was able to feed the buck convertor circuit that controls the voltage and current on the output using a PI controller (Divakaran, S. K., Krishna. 2018). The charging of only two kinds of batteries and the unavailability of addition of new kind of battery are considered as the major disadvantage of this system. A patented work by (Dolgov, A., Zane. 2019) demonstrated a universal charger in which RFID was used to select the battery profile for charging. Based on the RFID tag information, the current and voltage required to charge the device was selected by charging selector. But, the major disadvantage of the charger was in the selection of only predefined battery profiles. This survey has attempted to review and critically assess the reported design techniques for RFEH in the literature and battery charging techniques for various multichemistry batteries. The design techniques for antenna/absorber was classified into four categories – Miniaturization, polarization, reconfigurability and harmonic rejection. Also the design considerations for matching circuits are classified as diode and Mosfet based rectifier circuits. The survey has given an overview that,

metamaterial based circular absorbers with diode based rectifier circuits can play a significant role in reducing the electrical size of the antenna and exhibits circular polarization with wide absorption characteristics and provide maximum efficiency. Similarly from the literatures it was observed that, the requirement of different chargers for charging multichemistry batteries such as Lead-acid, Lithium-ion etc. is a major concern. The need for universal chargers are very essential which can support charging many kinds of batteries instead of buying separate chargers. So the need of the hour is a novel IoT based technique for charging existing chargers and upcoming any upcoming batteries. The present work was aimed at designing wide band metamaterial absorbers. Under wideband configuration three absorber designs have been proposed. At first, a polarization insensitive dual band square based metamaterial was designed (Escala, O. A. 2019). The proposed ultrathin absorber size was 16mm x 16mm and and documented to provide high absorptivity of more than 90% in the desired bands. Though, the proposed absorber was found insensitive to various incident angles of electromagnetic wave, it affords only linear polarization with dual bands. Thereafter, a circular shaped metamaterial absorber with concentric circular rings was proposed with optimized design to make the unit cell size smaller, compared to the operating wavelength (Falade, O. P., Rehman. 2018). The proposed circular absorber with compact size of 10mm x 10mm resonated at three different frequency bands with narrow bandwidth. The proposed absorber was found to have a good broad band characteristic and provided FWHM (Full Width Half Maximum). The absorber was thin in thickness with an overall thickness of ~ $\lambda/25$ corresponding to the absorption frequency of 7.6 GHz. The simulations were also carried out to determine the absorption characteristics for different values of incident wave as shown in the Figure.4. The proposed absorber provided wide angle insensitivity i.e. 00, 300, 450, 900 for various oblique angles of incident wave.

TRIPLE BAND METAMETERIAL ABSORBER

Besides, a compact triple band metamaterial with narrow band absorption characteristics was proposed. The proposed metamaterial structure contained three concentric circular rings with notch structures at the top and bottom with thickness of $0.04\lambda 0$ for the attained frequency. The size of the proposed circular absorber was reduced to 16% with resonant frequency at three different bands (triple band). The proposed metamaterial absorber was designed with patch (copper) at the top layer, having three asymmetrical concentric circular rings with notched structures etched at the top and bottom. Figure.6(a) represents the geometry of the proposed absorber (Fante, R. L., & Mccormack. 1988). Substrate size of 10 mm x10 mm, ground width size of 2 mm printed on the FR4 dielectric material with the thickness 1.6 mm, loss

Figure 6. (a) Unit cell absorber (b) Top View (c) Bottom View (d) Trimetric view

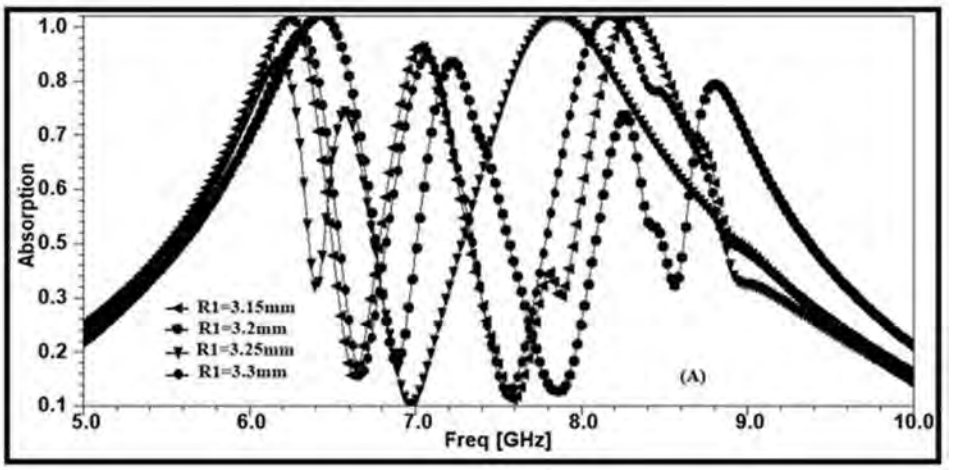


tangent 0.02 mm and permittivity 4.4 mm.The optimal radius of the three circular rings and the notched shape at the top and bottom was attained after the parametric analysis. Figure.6(a) illustrates the proposed absorber, Figure.6(b) and (c) represents the top view and the bottom view of the proposed absorber. Figure.6(d) represents the trimetric view of the proposed unit cell absorber.

Figure.7. represents the normalized absorption plot with various values for R1. When the radius R1 is 3.2 mm, two absorption peaks are obtained and their peak normalised values are 0.99 and 0.89. Similarly, when the value of R1 is 3.25 mm, similar two absorption peaks are obtained with normalised absorption values above 0.80. But when the value of R1 is 3.3mm, three absorption peaks with normalised absorption values of 0.96, 0.90 and 1 are obtained (Faruque, M. R. I. & Islam. 2018). The optimum R1 value is attained when the value of R1 is at 3.15 mm and hence the value of R1 is assigned at 3.15mm.

Figure.8. illustrates the normalized absorption plots for the various values of R2. From the figure, it is observed that when R2 value is 3.65 mm only two absorption peaks are attained with maximum absorption rates of more than 0.80. When the R2 value is 3.7 mm, three rates of 0.95, 0.90, and 1(unity) is observed. With 3.75 mm as the R2 value, the absorber exhibits the absorption rate of 1(unity), 0.90 and 0.96. But when R3 value is at 3.8 mm, the absorption peaks obtained are at 0.92, 1(unity)

Figure 7. Absorptivity plots optimizing R1



and 0.94. The absorption rates are almost similar to each other. From the parametric study, the value of R2 at 3.7 mm is assigned as the optimum value.

The parameters magnetic permeability (μ) and Electric permittivity (ε) are the parameters which determine the metamaterial behavior when exposed to a time varying electromagnetic field (Fowler, C., & Zhou, J. 2017). Based on the values of (μ) & (ε), metamaterials can be classified as DNG (both μ, ε <0), ENG (ε <0) and MNG (μ<0). To analyze the absorber's metamaterial characteristics, the values

Figure 8. Absorptivity plots optimizing R2

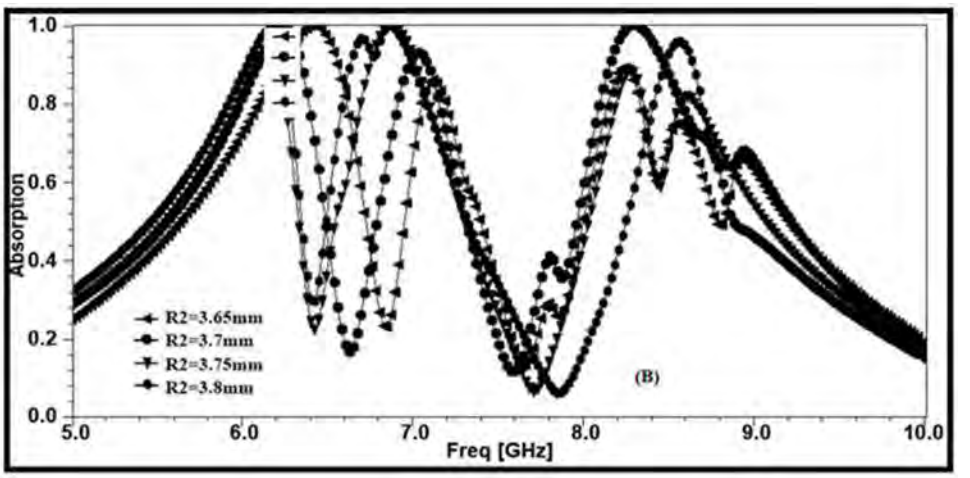
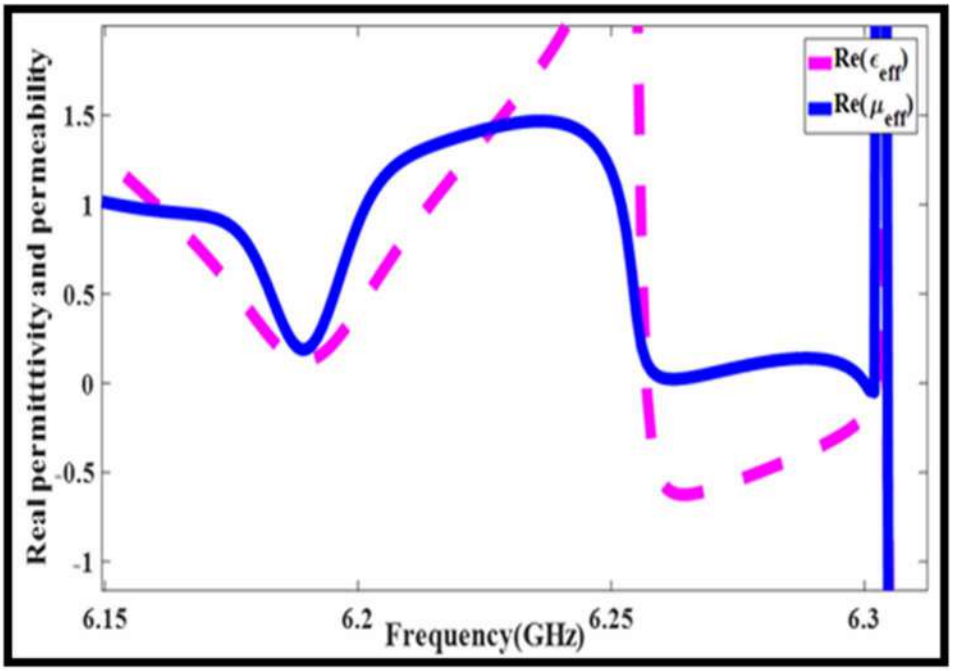


Figure 9. Simulated Real permittivity (Ɛ) and permeability (μ) plots at 6.3GHz frequency

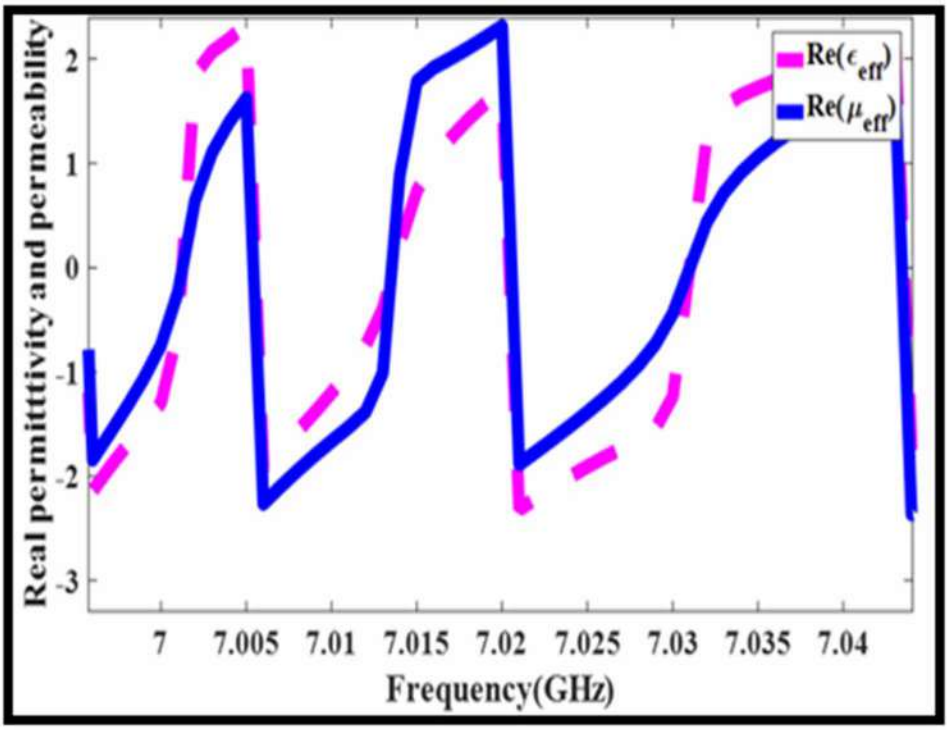


of μ and ε should be retrieved from the S parameters (S11 and S12). To retrieve the values of μ and ε the method proposed by Singh, Abegaonkar and Koul (2019).

It is obseved that at 6.3GHz both the permittivity and permeability values are negative with permeability value μ = -0.1 and the permittivity value Ɛ = -1.2. At 7 GHz both the permittivity and permeability values are in the negative region with permittivity Ɛ = -1.4 and the permeability μ = -0.4. At the peak frequency of 8.35 GHz the value of permittivity is around Ɛ= -0.2 and permeability is μ = +0.4. So it is concluded that at 6.3 GHz and 7 GHz, the material behaves as DNG metamaterial or a perfect metamaterial absorber and at 8.35 GHz it behaves as ENG metamaterial absorber (Ghosh, S., Bhattacharyya. 2018). From the above figures, it can be observed that in all the three frequencies (6.3GHz, 7GHz and 8.35GHz), the values of negative permittivity and the negative permeability lies in the positive region in their respective graphs (Gil, I., Martin, F. 2017). This clearly indicates that the imaginary part of permeability or permittivity always remains positive, irrespective to the metamaterial type such as MNG or ENG or DNG.

When Z (ω) is matched with free space impedance (Z0) then the material is said be a perfect absorber. Figure.9. and Figure.10. shows the plot of the refractive index

Figure 10. Simulated Real permittivity (Ɛ) and permeability. (μ) plots at 7 GHz frequency



and impedance at frequencies 6.3GHz, 7GHz and 8.35GHz respectively. The values of the refractive index (n) is negative at the frequencies 6.3GHz (n = - 0.1) and & 7GHz (n = -1.1), proving it to be DNG absorber and the value of n is positive (n = 0.1) at 8.35GHz proving it be ENG absorber.

The parameters μ and ε constitutes the metamaterial characteristics. These values were retrieved and analysed to realise the type of metamaterial absorber (MTM) (Gozel, M. A., Kahriman. 2018). The results prove that at frequencies 6.3 GHz and 7 GHz, the proposed absorber behaves as a perfect absorber with more than 90% absorption(Guha, D., Biswas. 2019). To achieves desirable size reduction compared to the existing dimensional structures (Guo, L., & Lu, Q. 2018). Though the work done by Singh (Hashemi, S. S., Sawan. 2018) has exhibited the same dimensions, the maximum absorption rate was comparatively low compared to the proposed work.

SINGLE BAND ABSORBER AND RECTIFIER

This section discusses a compact single band circular metamaterial absorber to harvest DC energy from very low ambient microwave signals in particular for indoor environments such as office, buildings etc. The most widely used 2.45 GHz frequency band which is used for Wi-Fi, Bluetooth, Zigbee, and RFID is preferred for RF energy harvesting. The RF energy harvested from the designed absorber is converted to dc power using a tuned single stage and a three stage voltage rectifier circuit. The absorber and the rectifying circuit are analysed for maximum dc output voltage and maximum efficiency. The L network is a simple circuit with few components compared to T or II network. The advantage of using L network is using few components for matching but with narrow band characteristics. The simulation is carried out with advanced design system (ADS) software with smart Smith Chart features for tuning the L matching circuits. For matching the input impedance of rectifier circuit (Zin) with normalised impedance Zo (50Ω), the value of Zin at the desired frequency (2.45 GHz) was computed. From the results, it was observed that the designed rectifier circuit with L-matching circuit, attained the maximum efficiency of 87.2% at the input power of 10dBm reported in the Figure 5.20. The maximum output volatage attained at -10 dBm, 0 dbm and 10 dBm were 1.07V 0.45V and 1.78V, respectively. The rectifier voltage doubler circuit designed for single stage gave maxium voltage around 1.8V and an efficiency of 87% at an input power of 10dBm. Further, doubler was increased to three stages. In each stage, two schottky diodes HSMS2850, filter capacitor and a load resistor of 100kΩ were used. The schottky diode is preferred due to its small resistance (Rs = 25Ω) and small barrier capacitance (Cp = 0.18pF) with high cut off frequency and high conversion efficiency. The Smith Chart utility provides synthesis of matching networks and enables impedance matching. Using Smith chart utility, L matching network was tuned and the values of L1 & L2 were obtained. The desired frequency 2.45GHz was attained after Zo is matched with Zin. The 3 stage rectifier voltage doubler was analysed for output voltage and efficiency. The output voltage was analysed for various input power levels. A peak voltage of 7.9V was attained around 10dBm with a maximum efficiency of 65% with a 100K load. From the efficiency plots, it is observed that despite improvement in the output voltage, the efficiency is reduced due to the dissipation losses in the schottky diode (Hassan, N., Hisham. 2018). In continuation, simulation of rectifier voltage doubler circuit for single stage and three stage was analyzed. Schotty diode HSMS 2850 was selected due to its high cut off frequency and high conversion efficiency. The single stage RFEH system showed a conversion efficiency of 87% at the input power of 10 dBm with an output voltage of 1.78V. As the stages are increased, the efficiency is reduced to 65%. But the output voltage is increased to 7.9V at 10dBm. This may be due to the dissipation of power due to the non linear elements in the circuit. The

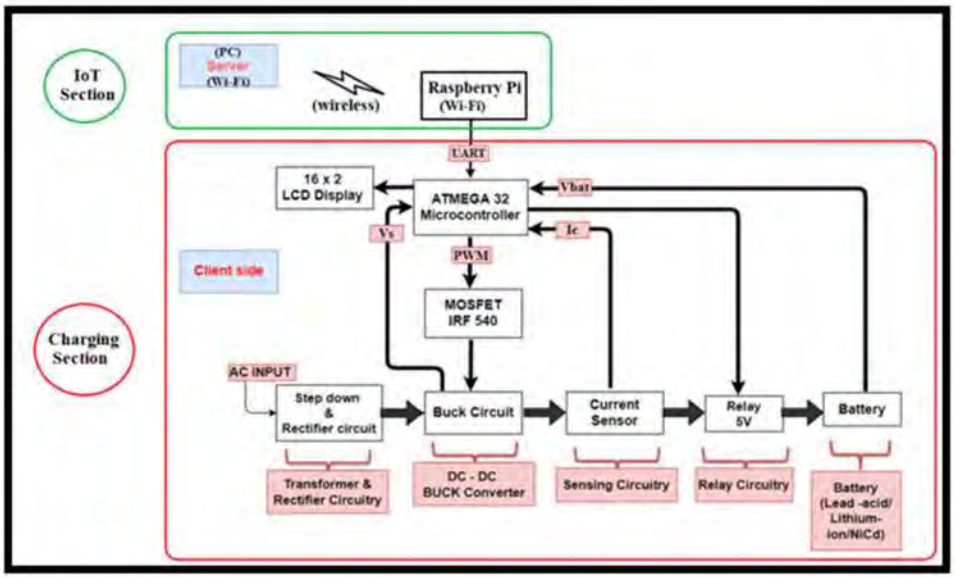
attained results are compared with the previous works and are tabulated. The results concludes that the RFEH system is a suitable candidate to harvest energy for indoor applications at the 2.45GHz frequency band.

IoT BASED SMART CHARGING SYSTEM

The intelligent charger is user friendly, which allows the user to select any of the existing battery profiles and also facilitates to update new battey profiles from the server. The proposed design consists of two sections: IoT section and charging Section. The IoT section consists of Raspberry pi controller as client and PC as server, both are remotely connected via internet. Both server and client contains the databases of various battery profiles (Hong, H., Cai, X., Shi, X. 2018). The charging section consists of a microcontroller based buck converter which monitors the nominal battery voltage, current and temperature. It also supplies the required output voltage and current for charging the batteries using pulse width modulation (PWM) technique. The charging profile of NiCd and NiMH batteries are same. The battery follows a Constant Current (CC) method for charging and the charging voltage is in the range of 1.4V-1.6V per cell. From the cell voltage of 0.075V to 0.625V, the battery follows a slow charging at 0.1C. From 0.625V onwards, it follows fast charging at 1C. Vpeak is represented as the maximum voltage at which the cell attains 1.4V. The maximum safe voltage (Vmax) attained by cell is 1.2V. It is imperative to determine the detection of full charge in the cell as overcharging could damage the cell. The detection of full charge was determined by the dropping of voltage from Vmax in a duration time of 1 minute. The proposed design is the combination of both hardware and software and is divided into two sections: IoT section and the charging section. In the IoT section, personal computer (PC) is server and it is connected to the client (raspberry pi) through Wi-Fi module. In the charging section, the modules used are Atmega 32 microcontroller, MOSFET IRF 540, DC – DC Buck converter, current sensor, relay module and the battery which is to be charged. Figure.11. illustrates proposed system.

The proposed block diagram has two main sections as shown in the Figure.11. In the IoT section, PC is the server and the client is the raspberry pi. The raspberry pi is a mini computer running on Linux based platform which facilitates both hardware and software applications. Both the server and client have the databases of different battery profiles. The server and client are connected remotely to the internet. The python script helps to connect, raspberry pi(client) to the servers database. Therefore, whenever a new battery profile is available at the server, the client can download and update the new battery profile to its database. In the charger section, the main blocks used are Atmega32 microcontroller, buck converter and a current sensor.

Figure 11. proposed system



The microcontroller communicates with the Raspberry pi through a serial UART interface. Also it generates the desired PWM voltage and current pulses by varying the duty cycle for CV and CC charging stages. Then the PWM pulses generated are fed to the buck converter circuit. The Buck converter circuit contains MOSFET IRF540, inductor and a capacitor. In the converter design, the design of L and C is very crucial. The voltage and the current values of the buck circuit has to be monitored continuously. So these values are fed back to the microcontroller. The current sensor ACS712 generates the equivalent analog voltage with respect to the current that is sensed. The voltage and current fed back is retained constantly by the PID controller for charging the battery. The PID controller continuously maintains the voltage and current constant in accordance to the set points of voltage and current. The battery charging is done in two modes i.e constant current and constant voltage, referred as charging algorithm. For charging a particular type of battery, the user enters the CC and CV values. Before the start of battery charging, the microcontroller checks for the battery voltage. If the battery voltage is less than the full discharge voltage, then the battery is detected as faulty. If the voltage is greater than the full discharge voltage, then the controller follows the constant current algorithm and the current is set according to the CC value. The PID control is applied to the CC values to maintain the current constant throughout the CC mode. As the battery voltage reaches more than 80% of full charge or Nominal voltage, the microcontroller enters CV mode and follows a constant voltage algorithm and the voltage is set according to the CV

value. The PID control is applied to the CV values to maintain constant voltage during the CV mode. The voltage (V) and current (I) are constantly updated in the LCD display. As the battery attains the maximum voltage value and if the current drops below 4% of the rated current, then the charge cycle is stopped and updated in the LCD display. In the proposed system, addition of upcoming new battery profiles is possible along with the existing battery profiles. This has an added davantage over the existing chargers reported in the literatures. In this proposed work, an intelligent charging system is implemented which can charge different batteries simultaneously. The novelty of the proposed system is the ability to update new charging profiles from the server. The proposed system is simulated and tested for various battery profiles and the results are found to be reliable. The results confirm that the designed charger can be implemented with differerent multichemistry batteries ranging from low capacity to high capacity. Therefore, it can be realized as a universal charging system for all kind of portable home appliances.

CONCLUSION

In this chapter, a wide band RF energy harvesting metamaterial absorber was proposed to harvest energy in the frequency range of 3.1 GHz to 10.6 GHz. Adding up, a single stage and three stage rectifier voltage doubler circuit were investigated for a single band circular absorber at 2.45 GHz frequency. Moreover, a novel IoT based battery charger has been designed as an alternative to the conventional chargers. From the attained results, the designed absorbers are proved to be a better candidate for energy harvesting applications. A novel IoT based intelligent universal charger was proposed in this study. The designed charger consisted of two sections such as IoT section and charging section. The IoT section consisted of Raspberry pi controller as client and PC as server. The server contains the databases of battery profiles and remotely connected to the controller through Internet. The charging section consists of microcontroller based buck converter which monitors the nominal battery voltage, current and temperature. The required output voltage was provided to charge the battery using Proportional Integral Derivative (PID) algorithm through PWM technique. Unlike the conventional chargers that could charge only the existing batteries, the proposed charger was able to charge diverse multichemistry batteries. The uniqueness of this charger is recognized based on the updation of existing battery profile as well as the upcoming new battey profile from the server. In the future, flexible electronics will play a major role in applications such as wearables, ingestible smart pills etc. The designs of various absorbers in this thesis considered is FR4 material with dielectric constant of 4.4. Hence flexible

dielectric materials like polydimethylsiloxane (PDMS), paper and jeans cloth will be concentrated in the future work.

ACKNOWLEDGMENT

We would like to acknowledge the support given by BMS Institute of Technology and Management, Bangalore, India.

REFERENCES

Almoneef, T. S., & Ramahi, O. M. (2014). Can Split Ring Resonators be viable for electromagnetic energy harvesting? IEEE Antennas and Propagation Society, 424-425.

Almutairi, A. F., Islam, M. S., Samsuzzaman, M., Islam, M. T., Misran, N., & Islam, M. T. (2019). A Complementary Split Ring Resonator Based Metamaterial with Effective Medium Ratio for C-band Microwave Applications. *Results in Physics*, *15*, 102675. doi:10.1016/j.rinp.2019.102675

Aminov, P., Jai, P., & Agrawal. (2018). RF Energy Harvesting. *IEEE 64th Electronic Components and Technology Conference (ECTC)*, 1838-1841.

Amiri, M., Tofigh, F., Shariati, N., Lipman, J., & Abolhasan, M. (2020). Wide angle metamaterial absorber with highly insensitive absorption for TE & TM modes. *Scientific Reports*, *10*(1), 2. doi:10.103841598-020-70519-8 PMID:32788706

Ang, K. H., Chong, G., & Li, Y. (2015). PID control system analysis, design, and technology. *IEEE Transactions on Control Systems Technology*, *13*(4), 559–576.

Arrawatia, M., Baghini, M. S., & Kumar, G. (2016). RF energy harvesting system at 2.67 and 5.8GHz. *Asia-Pacific Microwave Conference*, 900-903.

Bagmanci, M., Karaaslan, M., Altintaş, O., Karadag, F., Tetik, E., & Bakir, M. (2018). Wideband metamaterial absorber based on CRRs with lumped elements for microwave energy harvesting. *The Journal of Microwave Power and Electromagnetic Energy*, *52*(1), 45–59. doi:10.1080/08327823.2017.1405471

Bakır, M., Karaaslan, M., Unal, E., Akgol, O., & Sabah, C. (2017). Microwave metamaterial absorber for sensing applications. *Opto-Electronics Review*, *25*(4), 318–325. doi:10.1016/j.opelre.2017.10.002

Baqir, M. A., Ghasemi, M., Choudhury, P. K., & Majlis, B. Y. (2019). Design and analysis of nanostructured subwavelength metamaterial absorber operating in the UV and visible spectral range. *Journal of Electromagnetic Waves and Applications*, *29*(18), 2408–2419. doi:10.1080/09205071.2015.1073124

Benayad, A., & Tellache, M. (2020). A compact energy harvesting multiband rectenna based on metamaterial complementary split ring resonator antenna and modified hybrid junction ring rectifier. *International Journal of RF and Microwave Computer-Aided Engineering*, *30*(2), e22031. doi:10.1002/mmce.22031

Bently, W. F., & Heacock, D. K. (1996). Battery management considerations for multichemistry systems. *IEEE Aerospace and Electronic Systems Magazine*, *11*(5), 23–26. doi:10.1109/62.494184

Brown, W. C. (1980). The history of the development of the rectenna. *Proc. SPS Microwave Systems Workshop at JSC-NASA*, 271-280.

Chen, T., Li, S. J., Cao, X. Y., Gao, J., & Guo, Z. X. (2019). Ultra-wideband and polarization-insensitive fractal perfect metamaterial absorber based on a three-dimensional fractal tree microstructure with multi-modes. *Applied Physics. A, Materials Science & Processing*, *125*(4), 232. doi:10.100700339-019-2536-6

Chen, W. C., Totachawattana, A., Fan, K., Ponsetto, J. L., Strikwerda, A. C., Zhang, X., Averitt, R. D., & Padilla, W. J. (2018). Single-layer terahertz metamaterials with bulk optical constants. *Physical Review. B*, *85*(3), 035112. doi:10.1103/PhysRevB.85.035112

Chen, X., & Tomasz, M. (2017). A Robust method to retrieve the constitutive effective parameters of metamaterials. *Physical Review. E*, *70*(1), 016608. doi:10.1103/PhysRevE.70.016608 PMID:15324190

Dawar, P., Raghava, N. S., & De, A. (2019). UWB Metamaterial-Loaded Antenna for C-Band Applications. *International Journal of Antennas and Propagation*, *2019*, 1–13. doi:10.1155/2019/6087039

Derbal, M. C., & Nedil, M. (2020). A High Gain Dual Band Rectenna for RF Energy Harvesting Applications. *Progress in Electromagnetics Research Letters*, *90*, 29–36. doi:10.2528/PIERL19122604

Devi, K. K. A., Hau, N. C., Chakrabarty, C. K., & Din, N. M. (2018). Design of Patch Antenna Using Metamaterial at GSM 1800 for RF Energy Scavenging. *IEEE Asia Pacific Conference on Wireless and Mobile*, 157-161.

Devi, K. K. A., Sadasivam, S., Din, N. M., & Chakrabarthy, C. K. (2019). Design of a 377 Ω Patch Antenna for Ambient RF Energy Harvesting at Downlink Frequency of GSM 900. *The 17th Asia Pacific Conference on Communications,* 492-495.

Ding, F., Dai, J., Chen, Y., Zhu, J., Jin, Y., & Bozhevolnyi, S. I. (2018). Broadband near-infrared metamaterial absorbers utilizing highly lossy metals. *Scientific Reports*, *6*(1), 39445. doi:10.1038rep39445 PMID:28000718

Divakaran, S. K., & Krishna, D. D. & Nasimuddin. (2018). RF energy harvesting systems: An overview and design issues. *International Journal of RF and Microwave Computer-Aided Engineering*, *29*, 1.

Dolgov, A., Zane, R., & Popovic, Z. (2019). Power Management System for Online Low Power RF Energy Harvesting Optimization. *IEEE Transactions on Circuits and Systems*, *57*(7), 1802–1811. doi:10.1109/TCSI.2009.2034891

Escala, O. A. (2019). *Study of the Efficiency of Rectifying Antenna Systems for Energy Harvesting* [Thesis]. UPC-Barcelona, Spain.

Falade, O. P., Rehman, M. U., Gao, Y., Chen, X., & Parini, C. G. (2018). Single Feed Stacked Patch Circular Polarized Antenna for Triple Band GPS Receivers. *IEEE Transactions on Antennas and Propagation*, *60*(10), 4479–4484. doi:10.1109/TAP.2012.2207354

Fante, R. L., & Mccormack, M. T. (1988). Reflection Properties of the Salisbury Screen. *IEEE Transactions on Antennas and Propagation*, *36*(10), 1443–1454. doi:10.1109/8.8632

Faruque, M. R. I., & Islam, T. (2018). Novel triangular metamaterial design for electromagnetic absorption reduction in human head. *Progress in Electromagnetics Research*, *141*, 463–478. doi:10.2528/PIER13050603

Fowler, C., & Zhou, J. (2017). *A Highly Efficient Polarization-Independent Metamaterial-Based RF Energy-Harvesting Rectenna for Low-Power Applications.* arXiv, 1705.07717.

Ghosh, S., Bhattacharyya, S., Chaurasiya, D., & Srivastava, K. V. (2018). An Ultra-wideband Ultrathin Metamaterial Absorber Based on Circular Split Rings. *IEEE Antennas and Wireless Propagation Letters*, *14*, 1172–1175. doi:10.1109/LAWP.2015.2396302

Gil, I., Martin, F., Rottenberg, X., & De Raedt, W. (2017). Tunable stop-band filter at Q-band based on RF-MEMS metamaterials. *Electronics Letters*, *43*(21), 1153–1154. doi:10.1049/el:20072164

Gozel, M. A., Kahriman, M., & Kasar, O. (2018). Design of an efficiency-enhanced Greinacher rectifier operating in the GSM 1800 band by using rat-race coupler for RF energy harvesting applications. *International Journal of RF and Microwave Computer-Aided Engineering*, *21*(1), e21621. doi:10.1002/mmce.21621

Guha, D., Biswas, S., & Antar, Y. M. M. (2019). *Microstrip and Printed Antennas: New Trends, Techniques and Applications*. John Wiley & Sons.

Guo, L., & Lu, Q. (2018). Potentials of piezoelectric and thermoelectric technologies for harvesting energy from pavements. *Renewable & Sustainable Energy Reviews*, *72*, 761–773. doi:10.1016/j.rser.2017.01.090

Hashemi, S. S., Sawan, M., & Savaria, Y. (2018). A high-efficiency low-voltage CMOS rectifier for harvesting energy in implantable devices. *IEEE Transactions on Biomedical Circuits and Systems*, *6*, 326335. PMID:23853177

Hassan, N., Hisham, A. B., Fareq, A. M. M., Abidin, A. M. Z., Bakar, H., Noor, A. M. S., & Khairy, I. M. (2018). Radio frequency (RF) energy harvesting using metamaterial structure for antenna/rectenna communication network: A review. *Journal of Theoretical and Applied Information Technology*, *96*(6), 1538–1550.

Hong, H., Cai, X., Shi, X., & Zhu, X. (2018). Demonstration of a highly efficient RF energy harvester for Wi-Fi signals. *International Conference on Microwave and Millimetre Wave Technology (ICMMT)*, 1–4.

An Enhanced Method for Running Embedded Applications in a Power-Efficient Manner

N. M. G. Kumar (/affiliate/n-mg-kumar/439645/), Ayaz Ahmad, Dankan Gowda V., S. Lokesh, Kirti Rahul Rahul Kadam

Source Title: Energy Systems Design for Low-Power Computing (/book/energy-systems-design-low-power/293050)

Pages: 21
DOI: 10.4018/978-1-6684-4974-5.ch013

Abstract

Many modern items that are in widespread use have embedded systems. Due of embedded processing's ability to provide complex functions and a rich user experience, it has grown commonplace in many types of electronic products during the last 20 years. Power consumption in embedded systems is regarded as a crucial design criterion among other factors like area, testability, and safety. Low power consumption has therefore become a crucial consideration in the design of embedded microprocessors. The proposed new method takes into consideration both the spatial and temporal locality of the accessed data. In the chapter, the new cache replacement is combined with an efficient cache partitioning method to improve the cache hit rate. In this work, a new modification is proposed for the instruction set design to be used in custom made processors.

Chapter Preview

Top

Introduction

An Embedded system is a combination of an electronic and a computer system. It is a computer or processor-based system designed to perform a dedicated sequence of the task to control or operate a large system (mechanical or electrical) with real-time constraints. It is a system used to perform its functions without human intervention completely or partially. It is

also designed to accomplish a particular task in an efficient method. Mostly, embedded systems are used in operation where timing is very important (Aaron Lindsay and Binoy Ravindran. 2018). In the modern-day embedded systems applications, battery operated devices play an important role. The applications that run on those embedded devices rely more and more on powerful processors and are capable of running real-time applications. Because of the increase in number as well as the complexity of such applications, a significant amount of work has focused on the minimisation of power and energy consumed by the embedded processors (Chen Yang and Leibo Liu. 2018). Apart from giving a result in real-time, these embedded devices are to be designed to satisfy thermal limits as well as battery life limits, thereby directing the research towards low power and low energy enhancements. Along with performance, ease of use, and other such design metrics, power consumption is also a design metric for the present-day embedded systems. Embedded applications are demanding more processing power along with the ever-increasing need for more memory. It calls for active research to satisfy this multiple objective design challenge to develop an application for multicore embedded systems and their memory managing capabilities. The power consumption happening in memory by static and dynamic leakages is calling for a need to work on relevant solutions (Chenjie Yu and Peter Petrov. 2019). The power leakage in the bus and memory also create other side effects on the processor like thermal effects.

The cache contention happening in multicore processors can be addressed by the cache partitioning scheme to maximise the cache space utilisation. It may also increase the execution timing of the task. This work proposes a priority-based cache partitioning approach among the cores to improve both the cache performance and deadline avoidance. An embedded systems benchmark is used to select the set of applications from workloads to work on this

cache partitioning problem (Dan, A and Towsley. D. 2020).

This work also presents a method to reduce the power consumed in the instruction fetching data bus during the execution of the instruction. The instruction code fetched from memory is modified so that the bus is loaded less, and therefore the switching capacitance associated with the bus is reduced(Daniel Sanchez and Christos Kozyrakis. 2021). This results in decreasing the power consumed in the data bus during the instruction fetch cycle. To explore a working cache model for a multi-core processor with configurable cache details and to incorporate run time cache optimisation for power minimisation.

The research problem of improving cache design in the context of power and energy optimisation is important because, in today's processors, the cache memory has a share of over 30% of the total processor's power. Cache memory, which is accessed faster and also with a good hit rate, is considered to be more power-efficient (Dimitris Kaseridis and Jeffrey Stuechelix. 2009). By selecting a better cache replacement algorithm which can be implemented either as hardware maintained structure or as a program (Dongwoo Lee and Kiyoung Choi. 2019). This method reduces the bus power of the processor during program execution. Power consumption in embedded systems is one of the important issues. Since a substantial amount of power consumption in a processor happens inside the cache and memory operations. this chapter proposes a technique on that domain to reduce the power consumption in embedded systems

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=319982&ptid=293050&ctid=15&t=Table of Contents&isxn=9781668449745)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=319983&ptid=293050&ctid=15&t=Detailed Table of Contents&isxn=9781668449745)

Preface — View Full PDF (/pdf.aspx?tid=319984&ptid=293050&ctid=15&t=Preface&isxn=9781668449745)

Rathishchandra Ramachandra Gatti, Chandra Singh, Srividya P., Sandeep Bhat

Chapter 1

Low-Power Methodologies and Strategies in VLSI Circuits (/chapter/low-power-methodologies-and-strategies-in-vlsi-circuits/319986) (pages 1-16)

Preethi, Sapna R., Mohammed Mujeer Ulla

Preview Chapter (/viewtitlesample.aspx?id=319986&ptid=293050&t=Low-Power Methodologies and Strategies in VLSI Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 2

SOI Technology in Designing Low-Power VLSI Circuits (/chapter/soi-technology-in-designing-low-power-vlsi-circuits/319987) (pages 17-28)

Srividya P.

Preview Chapter (/viewtitlesample.aspx?id=319987&ptid=293050&t=SOI Technology in Designing Low-Power VLSI Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 3

Low-Power High-Speed Eight-Bit Universal Shift Register Design Using Clock Gating Technique (/chapter/low-power-high-speed-eight-bit-universal-shift-register-design-using-clock-gating-technique/319988) (pages 29-43)

Preeti Sahu

Preview Chapter (/viewtitlesample.aspx?id=319988&ptid=293050&t=Low-Power High-Speed Eight-Bit Universal Shift Register Design Using Clock Gating Technique&isxn=9781668449745) **$37.50** Add to Cart

Chapter 4

Dynamic Body Bias: A Transistor-Level Technique for the Design of Low-Voltage CMOS Analog Circuits (/chapter/dynamic-body-bias/319989) (pages 44-66)

Vandana Niranjan

Preview Chapter (/viewtitlesample.aspx?id=319989&ptid=293050&t=Dynamic Body Bias: A Transistor-Level Technique for the Design of Low-Voltage CMOS Analog Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 5

A Detailed Study on Single Electron Transistors in Nano Device Technologies (/chapter/a-detailed-study-on-single-electron-transistors-in-nano-device-technologies/319990) (pages 67-99)

S. Darwin, E. Fantin Irudaya Raj, M. Appadurai, M. Chithambara Thanu

Preview Chapter (/viewtitlesample.aspx?id=319990&ptid=293050&t=A Detailed Study on Single Electron Transistors in Nano Device Technologies&isxn=9781668449745) **$37.50** Add to Cart

Chapter 6

Electronic Cooling (/chapter/electronic-cooling/319991) (pages 100-122)

Shankara Murthy H. M., Niranjana Rai, Ramakrishna N. Hegde

Preview Chapter (/viewtitlesample.aspx?id=319991&ptid=293050&t=Electronic Cooling&isxn=9781668449745) **$37.50** Add to Cart

Chapter 7

AIoT and Deep Neural Network-Based Accelerators for Healthcare and Biomedical Applications (/chapter/aiot-and-deep-neural-network-based-accelerators-for-healthcare-and-biomedical-applications/319992) (pages 123-141)

Jothimani K., Bhagya Jyothi K. L.

Preview Chapter (/viewtitlesample.aspx?id=319992&ptid=293050&t=AIoT and Deep Neural Network-Based Accelerators for Healthcare and Biomedical Applications&isxn=9781668449745) **$37.50** Add to Cart

Chapter 8

Review of Applications of Energy Harvesting for Autonomous Wireless Sensor Nodes (/chapter/review-of-applications-of-energy-harvesting-for-autonomous-wireless-sensor-nodes/319994) (pages 143-165)

Wilma Pavitra Puthran, Sahana Prasad, Rathishchandra Ramachandra Gatti

Preview Chapter (/viewtitlesample.aspx?id=319994&ptid=293050&t=Review of Applications of Energy Harvesting for Autonomous Wireless Sensor Nodes&isxn=9781668449745) **$37.50** Add to Cart

Chapter 9

An Exhaustive Analysis of Energy Harvesting Absorbers and Battery Charging Systems for the Internet of Things (/chapter/an-exhaustive-analysis-of-energy-harvesting-absorbers-and-battery-charging-systems-for-the-internet-of-things/319995) (pages 166-186)

C. Padmavathy, Dankan Gowda V., Vaishali Narendra Agme, Algubelly Yashwanth Reddy, D. Palanikkumar

Preview Chapter (/viewtitlesample.aspx?id=319995&ptid=293050&t=An Exhaustive Analysis of Energy Harvesting Absorbers and Battery Charging Systems for the Internet of Things&isxn=9781668449745) **$37.50** Add to Cart

Chapter 10

Wireless Sensor and Actuator Networks-Based Reliable Data Acquisition Mechanism (/chapter/wireless-sensor-and-actuator-networks-based-reliable-data-acquisition-mechanism/319996) (pages 187-213)

Anil Sharma, Dankan Gowda V., A. Yasmine Begum, D. Nageswari, S. Lokesh

Preview Chapter (/viewtitlesample.aspx?id=319996&ptid=293050&t=Wireless Sensor and Actuator Networks-Based Reliable Data Acquisition Mechanism&isxn=9781668449745) **$37.50** Add to Cart

Chapter 11

Minimize the Energy Consumption for Communication Protocol in IoT (/chapter/minimize-the-energy-consumption-for-communication-protocol-in-iot/319997) (pages 214-234)

Manjula Gururaj Rao, Sumathi Pawar, Priyanka H., Hemant Kumar Reddy

Preview Chapter (/viewtitlesample.aspx?id=319997&ptid=293050&t=Minimize the Energy Consumption for Communication Protocol in IoT&isxn=9781668449745) **$37.50** Add to Cart

Chapter 12

An Energy-Efficient Keyless Approach to Home Security Using Internet of Things (/chapter/an-energy-efficient-keyless-approach-to-home-security-using-internet-of-things/319998) (pages 235-256)

Sandeep Kumar Hegde, Rajalaxmi Hegde

Preview Chapter (/viewtitlesample.aspx?id=319998&ptid=293050&t=An Energy-Efficient Keyless Approach to Home Security Using Internet of Things&isxn=9781668449745) **$37.50** Add to Cart

Chapter 13

An Enhanced Method for Running Embedded Applications in a Power-Efficient Manner (/chapter/an-enhanced-method-for-running-embedded-applications-in-a-power-efficient-manner/319999) (pages 257-277)

N. M. G. Kumar, Ayaz Ahmad, Dankan Gowda V., S. Lokesh, Kirti Rahul Rahul Kadam

Preview Chapter (/viewtitlesample.aspx?id=319999&ptid=293050&t=An Enhanced Method for Running Embedded Applications in a Power-Efficient Manner&isxn=9781668449745) **$37.50** Add to Cart

Chapter 14

Production and Use of Electric Vehicle Batteries (/chapter/production-and-use-of-electric-vehicle-batteries/320001) (pages 279-304)

Hasan Huseyin Coban

Preview Chapter (/viewtitlesample.aspx?id=320001&ptid=293050&t=Production and Use of Electric Vehicle Batteries&isxn=9781668449745) **$37.50** Add to Cart

Chapter 15

Stochastic Data Envelopment Analysis in Measuring the Efficiency of Electricity Distribution Companies (/chapter/stochastic-data-envelopment-analysis-in-measuring-the-efficiency-of-electricity-distribution-companies/320002) (pages 305-334)

Zühre Aydın, Bilal Toklu

Preview Chapter (/viewtitlesample.aspx?id=320002&ptid=293050&t=Stochastic Data Envelopment Analysis in Measuring the Efficiency of Electricity Distribution Companies&isxn=9781668449745) **$37.50** Add to Cart

About the Contributors

View Full PDF (/pdf.aspx?tid=320004&ptid=293050&ctid=17&t=About the Contributors&isxn=9781668449745)

Index

View Full PDF (/pdf.aspx?tid=320005&ptid=293050&ctid=17&t=Index&isxn=9781668449745)

Wireless Sensor and Actuator Networks-Based Reliable Data Acquisition Mechanism

Anil Sharma, Dankan Gowda V., A. Yasmine Begum, D. Nageswari, S. Lokesh

Source Title: Energy Systems Design for Low-Power Computing (/book/energy-systems-design-low-power/293050)

Pages: 27
DOI: 10.4018/978-1-6684-4974-5.ch010

Abstract

The objective of deploying a wireless sensor network is to collect data about the environment in which it will be utilised and then to transmit that data to a distant sink where it will be used to estimate or reconstruct the environment or event. In order for the centralised sink to be able to accurately reconstruct or estimate the event and take the necessary actions on time, the wireless sensor and actuator network must be able to guarantee delivery of a sufficient amount of the information gathered by the deployed sensor nodes in a time-bound and coherent manner. In addition to the above-mentioned fundamental problem, reliability also refers to the network's capacity to tolerate defects up to a certain point without compromising performance. This chapter introduces a brand-new, dependable data acquisition technique that makes use of wireless sensor and actuator networks.

Chapter Preview

Top

Introduction

A Wireless Sensor Network (WSN) is deployed with the purpose of acquiring information and then passing on the acquired information to a remote sink where the information can be used to estimate or re-construct the environment or event. This requires the WSN to have the ability to sense the parameter or event under consideration, in the region of deployment, and then ensure reliable delivery of the information to a centralized sink where the information sensed by it can be used to re-construct the events (Abroshan, S., & Moghaddam, M. H. Y. 2014). However, in context of the Wireless Sensor & Actuator Networks the Network Latency time assumes much greater importance since the action taken by the Actuator nodes has to be time coherent with the event sensed by the deployed nodes failing which the control loop will become un-stable and erratic.

This chapter presents a novel Reliable Data Acquisition methodology using Wireless Sensor & Actuator Networks which meets the criterion of reliability as mentioned above. The proposed methodology, ensures, regarding a sensed event is reported to the centralized sink / Actuators for acceptable estimation or re-construction of the event detected, within the time constraint fixed by the control action to be taken by the Actuator nodes.

The objective in this approach is for the Sensor Network to continue its operation with as high reliability as is feasible under the given fault condition in the network (Adelstein, 2005). The fault conditions could vary from temporary loss of communication between a set of nodes, to permanent loss of a node because of damage or end of battery life of the node. The focus of design of algorithm tends to be on ensuring detection of event and then delivery of sensed despite the fault condition. Some algorithms also focus on detection of fault condition itself and then managing it.

Majority of these solutions propose algorithms with the ability to dynamically find alternate network paths in case a given network becomes un-usable because of a fault condition. Many algorithms present methodology where the data sensed is transferred on multiple paths simultaneously to begin with thus ensuring delivery of information at sink despite some path becoming useless because of node failure (Ali, S., Fakoorian, A. & Taheri, H. 2017).

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=319982&ptid=293050&ctid=15&t=Table of Contents&isxn=9781668449745)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=319983&ptid=293050&ctid=15&t=Detailed Table of Contents&isxn=9781668449745)

Preface — View Full PDF (/pdf.aspx?tid=319984&ptid=293050&ctid=15&t=Preface&isxn=9781668449745)

Rathishchandra Ramachandra Gatti, Chandra Singh, Srividya P., Sandeep Bhat

Chapter 1

Low-Power Methodologies and Strategies in VLSI Circuits (/chapter/low-power-methodologies-and-strategies-in-vlsi-circuits/319986) (pages 1-16)

Preethi, Sapna R., Mohammed Mujeer Ulla

Preview Chapter (/viewtitlesample.aspx?id=319986&ptid=293050&t=Low-Power Methodologies and Strategies in VLSI Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 2

SOI Technology in Designing Low-Power VLSI Circuits (/chapter/soi-technology-in-designing-low-power-vlsi-circuits/319987) (pages 17-28)

Srividya P.

Preview Chapter (/viewtitlesample.aspx?id=319987&ptid=293050&t=SOI Technology in Designing Low-Power VLSI Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 3

Low-Power High-Speed Eight-Bit Universal Shift Register Design Using Clock Gating Technique (/chapter/low-power-high-speed-eight-bit-universal-shift-register-design-using-clock-gating-technique/319988) (pages 29-43)

Preeti Sahu

Preview Chapter (/viewtitlesample.aspx?id=319988&ptid=293050&t=Low-Power High-Speed Eight-Bit Universal Shift Register Design Using Clock Gating Technique&isxn=9781668449745) **$37.50** Add to Cart

Chapter 4

Dynamic Body Bias: A Transistor-Level Technique for the Design of Low-Voltage CMOS Analog Circuits (/chapter/dynamic-body-bias/319989) (pages 44-66)

Vandana Niranjan

Preview Chapter (/viewtitlesample.aspx?id=319989&ptid=293050&t=Dynamic Body Bias: A Transistor-Level Technique for the Design of Low-Voltage CMOS Analog Circuits&isxn=9781668449745) **$37.50** Add to Cart

Chapter 5

A Detailed Study on Single Electron Transistors in Nano Device Technologies (/chapter/a-detailed-study-on-single-electron-transistors-in-nano-device-technologies/319990) (pages 67-99)

S. Darwin, E. Fantin Irudaya Raj, M. Appadurai, M. Chithambara Thanu

Preview Chapter (/viewtitlesample.aspx?id=319990&ptid=293050&t=A Detailed Study on Single Electron Transistors in Nano Device Technologies&isxn=9781668449745) **$37.50** Add to Cart

Chapter 6

Electronic Cooling (/chapter/electronic-cooling/319991) (pages 100-122)

Shankara Murthy H. M., Niranjana Rai, Ramakrishna N. Hegde

Preview Chapter (/viewtitlesample.aspx?id=319991&ptid=293050&t=Electronic Cooling&isxn=9781668449745) **$37.50** Add to Cart

Chapter 7

AIoT and Deep Neural Network-Based Accelerators for Healthcare and Biomedical Applications (/chapter/aiot-and-deep-neural-network-based-accelerators-for-healthcare-and-biomedical-applications/319992) (pages 123-141)

Jothimani K., Bhagya Jyothi K. L.

Preview Chapter (/viewtitlesample.aspx?id=319992&ptid=293050&t=AIoT and Deep Neural Network-Based Accelerators for Healthcare and Biomedical Applications&isxn=9781668449745) **$37.50** Add to Cart

Chapter 8

Review of Applications of Energy Harvesting for Autonomous Wireless Sensor Nodes (/chapter/review-of-applications-of-energy-harvesting-for-autonomous-wireless-sensor-nodes/319994) (pages 143-165)

Wilma Pavitra Puthran, Sahana Prasad, Rathishchandra Ramachandra Gatti

Preview Chapter (/viewtitlesample.aspx?id=319994&ptid=293050&t=Review of Applications of Energy Harvesting for Autonomous Wireless Sensor Nodes&isxn=9781668449745) **$37.50** Add to Cart

Chapter 9

An Exhaustive Analysis of Energy Harvesting Absorbers and Battery Charging Systems for the Internet of Things (/chapter/an-exhaustive-analysis-of-energy-harvesting-absorbers-and-battery-charging-systems-for-the-internet-of-things/319995) (pages 166-186)

C. Padmavathy, Dankan Gowda V., Vaishali Narendra Agme, Algubelly Yashwanth Reddy, D. Palanikkumar

Preview Chapter (/viewtitlesample.aspx?id=319995&ptid=293050&t=An Exhaustive Analysis of Energy Harvesting Absorbers and Battery Charging Systems for the Internet of Things&isxn=9781668449745) **$37.50** Add to Cart

Chapter 10

Wireless Sensor and Actuator Networks-Based Reliable Data Acquisition Mechanism (/chapter/wireless-sensor-and-actuator-networks-based-reliable-data-acquisition-mechanism/319996) (pages 187-213)

Anil Sharma, Dankan Gowda V., A. Yasmine Begum, D. Nageswari, S. Lokesh

Preview Chapter (/viewtitlesample.aspx?id=319996&ptid=293050&t=Wireless Sensor and Actuator Networks-Based Reliable Data Acquisition Mechanism&isxn=9781668449745) **$37.50** Add to Cart

Chapter 11

Minimize the Energy Consumption for Communication Protocol in IoT (/chapter/minimize-the-energy-consumption-for-communication-protocol-in-iot/319997) (pages 214-234)

Manjula Gururaj Rao, Sumathi Pawar, Priyanka H., Hemant Kumar Reddy

Preview Chapter (/viewtitlesample.aspx?id=319997&ptid=293050&t=Minimize the Energy Consumption for Communication Protocol in IoT&isxn=9781668449745) **$37.50** Add to Cart

Chapter 12

An Energy-Efficient Keyless Approach to Home Security Using Internet of Things (/chapter/an-energy-efficient-keyless-approach-to-home-security-using-internet-of-things/319998) (pages 235-256)

Sandeep Kumar Hegde, Rajalaxmi Hegde

Preview Chapter (/viewtitlesample.aspx?id=319998&ptid=293050&t=An Energy-Efficient Keyless Approach to Home Security Using Internet of Things&isxn=9781668449745) **$37.50** Add to Cart

Chapter 13

An Enhanced Method for Running Embedded Applications in a Power-Efficient Manner (/chapter/an-enhanced-method-for-running-embedded-applications-in-a-power-efficient-manner/319999) (pages 257-277)

N. M. G. Kumar, Ayaz Ahmad, Dankan Gowda V., S. Lokesh, Kirti Rahul Rahul Kadam

Preview Chapter (/viewtitlesample.aspx?id=319999&ptid=293050&t=An Enhanced Method for Running Embedded Applications in a Power-Efficient Manner&isxn=9781668449745) **$37.50** Add to Cart

Chapter 14

Production and Use of Electric Vehicle Batteries (/chapter/production-and-use-of-electric-vehicle-batteries/320001) (pages 279-304)

Hasan Huseyin Coban

Preview Chapter (/viewtitlesample.aspx?id=320001&ptid=293050&t=Production and Use of Electric Vehicle Batteries&isxn=9781668449745) **$37.50** Add to Cart

Chapter 15

Stochastic Data Envelopment Analysis in Measuring the Efficiency of Electricity Distribution Companies (/chapter/stochastic-data-envelopment-analysis-in-measuring-the-efficiency-of-electricity-distribution-companies/320002) (pages 305-334)

Zühre Aydın, Bilal Toklu

Preview Chapter (/viewtitlesample.aspx?id=320002&ptid=293050&t=Stochastic Data Envelopment Analysis in Measuring the Efficiency of Electricity Distribution Companies&isxn=9781668449745) **$37.50** Add to Cart

About the Contributors — View Full PDF (/pdf.aspx?tid=320004&ptid=293050&ctid=17&t=About the Contributors&isxn=9781668449745)

Index — View Full PDF (/pdf.aspx?tid=320005&ptid=293050&ctid=17&t=Index&isxn=9781668449745)

Recent Advances in Metrology pp 19–27

Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future

V. Dankan Gowda, B. S. Puneeth Kumar, P. Ramesh Naidu & S. B. Sridhara

Conference paper | First Online: 28 July 2022

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 906)

Abstract

Intelligent technology to cope with bandwidth shortage problems has arisen as cognitive radio (CR). The CR is supposed to utilise the unoccupied spectrum band if the approved person may not use it. Since this technology was introduced in 1999, comprehensive research has taken place, addressing

diverse challenges including spectrum sensing, coordination among CR users and the applicability of CR networks. In this article, we introduce new applications of Internet of Things (IoT) CR technology and provide adequate solutions for the real CR technology problems that render IoT more accessible and more useful and also provide an overview of CR-based IoT systems. We illustrate feasible implementations of IoT systems focused on CR.

Keywords

Cooperative communication

Pervasive computing **Internet of Things**

Cognitive radio

References

1. Kakkavas G, Tsitseklis K, Karyotis V Papavassiliou S (2019) A software defined radio cross-layer resource allocation approach for cognitive radio networks: from theory to practice. IEEE Trans Cogn

Commun Network.
https://doi.org/10.1109/TCCN.2019.2963869

2. Ren J, Zhang Y, Deng R, Zhang N; Zhang D, Shen X (2016 April 21) Joint channel access and sampling rate control in energy harvesting cognitive radio sensor networks. IEEE Trans Emerg Top Comput PP(99) pp1–1

3. Shah G, Akan O (2015) Cognitive adaptive medium access control in cognitive radio sensor networks. IEEE Trans Veh Tech 64(2):757–767

4. Li N, Xiao M, Rasmussen LK (2017)Spectrum sharing with network coding for multiple cognitive users. IEEE Internet Things J.
https://doi.org/10.1109/JIOT.2017.2728626

5. Aijaz A, Aghvami AH (2015) Cognitive machine-to-machine communications for internet-of-things: a protocol stack perspective. IEEE Internet Things J 2(2):103–112

6. Chiwewe TM, Hancke GP (2017) Fast convergence cooperative dynamic spectrum access for cognitive radio networks. IEEE Trans Indust Inf https://doi.org/10.1109/TII.2017.2783973

7. Aslam S, Ejaz W, Ibnkahla M (2018) Energy and spectral efficient cognitive radio sensor networks for internet of things. IEEE Internet Things J https://doi.org/10.1109/JIOT.2018.2837354

8. Ejaz W, Shah GA, Kim HS et al (2015) Energy and throughput efficient cooperative spectrum sensing in cognitive radio sensor networks. Trans Emerg Telecommun Technol 26(7):1019–1030

9. Maghsudi S, Stanczak S (2016) Hybrid centralized-distributed resource allocation for device-to-device communication underlaying cellular networks. IEEE Trans Veh Technol

10. Xu Y, Wang J, Wu Q, Anpalagan A, Yao YD (2012) Opportunistic spectrum access in unknown dynamic environment: a game-theoretic stochastic learning solution. IEEE Trans Wirel Commun 1380–1391

11. Anandkumar A, Michael N, Tang A, Swami A (2012) Distributed algorithms for learning and cognitive medium access with logarithmic regret. IEEE J Sel Areas Commun Adv Cogn Radio Network Commun 29:731–745

12. Vakili S, Liu K, Zhao Q (2013) Deterministic sequencing of exploration and exploitation for multi-armed bandit problems. IEEE J Sel Top Sig Process 59(3):1902–1916

13. Gai Y, Krishnamachari B (2011) Decentralized online learning algorithms for opportunistic spectrum access. Proc IEEE Global Commun Conf (GLOBECOM)pp 1–6

14. Xu Y, Wu Q, Wang J, Shen L, Anpalagan A (2015) Robust multiuser sequential channel sensing and access in dynamic cognitive radio networks: potential games and stochastic learning. IEEE Trans Veh Technol 64(8):3594–3607

15. Zhou F, Beaulieu NC, Cheng J, Chu Z, Wang Y (2017) Robust Max–Min fairness resource allocation in sensing-based wideband cognitive radio with SWIPT: imperfect channel sensing. IEEE Syst J https://doi.org/10.1109/JSYST.2017.2698502

16. Jang H, Yun SY, Shin J, Yi Y (2018) IEEE Trans Wirel Commun 17(1):194–209. https://doi.org/10.1109/TWC.2017.2764081

17. Lu Y, Duel-Hallen A (2018) A sensing contribution-based two-layer game for channel selection and spectrum access in cognitive radio ad-hoc networks. IEEE Trans Wirel Commun 17(6):3631–3640

18. Ren J, Zhang Y, Zhang N, Zhang D, Shen X (2016) Dynamic channel access to improve energy efficiency in cognitive radio sensor networks. IEEE Trans Wireless Commun 15(5):3143–3156

19. Mansourkiaie F, Ahmed MH (2016) Optimal and near-optimal cooperative routing and power allocation for collision minimization in wireless sensor networks. IEEE Sens J 16(5):1398–1411

Author information

Authors and Affiliations

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka, India

V. Dankan Gowda

Department of Computer Science, St. Joseph's College, Bangalore, Karnataka, India

B. S. Puneeth Kumar

Department of Computer Science and Engineering, Nitte Meenakshi Institute of Technology, Bangalore, Karnataka, India

P. Ramesh Naidu

Department of Electronics and Communication Engineering, Vijaya Vittala Institute of Technology, Bangalore, Karnataka, India

S. B. Sridhara

Corresponding author

Correspondence to V. Dankan Gowda .

Editor information

Editors and Affiliations

CSIR-NPL, New Delhi, India

Dr. Sanjay Yadav

Maharaja Surajmal Institute of Technology, New Delhi, India

Prof. K.P. Chaudhary

Maharaja Surajmal Institute of Technology, New Delhi, India

Dr. Ajay Gahlot

Maharaja Surajmal Institute of Technology, New Delhi, India

Dr. Yogendra Arya

Maharaja Surajmal Institute of Technology, New Delhi, India

Dr. Aman Dahiya

CSIR-NPL, New Delhi, India

Dr. Naveen Garg

Rights and permissions

Copyright information

About this paper

Cite this paper

Dankan Gowda, V., Puneeth Kumar, B.S., Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (eds) Recent Advances in Metrology . Lecture Notes in Electrical

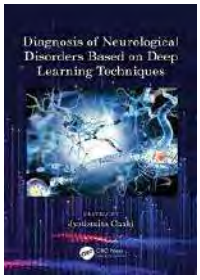
Engineering, vol 906. Springer, Singapore.
https://doi.org/10.1007/978-981-19-2468-2_3

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-19-2468-2_3	28 July 2022	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-2467-5	978-981-19-2468-2	Engineering Engineering (R0)

Diagnosis of Neurological Disorders Based on Deep Learning Techniques (https://www.taylorfrancis.com/books/mono/10.1201/9781003315452/diagnosis-neurological-disorders-based-deep-learning-techniques?refId=a911d9c6-5823-4158-aaf1-03b6920f4891&context=ubx)
Show Path

Chapter



Implementation of a Deep Neural Network-Based Framework for Actigraphy Analysis and Prediction of Schizophrenia

By Vijayalakshmi G V Mahesh (/search?contributorName=Vijayalakshmi G V Mahesh&contributorRole=author&redirectFromPDP=true&context=ubx), Alex Noel Joseph Raj (/search?contributorName=Alex Noel Joseph Raj&contributorRole=author&redirectFromPDP=true&context=ubx), R Chandraprabha (/search?contributorName=R Chandraprabha&contributorRole=author&redirectFromPDP=true&context=ubx)

Book Diagnosis of Neurological Disorders Based on Deep Learning Techniques (https://www.taylorfrancis.com/books/mono/10.1201/9781003315452/diagnosis-neurological-disorders-based-deep-learning-techniques?refId=a911d9c6-5823-4158-aaf1-03b6920f4891&context=ubx)

Edition	1st Edition
First Published	2023
Imprint	CRC Press
Pages	15
eBook ISBN	9781003315452

Share

ABSTRACT

< Previous Chapter (chapters/edit/10.1201/9781003315452-9/efficient-classification-schizophrenia-eeg-signals-using-deep-learning-methods-subha-puthankattil-marrapu-vynatheya-ahsan-ali?context=ubx)
Next Chapter > (chapters/edit/10.1201/9781003315452-11/evaluating-psychomotor-skills-autism-spectrum-disorder-deep-learning-ravi-kant-avvari?context=ubx)

Navigate This Page

Buy Hardcover

Qty: 1 **$125.00**

List Price: ~~$250.00~~

You Save: $125.00

Add to Cart

Benefits & Incentives

- 50% Contributor Discount
- Printed-On-Demand (POD)
- Usually ships one day from order

Available in GOBI® (http://www.gobi3.com/hx/Falcon.ashx?location=basicsearchsubmit&type=BusyPod&basicsearchinputtype=Isbn&basicsearchinputmode=AdvancedAny&basicsearchinpu

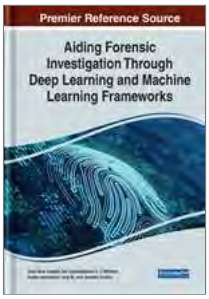
Available in OASIS® (http://oasis.proquest.com/openurl?isbn=9781668445587)

Share

More Information

Available In

Related Books



Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh (/affiliate/vijayalakshmi-gv-mahesh/381804/), Ruban Nerssison, Ang Yu, Jennifer Gentry

Release Date: June, 2022
Copyright: © 2022
Pages: 273
DOI: 10.4018/978-1-6684-4558-7
ISBN13: 9781668445587
ISBN10: 1668445581
EISBN13: 9781668445600

Hardcover: $125.00
List Price: ~~$250.00~~
(/book/aiding-forensic-investigation-through-deep/290354?f=hardcover&i=1) Available
Benefits & Incentives

E-Book: $125.00
List Price: ~~$250.00~~
(/book/aiding-forensic-investigation-through-deep/290354?f=e-book&i=1) Available
Benefits & Incentives

Hardcover + E-Book: $150.00
List Price: ~~$300.00~~
(/book/aiding-forensic-investigation-through-deep/290354?f=hardcover-e-book&i=1) Available
Benefits & Incentives

OnDemand:
(Individual Chapters) $18.75
List Price: ~~$37.50~~
(/book/aiding-forensic-investigation-through-deep/290354#table-of-contents) Available
Benefits & Incentives

Effective immediately, IGI Global has discontinued softcover book production. The softcover option is no longer available for direct purchase.

Description & Coverage

Description:

It is crucial that forensic science meets challenges such as identifying hidden patterns in data, validating results for accuracy, and understanding varying criminal activities in order to be authoritative so as to hold up justice and public safety. Artificial intelligence, with its potential subsets of machine learning and deep learning, has the potential to transform the domain of forensic science by handling diverse data, recognizing patterns, and analyzing, interpreting, and presenting results. Machine Learning and deep learning frameworks, with developed mathematical and computational tools, facilitate the investigators to provide reliable results. Further study on the potential uses of these technologies is required to better understand their benefits.

Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks provides an outline of deep learning and machine learning frameworks and methods for use in forensic science to produce accurate and reliable results to aid investigation processes. The book also considers the challenges, developments, advancements, and emerging approaches of deep learning and machine learning. Covering key topics such as biometrics, augmented reality, and fraud investigation, this reference work is crucial for forensic scientists, law enforcement, computer scientists, researchers, scholars, academicians, practitioners, instructors, and students.

Coverage:

The many academic areas covered in this publication include, but are not limited to:

- Augmented Reality
- Biometrics
- Cloud Forensics
- Cybersecurity
- Deep Learning
- Digital Forensics
- Document Analysis
- Forensic Pathology
- Fraud Investigation
- Machine Learning
- Video Tracking

Table of Contents

Table of Contents — View Full PDF (/pdf.aspx?tid=309770&ptid=290354&ctid=15&t=Table of Contents&isxn=9781668445587)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=309771&ptid=290354&ctid=15&t=Detailed Table of Contents&isxn=9781668445587)

Preface — View Full PDF (/pdf.aspx?tid=309772&ptid=290354&ctid=15&t=Preface&isxn=9781668445587)

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh, Ruban Nersisson, Ang Yu, Jennifer Gentry

Acknowledgment — View Full PDF (/pdf.aspx?tid=309773&ptid=290354&ctid=15&t=Acknowledgment&isxn=9781668445587)

Chapter 1

A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework (/chapter/a-review-on-forensic-science-and-criminal-investigation-through-a-deep-learning-framework/309774) (pages 1-72)

Anantanu Koley, Subhabrata Barman

Deep learning (DL) is a rising field that is applied in forensic science and criminal investigation (FSCI). FSCI specialists are confronting many difficulties...

Preview Chapter (/viewtitlesample.aspx?id=309774&ptid=290354&t=A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework&isxn=9781668445587) $37.50 Add to Cart

Chapter 2

Digital Forensic Techniques and Principles in a Cloud Environment (/chapter/digital-forensic-techniques-and-principles-in-a-cloud-environment/309775) (pages 73-91)

Ayaz Ahmad, Supriya Prashant Diwan, Sudhansu Sekhar Nanda, Lingaraj Sethi, Prashanta Kumar Patra

Digital forensic methods for cybercrime investigation in a cloud computing environment (i.e., cloud forensic) is a broad term that encompasses a number of...

Preview Chapter (/viewtitlesample.aspx?id=309775&ptid=290354&t=Digital Forensic Techniques and Principles in a Cloud Environment&isxn=9781668445587) $37.50 Add to Cart

Chapter 3

IoT-Deep Learning-Based Detection of Cyber Security Threats (/chapter/iot-deep-learning-based-detection-of-cyber-security-threats/309776) (pages 92-111)

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech, Ansuman Samal

The internet of things (IoT) is a revolutionary technology that links living and non-living devices all around the world. As a result, the frequency of...

Preview Chapter (/viewtitlesample.aspx?id=309776&ptid=290354&t=IoT-Deep Learning-Based Detection of Cyber Security Threats&isxn=9781668445587) $37.50 Add to Cart

Chapter 4

Deep Learning Models for Cyber Security in IoT Networks (/chapter/deep-learning-models-for-cyber-security-in-iot-networks/309777) (pages 112-127)

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech, Thangadurai N.

As the number of connected devices grows, the internet of things (IoT) poses new security challenges for network activity monitoring. Due to a lack of...

Preview Chapter (/viewtitlesample.aspx?id=309777&ptid=290354&t=Deep Learning Models for Cyber Security in IoT Networks&isxn=9781668445587) $37.50 Add to Cart

Chapter 5

Deep Learning-Based AI Modeling, Intrusion Detection (/chapter/deep-learning-based-ai-modeling-intrusion-detection/309778) (pages 128-143)

Madhab Paul Choudhury, Madhab Paul Choudhury, Chandrashekhar Azad

Machine learning techniques are being used to create an intrusion detection system (IDS) for detecting and classifying cyber-attacks at the network-level and...

Preview Chapter (/viewtitlesample.aspx?id=309778&ptid=290354&t=Deep Learning-Based AI Modeling, Intrusion Detection&isxn=9781668445587) $37.50 Add to Cart

Chapter 6

Face Detection and Recognition From Distance Based on Deep Learning (/chapter/face-detection-and-recognition-from-distance-based-on-deep-learning/309779) (pages 144-160)

Hui Wang, Wei Qi Yan

Face recognition is an important biometric in video surveillance. However, the conventional algorithms of face recognition are also susceptible to various...

Preview Chapter (/viewtitlesample.aspx?id=309779&ptid=290354&t=Face Detection and Recognition From Distance Based on Deep Learning&isxn=9781668445587) $37.50 Add to Cart

Chapter 7

Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks (/chapter/investigation-of-deep-fake-images-using-pre-trained-cnn-frameworks/309780) (pages 161-173)

Preview Chapter (/viewtitlesample.aspx?id=309780&ptid=290354&t=Investigation $37.50 Add to Cart

trained CNN frameworks/309780) (pages 151-173)

Anitha Ruth J., Uma R., Vijayalakshmi G. V. Mahesh, P. Ramkumar

Deep learning technologies, such as neural networks, have tackled complicated issues from large-scale data processing to computer vision and human-level...

Preview Chapter (/viewtitlesample.aspx?id=309781&ptid=290354&t=Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics&isxn=9781668445587) **$37.50** Add to Cart

Chapter 8

Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics (/chapter/implementation-of-machine-learning-aided-speech-analysis-for-speaker-accent-identification-applied-to-audio-forensics/309781) (pages 174-194)

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Accent recognition as a subset of speech recognition is crucial in audio forensics as it provides the authenticity of the speech that can be presented to the...

Preview Chapter (/viewtitlesample.aspx?id=309782&ptid=290354&t=Medical Image Encryption Using Distance-Based Biokey Generation&isxn=9781668445587) **$37.50** Add to Cart

Chapter 9

Medical Image Encryption Using Distance-Based Biokey Generation (/chapter/medical-image-encryption-using-distance-based-biokey-generation/309782) (pages 195-218)

Vijayarajan Rajangam, Shani S. Kumar, Sathiya Narayanan, Sangeetha N., Avudaiammal R.

As the internet of things and cloud computing create new technological revolutions in the field of healthcare, the issues such as security, privacy...

Preview Chapter (/viewtitlesample.aspx?id=309783&ptid=290354&t=Deep Learning-Based Automatic Student Authentication&isxn=9781668445587) **$37.50** Add to Cart

Chapter 10

Deep Learning-Based Automatic Student Authentication (/chapter/deep-learning-based-automatic-student-authentication/309783) (pages 219-236)

Sheetalrani R Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.

Classroom management relies heavily on the ability to keep track of student attendance. Attendance checks by calling names or handing out a sign-in sheet are...

About the Contributors — View Full PDF (/pdf.aspx?tid=309785&ptid=290354&ctid=17&t=About the Contributors&isxn=9781668445587)

Index — View Full PDF (/pdf.aspx?tid=309786&ptid=290354&ctid=17&t=Index&isxn=9781668445587)

Editor/Author Biographies

Vijayalakshmi G. V. Mahesh received the B. E. degree in Electronics and Communication Engineering from Bangalore University, India, in 1999, the M.Tech. degree in Digital Communication and Networking from Visvesvaraya Technological University in 2005 and Ph.D degree from VIT, Vellore, India. Currently she is working as an Associate Professor at BMS Institute of Technology and Management, Bangalore. Her research interests include Machine Learning, Image Processing, Signal Processing, Artificial Intelligence and Pattern Recognition.

Archiving

All of IGI Global's content is archived via the CLOCKSS and LOCKSS initiative. Additionally, all IGI Global published content is available in IGI Global's InfoSci® platform.

Segmentation of Optic Disc From Fundus Image Based on Morphology and SVM Classifier

Jiamin Luo, Alex Noel Joseph Raj, Nersisson Ruban, Vijayalakshmi G. V. Mahesh (/affiliate/vijayalakshmi-gv-mahesh/381804/)

Source Title: Handbook of Research on Deep Learning-Based Image Analysis Under Constrained and Unconstrained Environments (/book/handbook-research-deep-learning-based/256635)

Pages: 29

DOI: 10.4018/978-1-7998-6690-9.ch007

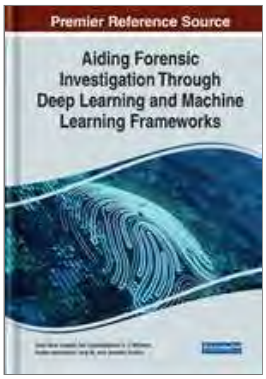
Abstract

Color fundus image is the most basic way to diagnose diabetic retinopathy, papillary edema, and glaucoma. In particular, since observing the morphological changes of the optic disc is conducive to the diagnosis of related diseases, accurate and effective positioning and segmentation of the optic disc is an important process. Optic disc segmentation algorithms are mainly based on template matching, deformable model and learning. According to the character that the shape of the optic disc is approximately circular, this proposed research work uses Kirsch operator to get the edge of the green channel fundus image through morphological operation, and then detects the optic disc by HOUGH circle transformation. In addition, supervised learning in machine learning is also applied in this chapter. First, the vascular mask is obtained by morphological operation for vascular erasure, and then the SVM classifier is segmented by HU moment invariant feature and gray level feature. The test results on the DRIONS fundus image database with expert-labeled optic disc contour show that the two methods have good results and high accuracy in optic disc segmentation. Even though seven different assessment parameters (sensitivity [Se], specificity [Sp], accuracy [Acc], positive predicted value [Ppv], and negative predicted value [Npv]) are used for performance assessment of the algorithm. Accuracy is considered as the criterion of judgment in this chapter. The average accuracy achieved for the nine random test set is 97.7%, which is better than any other classifiers used for segmenting Optical Disc from Fundus Images.

Chapter Preview

Top

1. Introduction



Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Source Title: Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks (/book/aiding-forensic-investigation-through-deep/290354)

Pages: 21
DOI: 10.4018/978-1-6684-4558-7.ch008

Abstract

Accent recognition as a subset of speech recognition is crucial in audio forensics as it provides the authenticity of the speech that can be presented to the judicature as evidence. The challenge here is to design a speaker accent identifying system that can provide significant information to match with the standards of the court of justice. This process requires robust descriptors to represent speech signal with good discrimination ability. This chapter proposes to use Mel Frequency Cepstral Coefficients to identify and represent the human utterances by transforming the frequency from normal scale to Mel scale. The work utilized support vector machine, k nearest neighbors, XG boost, linear discriminant analysis, quadratic discriminant analysis, and decision tree algorithms to recognize the accent of the speaker. The experiment conducted on the accent recognition dataset demonstrated the ability of Mel Frequency Cepstral Coefficients and kNN classifier in identifying and discriminating six accents belonging to different speakers with better accuracy.

Chapter Preview

Top

Introduction

With the advancements in Artificial intelligence, signal processing and machine learning algorithms, there is growing demand for speech processing and speech/accent recognition applications in banking sector, tourism industry, marketing, healthcare, law enforcement, forensics, language learning etc., Further the introduction of voice assistants with AI and Internet of Things have transformed speech recognition technology and have changed the way human interact with devices. Speech or accent recognition involves understanding, analyzing the human speech and make the right decision.

Speech processing and speech/accent recognition plays a significant role in audio forensics (AF). AF is a part of forensic science that serves as evidence for investigation as required for law enforcement. This involves (i) Audio/Speech data acquisition, (ii) Analysis, (iii) Interpretation and (iv) Evaluation as major processing steps. The overview of the process is displayed in Fig.1.The acquired data has to be preprocessed to improve its quality by using filters and enhancement techniques. The preprocessing improves the speech intelligibility by removing background noise, distortions and artifacts thus providing authenticity of the evidence. Further it is very important to note the key frequency components and frequency range in the audio/speech signal to observe and gather required information to make it available as evidence.

Figure 1. Overview of Audio Forensic Process
978-1-6684-4558-7.ch008.f01(https://igiprodst.blob.core.windows.net:443/source-content/9781668445587_290354/978-1-6684-4558-7.ch008.f01.png?sv=2015-12-11&sr=c&sig=TPMOH5btWi1nHwY0SawQrVcscZdmOWTeuTdo822A7jI%3D&se=2022-09-23T05%3A01%3A05Z&sp=r)
Digital signal processing algorithms such as Fourier transform, Wavelet transform, and Mel Frequency Cepstral Coefficients serve as a better choice for frequency analysis of the audio signals. This analysis later serves in identifying the suspect based on identifying the speech and accent of the speaker. This chapter emphasizes on applying frequency analysis to extract MFCC features from the speech signals and provide them to the machine learning algorithm and evaluate its output to interpret the result.

Related Work

Top

A lot of research has been carried out on speech recognition and to identify the accent of the speaker. A critical review on the various methods used for accent-based speech recognition is provided in (Thandil and Basheer, 2020; Bhagath and Das, 2004). Automatic speech recognition framework can be developed based on acoustic phonetic method and pattern recognition (Jahangir et al., 2021).

Acoustic phonetic method explores the acoustic properties of speech signal such as time domain characteristics of the speech that include the fundamental frequency, its duration, mean squared amplitude of the signal or transform domain characteristics such as frequency spectrum to represent and identify the speech. Whereas the pattern recognition (Rabiner, 1992) methods involve extracting the patterns from the speech signal, training the pattern to generate the model and finally adapting the model for speech recognition. Numerous approaches have been developed to extract the patterns or features based on independent component analysis (ICA), principal component analysis (PCA), Liner predictive coding (LPC), zero crossing detection (ZCD), Wavelet transform, Mel Frequency Cepstral coefficient (MFCC), Guassian mixture modeling (GMM), Vector quantization (VQ), Dynamic time warping (DTW), Hidden Markov modeling (HMM) and deep learning architectures (Nassif et al., 2019) to be applied in speech recognition.

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=309770&ptid=290354&ctid=15&t=Table of Contents&isxn=9781668445587)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=309771&ptid=290354&ctid=15&t=Detailed Table of Contents&isxn=9781668445587)

Preface — View Full PDF (/pdf.aspx?tid=309772&ptid=290354&ctid=15&t=Preface&isxn=9781668445587)

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh, Ruban Nersisson, Ang Yu, Jennifer Gentry

Acknowledgment — View Full PDF (/pdf.aspx?tid=309773&ptid=290354&ctid=15&t=Acknowledgment&isxn=9781668445587)

Chapter 1

A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework (/chapter/a-review-on-forensic-science-and-criminal-investigation-through-a-deep-learning-framework/309774) (pages 1-72)

Pinaki Pratim Acharjya, Santanu Koley, Subhabrata Barman

Preview Chapter (/viewtitlesample.aspx?id=309774&ptid=290354&t=A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework&isxn=9781668445587) $37.50 Add to Cart

Chapter 2

Digital Forensic Techniques and Principles in a Cloud Environment (/chapter/digital-forensic-techniques-and-principles-in-a-cloud-environment/309775) (pages 73-91)

Ayaz Ahmad, Supriya Prashant Diwan, Sudhansu Sekhar Nanda, Lingaraj Sethi, Prashanta Kumar Patra

Preview Chapter (/viewtitlesample.aspx?id=309775&ptid=290354&t=Digital Forensic Techniques and Principles in a Cloud Environment&isxn=9781668445587) $37.50 Add to Cart

Chapter 3

IoT-Deep Learning-Based Detection of Cyber Security Threats (/chapter/iot-deep-learning-based-detection-of-cyber-security-threats/309776) (pages 92-111)

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech, Ansuman Samal

Preview Chapter (/viewtitlesample.aspx?id=309776&ptid=290354&t=IoT-Deep Learning-Based Detection of Cyber Security Threats&isxn=9781668445587) $37.50 Add to Cart

Chapter 4

Deep Learning Models for Cyber Security in IoT Networks (/chapter/deep-learning-models-for-cyber-security-in-iot-networks/309777) (pages 112-127)

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech, Thangadurai N.

Preview Chapter (/viewtitlesample.aspx?id=309777&ptid=290354&t=Deep Learning Models for Cyber Security in IoT Networks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 5

Deep Learning-Based AI Modeling, Intrusion Detection (/chapter/deep-learning-based-ai-modeling-intrusion-detection/309778) (pages 128-143)

Madhab Paul Choudhury, Madhab Paul Choudhury, Chandrashekhar Azad

Preview Chapter (/viewtitlesample.aspx?id=309778&ptid=290354&t=Deep Learning-Based AI Modeling, Intrusion Detection&isxn=9781668445587) **$37.50** Add to Cart

Chapter 6

Face Detection and Recognition From Distance Based on Deep Learning (/chapter/face-detection-and-recognition-from-distance-based-on-deep-learning/309779) (pages 144-160)

Hui Wang, Wei Qi Yan

Preview Chapter (/viewtitlesample.aspx?id=309779&ptid=290354&t=Face Detection and Recognition From Distance Based on Deep Learning&isxn=9781668445587) **$37.50** Add to Cart

Chapter 7

Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks (/chapter/investigation-of-deep-fake-images-using-pre-trained-cnn-frameworks/309780) (pages 161-173)

Anitha Ruth J., Uma R., Vijayalakshmi G. V. Mahesh, P. Ramkumar

Preview Chapter (/viewtitlesample.aspx?id=309780&ptid=290354&t=Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 8

Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics (/chapter/implementation-of-machine-learning-aided-speech-analysis-for-speaker-accent-identification-applied-to-audio-forensics/309781) (pages 174-194)

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Preview Chapter (/viewtitlesample.aspx?id=309781&ptid=290354&t=Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics&isxn=9781668445587) **$37.50** Add to Cart

Chapter 9

Medical Image Encryption Using Distance-Based Biokey Generation (/chapter/medical-image-encryption-using-distance-based-biokey-generation/309782) (pages 195-218)

Vijayarajan Rajangam, Shani S. Kumar, Sathiya Narayanan, Sangeetha N., Avudaiammal R.

Preview Chapter (/viewtitlesample.aspx?id=309782&ptid=290354&t=Medical Image Encryption Using Distance-Based Biokey Generation&isxn=9781668445587) **$37.50** Add to Cart

Chapter 10

Deep Learning-Based Automatic Student Authentication (/chapter/deep-learning-based-automatic-student-authentication/309783) (pages 219-236)

Sheetalrani R Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.

Preview Chapter (/viewtitlesample.aspx?id=309783&ptid=290354&t=Deep Learning-Based Automatic Student Authentication&isxn=9781668445587) **$37.50**

Add to Cart

About the Contributors — View Full PDF (/pdf.aspx?tid=309785&ptid=290354&ctid=17&t=About the Contributors&isxn=9781668445587)

Index — View Full PDF (/pdf.aspx?tid=309786&ptid=290354&ctid=17&t=Index&isxn=9781668445587)

Chapter 7
Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks

Anitha Ruth J.
https://orcid.org/0000-0003-1900-6173
SRM Institute of Science and Technology, India

Uma R.
https://orcid.org/0000-0002-0053-0162
Sri Sairam Engineering College, India

Vijayalakshmi G. V. Mahesh
B. M. S. Institute of Technology and Management, India

P. Ramkumar
Sri Sairam College of Engineering, India

ABSTRACT

Deep learning technologies, such as neural networks, have tackled complicated issues from large-scale data processing to computer vision and human-level control. Editing photographs with software designed for commercial purposes makes it easy for anyone to manipulate images in order to create fictitious ones. Generative adversarial networks (GANs) are currently being utilised to generate superficially realistic photographs, as opposed to the old methods that were previously used to create phony images. The GANs images are referred to as deep fakes. To the unaided eye, they appear to be real photographs. As a result, it's impossible to spot a phony image produced with GANs. The social network is made less safe as a result of the uploading of these bogus photographs. It is therefore critical that the digital image's legitimacy be detected before it is uploaded. Thus, in this chapter, the authors suggest a few pre-trained deep learning frameworks that may be used to effectively detect deep fake images.

DOI: 10.4018/978-1-6684-4558-7.ch007

INTRODUCTION

Smart phones' rising complexity and the rise of social networks' digital object content have led to an enormous number of new digital object content in recent years. As a result of the widespread use of digital photographs, a proliferation of tools and methods for manipulating image data has occurred. Because these methods were tedious and time-consuming, they had previously been out of the reach of the majority of people. Artificial intelligence-based false picture production is now possible thanks to deep generative models, which allow for the rapid transmission of high-quality tampered media material (Li et al., 2020). In the field of image improvement and manipulation, one of the most promising recent developments has been the use of Generative Adversarial Networks (GANs). Splicing, resampling, and copy moving are all examples of digital image modifications that can be detected using digital image forensics tools and methodologies (Nataraj et al., 2019*)*. Image augmentation and manipulation have been greatly improved by the use of generative adversarial networks (GANs). GANs are made up of two networks, one that generates false pictures and the other that assesses whether an image is real or fake (Chen et al., 2022). Digital picture forensics is devoted to identifying and preventing image forgeries in order to control the circulation of such fabricated material. " According to various estimates, the internet receives over two billion photographs each day. Image forgery detection has made significant progress, but digital video falsification detection is still a challenging problem. Due to the significant deterioration of video frames when they are compressed, many image processing techniques cannot be applied to video. Digital image forensics has had success using deep learning. However, deep learning may be used to fake videos as well (Bonettini et al., 2020). There are inherent picture disparities across the borders that make face editing procedures fundamentally recognizable despite the standard process of blending a changed face into an existing backdrop (Durall et al., 2019). One of the most important issues in today's digital society is the manipulation of visual material. Face expression manipulation and identity manipulation are two of the most common ways of facial manipulation now in use. It is possible to communicate the facial emotions of one person to another in real time using only inexpensive equipment. This is the second type of facial forgery in which a person's face is replaced with an image of another person. However, DeepFakes uses deep learning to conduct face swapping as well. DeepFakes, on the other hand, must be trained for each pair of movies in order to do face swapping in real time. (de Lima et al., 2020). Massive advancements in the area of facial modification have shown a difficult footprint that may be exposed to the specifics of automated video editing techniques in the last few years. Forensic footprints are frequently quite delicate and difficult to detect, which has sparked a lot of interest in the procedures. New artistic possibilities are opened up by facial

modification (e.g., movie creation, visual effects). However, criminal individuals may easily create video forgeries by manipulating the facial expressions. Detecting face modification using recent technologies is the focus of this article (Afchar, 2019).

This Chapter is organized as follows, background section gives a brief synopsis of the underlying research, main focus of the Chapter section discusses about the Deep learning frameworks that can be used to detect Deep fake images. solutions and recommendations section provide the results from the application of TENSOR FLOW and KERAS frameworks and the final section with future scope and conclusion.

BACKGROUND

Deepfake detection may outperform existing state-of-the-art approaches. This work approaches generalizable deep fake detection from a straightforward premise: a representation should be susceptible to a variety of different forms of forgery. Synthesising augmented forgeries from a pool of forgery configurations and guaranteeing the model's sensitivity to forgery configurations has been proposed. In this research offers a method for detecting deepfakes that is capable of generalising well to previously encountered contexts. To construct forgeries, a synthesiser and an adversarial training system has been used. The primary restriction of the existing technology is that the synthesizer's augmentation type remains limited (Chen et al., 2022). Face manipulation detection in video sequences that target contemporary facial manipulation techniques is the subject of this article. On two publicly available datasets with over 119,900 videos, is has been demonstrated how various models may be generated from a single base network. The prevalence of videos in our daily lives and in mass communication means that the ability to tell whether or not a video has been altered is more important than ever. The method that has been proposed is based on the family of models and improves upon an ensemble of models trained using the fundamental concepts (Bonettini et al., 2020). Despite the fact that many face-editing algorithms create human-like results, they contain artefacts in particular areas that are difficult to detect. Classic frequency domain analysis is the foundation of the approach, followed by a simple classifier. When trained on just 20 annotated examples, it achieves a flawless classification accuracy of 100%. Using two datasets, including data from multiple GANs, it has been demonstrated that the proposed technique can recognise high and medium resolution deepfake images with 100% accuracy. Low-resolution material is more difficult to recognise since the accessible frequency spectrum is smaller, making it more difficult. With a 91% accuracy rate, it was able to detect low-resolution forgeries of an industry standard (Durall et al., 2019). The fast advancement in the creation and manipulation of synthetic images raises serious concerns about the consequences for society. This research

investigates the realism of current image modifications and the difficulty of spotting them in their original form. In addition to a database of over 1:8 million modified photos, the benchmark is accessible to the public. Trained forgery detectors can spot even the most advanced alterations to a person's facial features. A new dataset of films of modified faces has been proposed to train detectors using domain-specific information. This dataset dwarfs all other publicly accessible forensic datasets by an order of magnitude. It has been shown that this dataset and benchmark will serve as a springboard for further investigation (Rossler et al., 2019). Detection approaches for deepfakes rely solely on individual video frames, so they are unable to learn from temporal information. Using the Celeb-DF dataset, a benchmark for the performance of spatiotemporal convolutional algorithms has been established. Modern frame-based detection algorithms were outperformed by the proposed one. AI-generated Deepfakes may be easily detected using action detection methods, which has been explained and assess in detail in this study. Even I3D, which was superior at action. recognition, was no match for R3D in terms of performance. Researchers in the future may benefit from the findings by better understanding how Deepfake can be identified (de Lima, 2019). This research proposes Face X-ray as a new picture format for detecting fraud in face photos. When an image is fake, it shows where the blending line is, and when it's real and is not. Facial X-rays can be taught without the use of any of the current face manipulation methods. There are inherent image inconsistencies across the blending zones in most known face manipulation methods, and a new method to detect face forgeries based on this fact has been suggested. It is now possible to train a face forgery detector using face X-ray instead of the current state-of-the-art face manipulation techniques (Li et al., 2020). Face tampering in videos, including the more current Deepfake and Face2Face approaches, may be detected automatically and effectively using the technology shown in this study. In most cases, image forensics approaches are ineffective. Videos suffer greatly as a result of the compression used. Under real-world internet dispersion conditions, Deepfake detects Deepfake videos with a 98 percent accuracy rate and Face2Face videos with a 95 percent accuracy rate. Detection of phoney faces relies heavily on the eyes and mouth, according to the tests (Afchar et al., 2019). In this paper, pixel co-occurrence matrices and deep learning are combined to present a unique method for detecting GAN-generated bogus images. It is possible to train a Convolutional Neural Network to tell the difference between images created by GAN and those that are real. GAN experiments reveal that the approach is both effective and generalizable, as evidenced by the results (Nataraj et al., 2019). This study is based on deep learning methodologies, here, the author (Almars, 2021) presents an in-depth analysis of deepfake production and detection systems. A full examination of the various technologies and their applications in the identification of deepfakes is also provided. Because it covers the most recent state-of-the-art approaches for discovering

deepfakes in social content. Using neural networks, this research (Sai Vamsi et al., 2022) describes how to distinguish between bogus and real videos. In light of the results, it may be determined that the project was effective in its implementation based on the results. a consequence of training the model on the Celeb-DF dataset using ResNet CNN for frame-level feature recognition, screenshots of its performance and predicted outcomes are provided. For comparing video frames, LSTM had an average accuracy of 91% on the comparisons tested.

MAIN FOCUS OF THE CHAPTER

The term "Deepfake" is a combination of the terms "Deep learning" and "fake," and it refers to information created by an artificial neural network, a field of machine learning that is increasingly popular.
The primary points of the Chapter are summarized as follows,

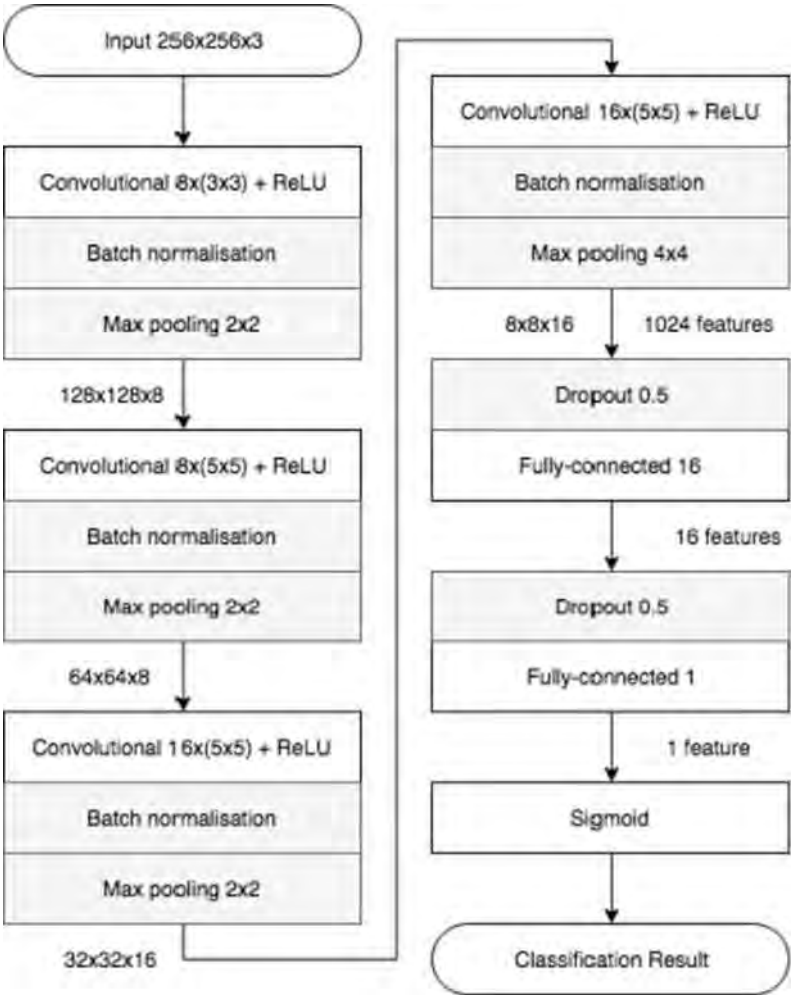
1. The Chapter contributes to discussing many successful Deep learning strategies such as MesoNet, ResNet50 and CNN for dealing with Deepfake images
2. We can detect fabricated face videos by utilising a pre-trained Deep learning techniques at the mesoscopic level of scrutiny.
3. The MESONET, RESNET50 are implemented using TENSOR flow and KERAS framework.
4. The experimental results prove that MESONET framework provides 96% of accuracy in detecting the fake images while using the customized data set.
5. Whereas RESENET50 and CNN provides 94% and 90% accuracy in detecting the fake images while using the Kaggle (2019) dataset.
6. Finally, in comparison to other frameworks, the results show that the MESONET framework is more effective at detecting fake images

The Chapter uses a range of strategies to detect manufactured face movies at the mesoscopic level. Microscopy is not applicable because compressed video has worsened the picture noise to such a degree that microscopic analysis based on image noise is not possible. When an image depicts a person's face features, the human eye has even more difficulty recognising it at a higher semantic level. Deep Neural Network algorithms are used to detect Deep Fake Images efficiently in order to overcome this issue. Deepfake is a video manipulation technique that seeks to replace the face of a targeted individual with the face of another person in a video clip. Pre-trained framework's approach and architecture are discussed in the following paragraphs

MESONET

This network begins with a sequence of four layers of successive convolutions and pooling and is followed by a dense network with one hidden layer. In order to improve generalisation, the convolutional layers employ ReLU activation functions, which introduce nonlinearities, as well as Batch Normalization (Ioffe and Szegedy, 2015), which regularises their output and prevents the vanishing gradient effect. The fully connected layers employ Dropout (He et al., 2015), which regularises their output and improves their robustness. The architecture diagram for MESO-4 is represented in the figure 1 (Gong et al, 2021).

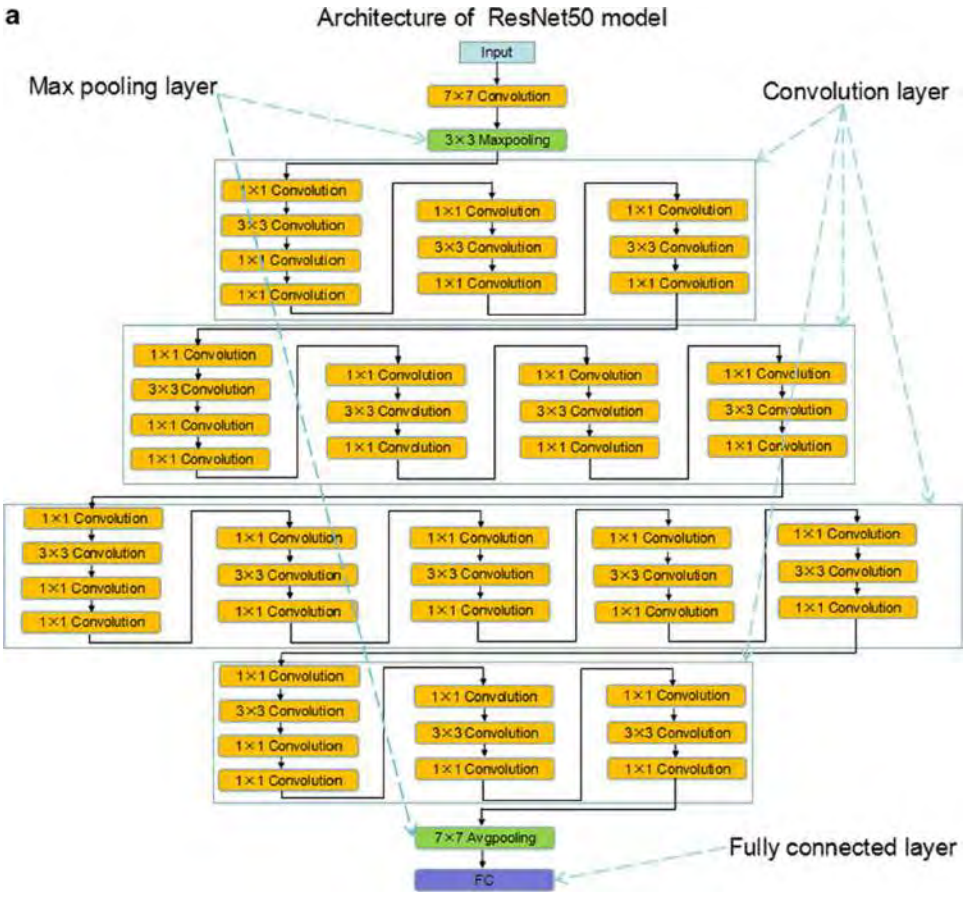
Figure 1. The Architecture of MESO-4



RESNET

In this study we use ResNet model presented by Shaoqing Ren, Kaiming He, Jian Sun, and Xiangyu (He et al., 2015, Goodfellow et al., 2014) is one of the most popular Deep learning models in detecting the Deep fake images. This research uses ResNet to evaluate the performance of deepfake detection because stacking more layers raises the issue of disappearing/exploding gradients while training ultra-deep neural networks Here, ResNet50 is a ResNet version that consists of 48 Convolution layers as well as 1 MaxPool and 1 Average Pool layer is used in detecting the fake images. It is a variant of the ResNet model that is used for image recognition. The following figure 3 shows the architecture of Res Net 50 (Gong et al., 2021, Goodfellow et al., 2014).

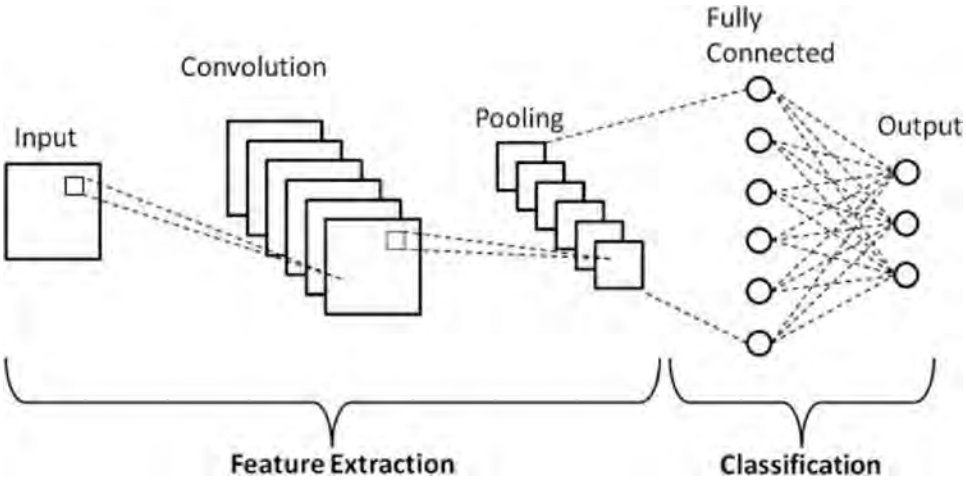
Figure 2. Architecture of RESTNET 50



CONVOLUTIONAL NEURAL NETWORK (CNN)

An analysis was done using CNN (Peng et al., 2019) to determine whether or not a face images are authentic or a hoax. To do this experiment data was gathered via Kaggle (2019) and the images for training are around 1450, and images for validation are around 300. A convolutional layer divides a given input into smaller parts, uses a kernel to perform a convolution operation that extracts facial patterns from each part, and finally outputs a feature map. A 3x3 kernel size is being considered for the proposed work. Typically, a pooling layer follows a convolutional layer to reduce complexity by down sampling the convolutional output using a pooling function like average pooling or max pooling, because the generated features from these layers can be quite large. In a CNN, the convolutional and pooling layers serve as the foundation for a classification layer that may include dropout layers as well as fully connected layers. Specifically, a dropout layer serves to prevent neural networks from becoming overfitted. The layer "drops out" a random group of neurons, causing the neural network to ensure that it can deliver proper classifications without relying too much on individual activated neurons, preventing it from in a CNN, the convolutional and pooling layers serve as the foundation for a classification layer that may include dropout layers as well as fully connected layers. Specifically, a dropout layer serves to prevent neural networks from becoming overfitted. The layer "drops out" a random group of neurons, causing the neural network to ensure that it can deliver proper classifications without relying too much on individual activated neurons, preventing it from memorizing the training data completely. Following figure 3 shows the architecture of Convolutional neural network.

Figure 3. Architecture of Convolutional neural network



SOLUTIONS AND RECOMMENDATIONS

Pre-trained Deep learning techniques were used in this experiment to improve the accuracy of Deep fake picture detection This work proposed to use MESONET, RESNET50 and CNN for detecting Deep fake images. Both custom data and data from the Kaggle Repository (2019) are used to train the pre-trained Deep Neural Network architectures. The custom data has 100 Deep fake images and 100 Real images for training and testing data. We have 2041 Kaggle Repository files, 960 of which are training false photos and 1081 of which are real photographs for testing. The images are provided as input to the pre-trained framework for training and then predicting. The performance of the model is later tested for validation. Using the performance metrices.

Verification of the proposed procedure is carried out using images from the dataset that have been meticulously altered and of extremely high quality. The images are composites of different faces, separated by the eyes, the nose, the mouth, or the whole of the facial structure are taken as input to the pre-trained models.

In the proposed work training and testing set for pre-trained models are split into three and two sets, respectively. There were two classes of images used throughout the training phase: "0" denoted the REAL image and "1" indicated the FAKE image, with X = "0" as the default. A set of images labelled with classes were fed into the pre-trained models, which generates the classifier models.

Pre-trained models such as MESONET, RESNET50 and CNN were developed by fine-tuning the learning parameters. The models were then tested using characteristics from the testing data set, and the proposed method was evaluated using standard metrics derived from confusion matrix. The following Table 1 shows the real and fake images that were identified by the pretrained models are a variety of computational metrics used to evaluate the pre-trained models' performance, including Accuracy, Specificity/Precision, RECALL, and F1-Score. True Positive (TP), True Negative (TN), and False Positive (FP) and False Negative (FN) images are used in the metrics for evaluation.

Table 1. Real and fake images

Real Images	Fake Images
	
	
	

In this work performance of the pre-trained models is determined by calculating its Accuracy of detecting Fake images from Real images. The following equation determines the Classification accuracy

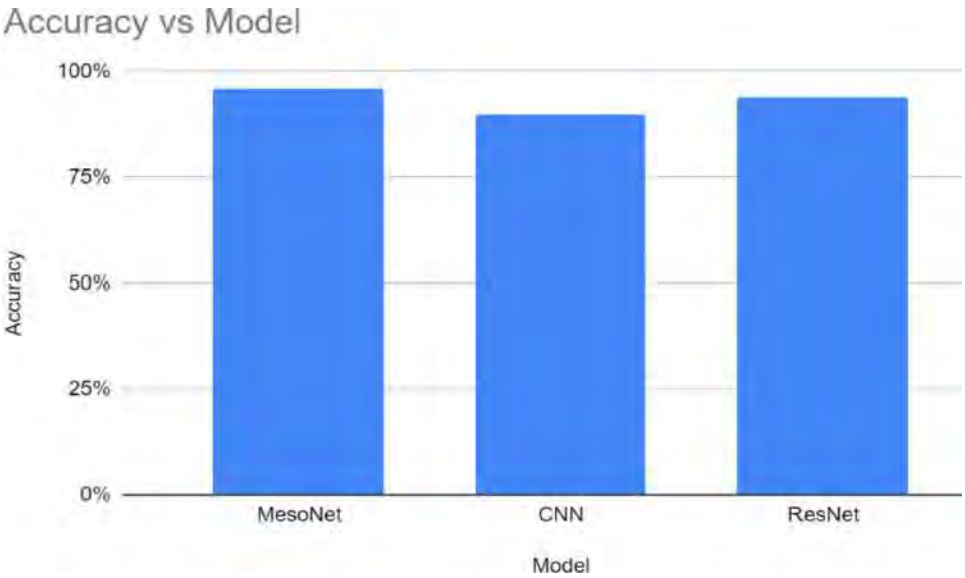
$$\text{Classification Accuracy} = \frac{(\text{TP} + \text{TN})}{(\text{TP} + \text{TN} + \text{FP} + \text{FN})} \tag{1}$$

The Classification Accuracy of the three pre-trained model is represented in the Table 2 and the graph the figure 4 pictorially represents the performance of each pre- trained models (MESONET, RESNET50 and CNN)

Table 2. Classification accuracy

Model	Accuracy
MesoNet	96%
CNN	90%
ResNet	94%

Figure 4. Accuracy of FAKE IMAGE DETECTION graph of different models



According to the findings, pre-trained models such as MESONET, RESENT50, and CNN do a superior job of distinguishing between fictitious and authentic images.

FUTURE RESEARCH DIRECTIONS

The further performance of Deep fake reduction can be improved by combining RESNET with LSTM and also Computer Vision algorithm can be applied for detection of Deep Fake images.

CONCLUSION

Face alteration in video is now widely acknowledged as a hazard. As a result of our research, we have used three pre-trained network architectures to detect forgeries quickly and cheaply. Deep networks are currently used to extract features in the majority of popular algorithms for detecting fraudulent face videos. In this study the paper highlights that pre-trained model such as MESONET.RESNET50 and CNN can be used in detecting the deep fake images. The experimental results prove that pre-trained models provide high accuracy in detecting the FAKE images from the given dataset.

REFERENCES

Afchar, D., Nozick, V., Yamagishi, J., & Echizen, I. (2018, December). Mesonet: a compact facial video forgery detection network. In 2018 IEEE international workshop on information forensics and security (WIFS) (pp. 1-7). IEEE.

Almars, A. M. (2021). Deepfakes detection techniques using deep learning: A survey. *Journal of Computer and Communications*, *9*(5), 20–35.

Bonettini, N., Cannas, E. D., Mandelli, S., Bondi, L., Bestagini, P., & Tubaro, S. (2021, January). Video face manipulation detection through ensemble of cnns. In *2020 25th international conference on pattern recognition (ICPR)* (pp. 5012-5019). IEEE.

Chen, L., Zhang, Y., Song, Y., Liu, L., & Wang, J. (2022). Self-supervised Learning of Adversarial Example: Towards Good Generalizations for Deepfake Detection. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition* (pp. 18710-18719). IEEE.

de Lima, O., Franklin, S., Basu, S., Karwoski, B., & George, A. (2020). *Deepfake detection using spatiotemporal convolutional networks.* arXiv preprint arXiv:2006.14749.

Durall, R., Keuper, M., Pfreundt, F. J., & Keuper, J. (2019). *Unmasking deepfakes with simple features.* arXiv preprint arXiv:1911.00686.

Gong, D., Yogan, J. K., Goh, O. S., Ye, Z., & Chi, W. (2021). DeepfakeNet, an efficient deepfake detection method. *International Journal of Advanced Computer Science and Applications*, *12*(6).

Goodfellow, I., Pouget-Abadie, J., Mirza, M., Xu, B., Warde-Farley, D., Ozair, S., ... Bengio, Y. (2014). Generative adversarial nets. *Advances in Neural Information Processing Systems*, 27.

He, K., Zhang, X., Ren, S., & Sun, J. (2016). Deep residual learning for image recognition. In *Proceedings of the IEEE conference on computer vision and pattern recognition* (pp. 770-778). IEEE.

Ioffe, S., & Szegedy, C. (2015, June). Batch normalization: Accelerating deep network training by reducing internal covariate shift. In *International conference on machine learning* (pp. 448-456). PMLR.

Li, L., Bao, J., Zhang, T., Yang, H., Chen, D., Wen, F., & Guo, B. (2020). Face x-ray for more general face forgery detection. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition* (pp. 5001-5010). IEEE.

Li, Y., Chang, M. C., & Lyu, S. (2018, December). In ictu oculi: Exposing ai created fake videos by detecting eye blinking. In *2018 IEEE International workshop on information forensics and security (WIFS)* (pp. 1-7). IEEE.

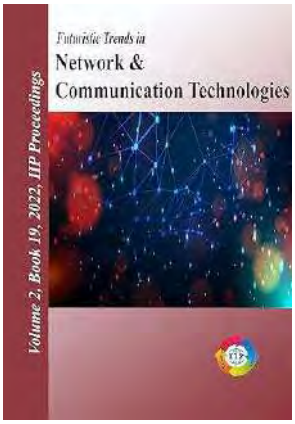
Nataraj, L., Mohammed, T. M., Manjunath, B. S., Chandrasekaran, S., Flenner, A., Bappy, J. H., & Roy-Chowdhury, A. K. (2019). Detecting GAN generated fake images using co-occurrence matrices. *Electronic Imaging*, *2019*(5), 532–1.

Peng, J., Kang, S., Ning, Z., Deng, H., Shen, J., Xu, Y., ... Liu, L. (2020). Residual convolutional neural network for predicting response of transarterial chemoembolization in hepatocellular carcinoma from CT imaging. *European Radiology*, *30*(1), 413–424.

Real and fake face detection. Kaggle. (2019, January 14). Retrieved from https://www.kaggle.com/datasets/ciplab/real-and-fake-face-detection?resource=download

Rossler, A., Cozzolino, D., Verdoliva, L., Riess, C., Thies, J., & Nießner, M. (2019). Faceforensics++: Learning to detect manipulated facial images. In *Proceedings of the IEEE/CVF international conference on computer vision* (pp. 1-11). IEEE.

Vamsi, V. V. V. N. S., Shet, S. S., Reddy, S. S. M., Rose, S. S., Shetty, S. R., Sathvika, S., & Shankar, S. P. (2022). *Deepfake Detection in Digital Media Forensics*. *Global Transitions Proceedings*.



Roll over image to zoom in

Futuristic Trends in Network & Communication Technologies

Paperback – 6 July 2023

by IIP Series Editors (Author)

See all formats and editions

Paperback
₹700.00

1 New from ₹700.00

Save Extra with 2 offers

Bank Offer (2): 7.5% Instant Discount up to INR 1000 on Bank of Baroda Credit Card EMI Trxns. Min purchase value INR 8000 | See All

Partner Offers: Get GST invoice and save up to 28% on business purchases. Sign up for free | Details

10 days Replacement | Amazon Delivered | Secure transaction

Computer network & communication book series provides a premier interdisciplinary platform for researchers, practitioners and educators to publish not only the most recent innovations, trends, and concerns but also practical challenges encountered and solutions adopted in the fields of networks and communication. This book will provide an excellent

Read more

ISBN-10	ISBN-13	Publisher
8195935613	978-8195935611	Iterative International Publisher IIP

Buy new: ₹700.00
Inclusive of all taxes

₹57 delivery **Sunday, 26 November**. Order within 13 hrs 25 mins. Details

Delivering to Bengaluru 560076 - Update location

In stock

Sold by Iterative International Publisher IIP and Delivered by Amazon.

Quantity: 1

Add to Cart

Buy Now

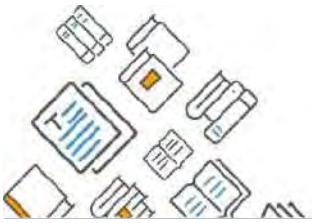
Secure transaction

Add to Wish List

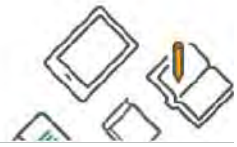
Product details

Publisher : Iterative International Publisher IIP (6 July 2023)
Paperback : 319 pages
ISBN-10 : 8195935613
ISBN-13 : 978-8195935611
Reading age : 16 years and up
Item Weight : 200 g
Dimensions : 22.86 x 17.78 x 1 cm
Country of Origin : India

How would you rate your experience shopping for books on Amazon today?



Very poor ---------- Neutral ---------- Great



Customer reviews

No customer reviews

5 star 0%

4 star 0%

3 star 0%

2 star 0%

1 star 0%

How are ratings calculated?

Review this product

Share your thoughts with other customers

Write a product review

Back to top

English

Australia Brazil Canada China France Germany Italy Japan Mexico Netherlands Poland Singapore Spain Turkey United Arab Emirates United Kingdom United States

AbeBooks
Books, art & collectibles

Amazon Web Services
Scalable Cloud Computing Services

Audible
Download Audio Books

IMDb
Movies, TV & Celebrities

Shopbop
Designer Fashion Brands

Amazon Business
Everything For Your Business

Prime Now
2-Hour Delivery on Everyday Items

Amazon Prime Music
100 million songs, ad-free
Over 15 million podcast episodes

Artificial Intelligence for Smart Cities and Villages: Advanced Technologies, Development, and Challenges

Back Book Details

Intelligent IoT Enabled Lighting System

Author(s): Dankan G. Veeranna*, Ramesha Muniyappa, Sridhara S. Boraiah, Devananda S. Narayanappa and Manoj K. S. Boraiah

Pp: 157-173 (17)

DOI: 10.2174/9789815049251122010012

Abstract

The term Smart is an acronym for the five components, which are unique, observable, practicable, appropriate and time-based. IoT identifies the widespread and increasing number of wireless devices as the trillions of networks of a potentially global scale running today. When the planet expands quicker, people are drawn to this clever term. India is one of the world's fastest-growing economies. The Internet of Things (IoT) offers innovative services and stimulates market shifts. The lighting industry often welcomes this transition by developing an Internet of Lighting (IoL). The term smart lighting describes three important design attributes: advanced control, solid-state, and broad network connectivity in accordance with international standards. It provides optical networking gateways as an additional function that enables coexistence with traditional Wi-Fi. In this context, various IoT-compatible wireless communication interfaces are being used to meet the lighting's needs. Related IoT systems are forecasted to expand dramatically in the entire smart housing and industrial sectors in the coming years. These instruments use advanced features, such as the use of multiple spectral light sources and transitive spectral light to provide. and they use RF and optical networking mainly for advanced sensing and more recently, they've started using the Wi-Fi interface as well. Visible Contact of Light. This chapter presents the numerous options for creating and implementing IoT-based smart lighting solutions in this chapter. As a result, issues related to interoperability and online services are addressed, and industrial lighting platforms are also included.

Keywords: Creative Network, Internet of Things, Lighting systems, Smart lighting, Visible light communications.

Related Journals

Micro and Nanosystems

View More »

Related Books

Advances in III-V Semiconductor Nanowires and Nanodevices

View More »

Artificial Intelligence for Smart Cities and Villages: Advanced Technologies, Development, and Challenges

Back Book Details

IoT Enabled Smart Village for Sustainable Development

Author(s): Dankan G. Veeranna*, Sridhara S. Boraiah, Ramesha Muniyappa, Manoj K. S. Boraiah and Devananda S. Narayanappa

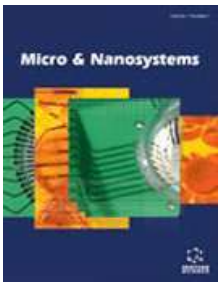
Pp: 194-215 (22)

DOI: 10.2174/9789815049251122010014

Abstract

In recent years, large-scale urbanisation has been on the rise, with cities being the hubs for growth, but rural business continues to play an important role in any country's overall development. A recent report reveals that almost 69% of India's population resides in their villages, accounting for almost 50% of the nation's GDP. This agricultural region varies from small towns with less than 500 residents to small towns. There are also similar situations in western nations. Despite their economic contribution, smaller villages seem to earn less in terms of infrastructural expenditure. The major problems addressed in these small towns include a shortage of adequate public transit, emergency care, and limited knowledge of federal subsidies for rural areas. Our vision is to expand the smart city to the smart village by allowing the use of recent technical advances and giving more attention to the problems in rural areas. This chapter aims to bring IoT technology to Villages by literally showing a network of linked sensors and knowledge dissemination devices, controlling energy use, and ensuring infrastructure protection. It gives a wide-ranging vision of enhancing the standard of living in villages and encourages them to meet the essential needs of domestic villagers.

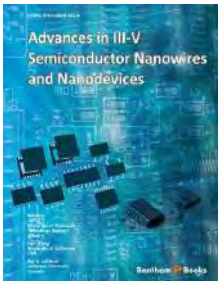
Keywords: Big data, Cloud computing, Information and communication technology, Internet of Things, Smart city, Smart village.



Micro and Nanosystems

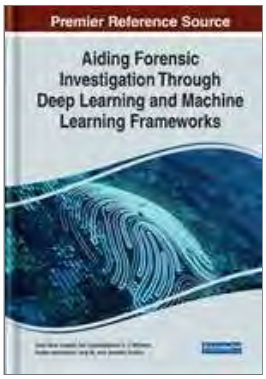
View More

Related Books



Advances in III-V Semiconductor Nanowires and Nanodevices

View More



Deep Learning-Based Automatic Student Authentication

Sheetalrani R. Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.

Source Title: Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks (/book/aiding-forensic-investigation-through-deep/290354)

Pages: 18

DOI: 10.4018/978-1-6684-4558-7.ch010

Abstract

Classroom management relies heavily on the ability to keep track of student attendance. Attendance checks by calling names or handing out a sign-in sheet are time-consuming and vulnerable to fraud, especially the latter. Data science and image processing are the focus of this regarding counting the number of people present at a gathering for various objectives, such as determining a person's duty status, determining a person's physical presence in a classroom, determining a person's security clearance to enter a meeting hall, etc. It takes a lot of time and effort to maintain a database for future use when generating attendance records using the standard technique. With the help of the latest technology, attendance can be entered automatically. It is a frequently used face recognition technique that generates a binary code for each cell and compares it to the reference image. With the use of deep learning, this LBP method has been reworked in order to automate attendance generation.

Chapter Preview

Top

Introduction

Humans have a certain amount of time in which to fulfill their goals. Spending time on mindless, mindless physical labour is a waste of time and a distraction from more important matters. Using Deep Neural Networks to solve the mass surveillance issue, offer an automation-based strategy to avoid such pointless effort, focusing primarily on the sub problem of enhancing teacher efficiency by automating the attendance taking process (Girshick et al., 2014). Intensely complex neural networks Using Deep Learning frameworks for face detection and recognition, this chapter proposes ways to lessen the burden on teachers by delivering fast and accurate results while also overcoming existing method shortcomings like multiclass identification problems, multiple appearances of the same person, illumination variance, and occlusion (Ren et al., 2015). Using deep learning, data is routed through several layers of non-linear mappings before the computer can construct a representation for it. Finally, Deep Learning Models are nothing but a function approximator capable of learning each and every function that depends on the input dataset. In the end, An extensive variety of subjects is included under the umbrella term “deep learning,” which includes anything from supervised regression models to reinforcement learning-based AI. In Deep Learning, the problem of face detection and recognition is extremely challenging and contentious (Redmon et al., 2016). Face recognition on a broad scale has proven difficult due to a scarcity of properly labelled training data. Current state-of-the-art methods have been pushed aside by Nvidia Graphics cards, which can perform matrix operations very quickly. For a long time, Image and Vision Computing was dominated by non-deep learning algorithms as Viola Jones (Viola and Jones, 2004). Clustering and HOG, but noise in the data, such as occlusion and luminance variation, has greatly hindered these methods. Before Deep Learning, only Feature Engineering, Data Preprocessing (Zhu et al., 2006), Output Parsing, and KDD Tree Structures could provide an acceptable result in face recognition. RCNN, Faster and Faster RCNN, Overfeat, and Yolo can now operate in near-real time on commercial workstations because of the improved processing capability. The application of Deep Learning techniques for image and vision computing skyrocketed with the invention of Convolutional Neural Networks (Ding et al., 2014). For example, in Deep Learning, the use of data augmentation procedures during training may ensure that the model is invariant to noise and has a good generalizability so that it can overcome challenges with noise and erroneous predictions caused by things like occlusion.

There are two parts to a Face Recognition challenge in general: detecting a face in the picture and accurately classifying the image (Krizhevsky et al., 2012). To put it another way, we're dealing with an issue with the detection and classification of objects. How can an object detection framework be described? First, the neural network must be able to locate the item to be identified in an image, and secondly, it must be able to predict with high certainty that this object belongs to a certain class. There are a variety of ways to solve the issue, but the two most common are branch networks and a single network. Multiple Neural Network topologies are used in Branch Networks to conduct the many subtasks necessary to complete the pen-ultimate task, as the name indicates In nature, the branching sub-networks are decoupled, and each of them has been taught to mimic a unique function (Liu et al., 2016). Subnetworks that focus on the issue of region proposal are used in object detection networks. This means that the network is trying to learn which area in the picture is most likely to contain an item of interest. Classification of the observed item is taught to the second branching sub-network. The proposal of (Lowe, 2004) R-CNN with deep learning features a rich hierarchical feature set for accurate object detection and semantic segmentation. Region extraction is the initial phase of R-CNN, which may be done using approaches like objectness and selective search, limited parametricmin cuts, and so on. Selective search has been utilised to obtain region ideas for all of the experiment's goals.

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=309770&ptid=290354&ctid=15&t=Table of Contents&isxn=9781668445587)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=309771&ptid=290354&ctid=15&t=Detailed Table of Contents&isxn=9781668445587)

Preface — View Full PDF (/pdf.aspx?tid=309772&ptid=290354&ctid=15&t=Preface&isxn=9781668445587)

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh, Ruban Nersisson, Ang Yu, Jennifer Gentry

Acknowledgment — View Full PDF (/pdf.aspx?tid=309773&ptid=290354&ctid=15&t=Acknowledgment&isxn=9781668445587)

Chapter 1

A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework (/chapter/a-review-on-forensic-science-and-criminal-investigation-through-a-deep-learning-framework/309774) (pages 1-72)

Pinaki Pratim Acharjya, Santanu Koley, Subhabrata Barman

Preview Chapter (/viewtitlesample.aspx?id=309774&ptid=290354&t=A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework&isxn=9781668445587) **$37.50** Add to Cart

Chapter 2

Digital Forensic Techniques and Principles in a Cloud Environment (/chapter/digital-forensic-techniques-and-principles-in-a-cloud-environment/309775) (pages 73-91)

Ayaz Ahmad, Supriya Prashant Diwan, Sudhansu Sekhar Nanda, Lingaraj Sethi, Prashanta Kumar Patra

Preview Chapter (/viewtitlesample.aspx?id=309775&ptid=290354&t=Digital Forensic Techniques and Principles in a Cloud Environment&isxn=9781668445587) **$37.50** Add to Cart

Chapter 3

IoT-Deep Learning-Based Detection of Cyber Security Threats (/chapter/iot-deep-learning-based-detection-of-cyber-security-threats/309776) (pages 92-111)

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech, Ansuman Samal

Preview Chapter (/viewtitlesample.aspx?id=309776&ptid=290354&t=IoT-Deep Learning-Based Detection of Cyber Security Threats&isxn=9781668445587) **$37.50** Add to Cart

Chapter 4

Deep Learning Models for Cyber Security in IoT Networks (/chapter/deep-learning-models-for-cyber-security-in-iot-networks/309777) (pages 112-127)

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech, Thangadurai N.

Preview Chapter (/viewtitlesample.aspx?id=309777&ptid=290354&t=Deep Learning Models for Cyber Security in IoT Networks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 5

Deep Learning-Based AI Modeling, Intrusion Detection (/chapter/deep-learning-based-ai-modeling-intrusion-detection/309778) (pages 128-143)

Madhab Paul Choudhury, Madhab Paul Choudhury, Chandrashekhar Azad

Preview Chapter (/viewtitlesample.aspx?id=309778&ptid=290354&t=Deep Learning-Based AI Modeling, Intrusion Detection&isxn=9781668445587) **$37.50** Add to Cart

Chapter 6

Face Detection and Recognition From Distance Based on Deep Learning (/chapter/face-detection-and-recognition-from-distance-based-on-deep-learning/309779) (pages 144-160)

Hui Wang, Wei Qi Yan

Preview Chapter (/viewtitlesample.aspx?id=309779&ptid=290354&t=Face Detection and Recognition From Distance Based on Deep Learning&isxn=9781668445587) **$37.50** Add to Cart

Chapter 7

Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks (/chapter/investigation-of-deep-fake-images-using-pre-trained-cnn-frameworks/309780) (pages 161-173)

Anitha Ruth J., Uma R., Vijayalakshmi G. V. Mahesh, P. Ramkumar

Preview Chapter (/viewtitlesample.aspx?id=309780&ptid=290354&t=Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 8

Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics (/chapter/implementation-of-machine-learning-aided-speech-analysis-for-speaker-accent-identification-applied-to-audio-forensics/309781) (pages 174-194)

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Preview Chapter (/viewtitlesample.aspx?id=309781&ptid=290354&t=Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics&isxn=9781668445587) **$37.50** Add to Cart

Chapter 9

Medical Image Encryption Using Distance-Based Biokey Generation (/chapter/medical-image-encryption-using-distance-based-biokey-generation/309782) (pages 195-218)

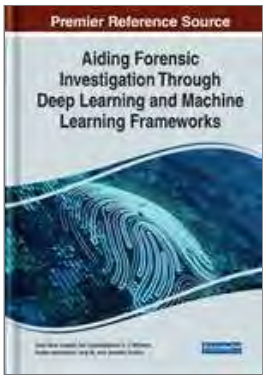
Vijayarajan Rajangam, Shani S. Kumar, Sathiya Narayanan, Sangeetha N., Avudaiammal R.

Preview Chapter (/viewtitlesample.aspx?id=309782&ptid=290354&t=Medical Image Encryption Using Distance-Based Biokey Generation&isxn=9781668445587) **$37.50** Add to Cart

Chapter 10

Deep Learning-Based Automatic Student Authentication (/chapter/deep-learning-based-automatic-student-authentication/309783) (pages 219-236)

Sheetalrani R Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.



IoT-Deep Learning-Based Detection of Cyber Security Threats

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech (/affiliate/pankaj-dadheech/428137/), Ansuman Samal

Source Title: Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks (/book/aiding-forensic-investigation-through-deep/290354)

Pages: 20

DOI: 10.4018/978-1-6684-4558-7.ch003

Abstract

The internet of things (IoT) is a revolutionary technology that links living and non-living devices all around the world. As a result, the frequency of cyber-attacks against IoT deployments is expected to rise. As a result, each system must be absolutely secure; otherwise, consumers may opt not to utilize the technology. DDoS assaults that recently attacked various IoT networks resulted in massive losses. There is only one way to detect stolen data from software and malware on the IoT network that is discussed in this chapter. To categorize stolen programming with source code literary theft, the tensor flow deep neural system is offered. To communicate raucous information while simultaneously emphasizing the value of each token in terms of source code forgery, tokenization, and measurement, the malware samples were gathered using the Malign dataset. The results show that the methodology proposed for analyzing IoT cyber security threats has a higher classification efficiency than current methodologies.

Chapter Preview

Top

Introduction

IoT expansion has expanded significantly in social orders worldwide lately. As how much interconnected IoT hardware came to $27 billion out of 2017, these IoT gear will increment dramatically with buyer interest, with a limit expected to reach $125 billion by 2030. Explicit clever city applications are connected to monstrous, true gadgets, which actually have extremely critical metropolitan advantages. The enormous quantities of Iot gadgets in different assistance types, structures, applications and conventions (e.g., WIFI, wired, versatile, cell, Bluetooth) add to the test of potential IoT network the board. There are accordingly not kidding digital protection dangers and weaknesses to these Internet reconciliation conventions to assault data about ordinary resident exercises (Srinivasan et al., 2019). Such digital dangers may not be permitted on the LOT framework except if the approved customer or leader (for instance Miria botnet) is familiar with them. In splendid city adventures, there are two significant security issues. The chief test is the means by which to recognize zero-day attacks from a combination of IoT shows in the sharp cloud server ranch assuming that critical perils are stayed away from IoT systems (Joyia et al., 2017). The second is the best way to deal with recognize advanced attacks adroitly (for instance IoT malware attacks, etc.) in the IoT arrange already obliterating a shrewd local area. Today, most IoT sensors gather all data through the enormous measure of information gathered on cloud servers (Zanella et al., 2014).

As of now, the IOT network frameworks have insignificant assets and less elements (for example shrewd watches, keen lights, savvy locks, and so on) which are not utilized by IOT network applications. An amazing new technology known as the Thing Internet ties the whole world to the Internet. Personal, professional, and social goals may all be met with the support of Internet of Things (IoT) solutions (Karbab et al., 2017). The Internet of Things (IoT) is a global network of intelligent devices that does not need human involvement, which is great but it is susceptible to cyber-attacks. Intruder detection systems are an essential part of any network's defense against cyber attacks (IDS). A machine-learning network is used by many of the latest IDS to help them prepare for and identify cyber attacks (Jabbar et al., 2018). Centralized distributed computing development is facilitated by fog computing, which tackles scalability obstacles, poorer QoS (Quality of Service) high transmission capacity utilisation and distributed computing's high inactivity need in the IoT framework by using mist-based distributed computing. Consolidating and achieving IoT arrangements effectively is possible with the use of Mist to-hub. Figure 1 depicts the mist to-hub model design with applicable equal figuring for offering IDS measures, controls, and stocking closer to IoT network arte facts for circulated mist.

Figure 1. IoT network design based on fog-to-node communication


978-1-6684-4558-7.ch003.f01(https://igiprodst.blob.core.windows.net:443/source-content/9781668445587_290354/978-1-6684-4558-7.ch003.f01.png?sv=2015-12-11&sr=c&sig=qUbswynDLt2NUwIEMy4AA9wZzMQCtkUq23mqLyqcyGU%3D&se=2022-09-23T04%3A24%3A59Z&sp=r)

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=309770&ptid=290354&ctid=15&t=Table of Contents&isxn=9781668445587)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=309771&ptid=290354&ctid=15&t=Detailed Table of Contents&isxn=9781668445587)

Preface — View Full PDF (/pdf.aspx?tid=309772&ptid=290354&ctid=15&t=Preface&isxn=9781668445587)

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh, Ruban Nersisson, Ang Yu, Jennifer Gentry

Acknowledgment — View Full PDF (/pdf.aspx?tid=309773&ptid=290354&ctid=15&t=Acknowledgment&isxn=9781668445587)

Chapter 1

A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework (/chapter/a-review-on-forensic-science-and-criminal-investigation-through-a-deep-learning-framework/309774) (pages 1-72)

Pinaki Pratim Acharjya, Santanu Koley, Subhabrata Barman

Preview Chapter (/viewtitlesample.aspx?id=309774&ptid=290354&t=A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework&isxn=9781668445587) **$37.50** Add to Cart

Chapter 2

Digital Forensic Techniques and Principles in a Cloud Environment (/chapter/digital-forensic-techniques-and-principles-in-a-cloud-environment/309775) (pages 73-91)

Ayaz Ahmad, Supriya Prashant Diwan, Sudhansu Sekhar Nanda, Lingaraj Sethi, Prashanta Kumar Patra

Preview Chapter (/viewtitlesample.aspx?id=309775&ptid=290354&t=Digital Forensic Techniques and Principles in a Cloud Environment&isxn=9781668445587) **$37.50** Add to Cart

Chapter 3

IoT-Deep Learning-Based Detection of Cyber Security Threats (/chapter/iot-deep-learning-based-detection-of-cyber-security-threats/309776) (pages 92-111)

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech, Ansuman Samal

Preview Chapter (/viewtitlesample.aspx?id=309776&ptid=290354&t=IoT-Deep Learning-Based Detection of Cyber Security Threats&isxn=9781668445587) **$37.50** Add to Cart

Chapter 4

Deep Learning Models for Cyber Security in IoT Networks (/chapter/deep-learning-models-for-cyber-security-in-iot-networks/309777) (pages 112-127)

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech, Thangadurai N.

Preview Chapter (/viewtitlesample.aspx?id=309777&ptid=290354&t=Deep Learning Models for Cyber Security in IoT Networks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 5

Deep Learning-Based AI Modeling, Intrusion Detection (/chapter/deep-learning-based-ai-modeling-intrusion-detection/309778) (pages 128-143)

Madhab Paul Choudhury, Madhab Paul Choudhury, Chandrashekhar Azad

Preview Chapter (/viewtitlesample.aspx?id=309778&ptid=290354&t=Deep Learning-Based AI Modeling, Intrusion Detection&isxn=9781668445587) $37.50 Add to Cart

Chapter 6

Face Detection and Recognition From Distance Based on Deep Learning (/chapter/face-detection-and-recognition-from-distance-based-on-deep-learning/309779) (pages 144-160)

Hui Wang, Wei Qi Yan

Preview Chapter (/viewtitlesample.aspx?id=309779&ptid=290354&t=Face Detection and Recognition From Distance Based on Deep Learning&isxn=9781668445587) $37.50 Add to Cart

Chapter 7

Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks (/chapter/investigation-of-deep-fake-images-using-pre-trained-cnn-frameworks/309780) (pages 161-173)

Anitha Ruth J., Uma R., Vijayalakshmi G. V. Mahesh, P. Ramkumar

Preview Chapter (/viewtitlesample.aspx?id=309780&ptid=290354&t=Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks&isxn=9781668445587) $37.50 Add to Cart

Chapter 8

Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics (/chapter/implementation-of-machine-learning-aided-speech-analysis-for-speaker-accent-identification-applied-to-audio-forensics/309781) (pages 174-194)

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Preview Chapter (/viewtitlesample.aspx?id=309781&ptid=290354&t=Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics&isxn=9781668445587) $37.50 Add to Cart

Chapter 9

Medical Image Encryption Using Distance-Based Biokey Generation (/chapter/medical-image-encryption-using-distance-based-biokey-generation/309782) (pages 195-218)

Vijayarajan Rajangam, Shani S. Kumar, Sathiya Narayanan, Sangeetha N., Avudaiammal R.

Preview Chapter (/viewtitlesample.aspx?id=309782&ptid=290354&t=Medical Image Encryption Using Distance-Based Biokey Generation&isxn=9781668445587) $37.50 Add to Cart

Chapter 10

Deep Learning-Based Automatic Student Authentication (/chapter/deep-learning-based-automatic-student-authentication/309783) (pages 219-236)

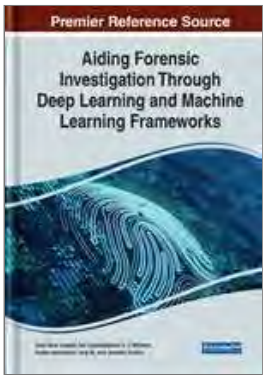
Sheetalrani R Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.

Preview Chapter (/viewtitlesample.aspx?id=309783&ptid=290354&t=Deep Learning-Based Automatic Student Authentication&isxn=9781668445587) $37.50 Add to Cart

About the Contributors

View Full PDF (/pdf.aspx?tid=309785&ptid=290354&ctid=17&t=About the Contributors&isxn=9781668445587)

Index



Deep Learning Models for Cyber Security in IoT Networks

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech (/affiliate/pankaj-dadheech/428137/), Thangadurai N.

Source Title: Aiding Forensic Investigation Through Deep Learning and Machine Learning Frameworks (/book/aiding-forensic-investigation-through-deep/290354)

Pages: 16

DOI: 10.4018/978-1-6684-4558-7.ch004

Abstract

As the number of connected devices grows, the internet of things (IoT) poses new security challenges for network activity monitoring. Due to a lack of security understanding on the side of device producers and end users, the majority of internet of things devices are vulnerable. As a result, virus writers have found them to be great targets for converting them into bots and using them to perform large-scale attacks against a variety of targets. The authors provide deep learning models based on deep reinforcement learning for cyber security in IoT networks in this chapter. The IoT is a potential network that connects both living and nonliving things all around the world. As the popularity of IoT grows, cyber security remains a shortcoming, rendering it exposed to a variety of cyber-attacks. It should be emphasized, however, that while there are numerous DL algorithms presently, the scientific literature does not yet include a comprehensive catalogue of them. This chapter provides a complete list of DL algorithms as well as their many application areas.

Chapter Preview

Top

Introduction

The newest developing promising technology known as the Internet of Things links everything on Earth via the internet. IoT (the technology) has the potential to enhance and assist in both our personal and professional lives, as well as our societal well-being. The internet of things, otherwise known as IoT, includes millions of smart devices all over the globe, all linked via the internet without human intervention. Unfortunately, because of this, it is vulnerable to cyber assaults just like any other network. A powerful tool for detecting cyber-attacks is an intrusion detection system (IDS) (Roopak et al., 2019). Machine learning has formed the core of all current IDS (including the state-of-the-art models). Fog computing extends cloud computing, where fog nodes are placed closer to IoT network devices to increase scalability, lower bandwidth usage, lower latency in QoS and overall enhance overall functionality. In fog-to-node computing, real IoT networks may achieve success. Figure 1 shows the architecture of a fog-to-node model, where fog nodes are dispersed and each node hosts distributed parallel computing that provides distributed fog systems with processing, control, and storing of IDS components that are more central to IoT network objects. In contrast to cloud services, fog nodes can identify a cyber-attack effectively and promptly. Internet of Things (IoT) refers to a collection of diverse technologies ranging from supercomputers to small devices with extremely little computational capacity, therefore protecting an IoT network poses major security challenges (Yuan et al., 2017). DDoS was a significant cyber-attack on the Internet of Things (IoT) that caused considerable damages. Attacks against web servers often use many hosts to overwhelming target systems, resulting in a total crash. This causes genuine users to be unable to use the system. It is expected that the denial-of-service attack would grow to around 17 million by 2020 as predicts. DL is a wider subfield of machine learning known as machine learning and machine learning (ML) is deeper than that, it is in fact a much larger network known as a deep neural network. This research, published in 2016, shows that deep learning methods such as deep belief networks and also convolutional neural networks (CNN) were developed in the 1980s and have since proved successful in a variety of areas such as image processing, natural language processing, and self-driving cars. One of the main challenges of deep learning techniques is the bigger the training dataset, the longer the training period.

Figure 1. Depicts the architecture of an Internet of Things network from fog to node.
978-1-6684-4558-7.ch004.f01(https://igiprodst.blob.core.windows.net:443/source-content/9781668445587_290354/978-1-6684-4558-7.ch004.f01.png?sv=2015-12-11&sr=c&sig=2qh%2F5gAair%2FdEYiy%2Fd7gHx19tumxv9koH5784%2BQwZyU%3D&se=2022-09-23T04%3A24%3A28Z&sp=r)

Figure. 2 shows the fundamental architecture of a deep learning model, which consists of one input layer followed by several hidden layers. The last layer of the model has one input and produces an output. CNN is a deep learning model that has been widely used in image and language processing (Sharif Razavian et al., 2014). The raw picture is immediately supplied to CNN models without any pre-processing, and the model then does various convolution operations on the image data (Hao et al., 2018). In the areas of NLP (natural language processing) and text processing, RNNs (Recurrent neural networks) have shown positive results. LSTM (Long Short-Term Memory) is a kind of RNN. Because LSTM is applied directly to raw data, it has a number of advantages.

Figure 2. The Deep Learning Model's Architectural Design

 *978-1-6684-4558-7.ch004.f02(https://igiprodst.blob.core.windows.net:443/source-content/9781668445587_290354/978-1-6684-4558-7.ch004.f02.png?sv=2015-12-11&sr=c&sig=2qh%2F5gAair%2FdEYiy%2Fd7gHx19tumxv9koH5784%2BQwZyU%3D&se=2022-09-23T04%3A24%3A28Z&sp=r)*

Complete Chapter List

Search this Book: Full text search terms

Reset

Table of Contents — View Full PDF (/pdf.aspx?tid=309770&ptid=290354&ctid=15&t=Table of Contents&isxn=9781668445587)

Detailed Table of Contents — View Full PDF (/pdf.aspx?tid=309771&ptid=290354&ctid=15&t=Detailed Table of Contents&isxn=9781668445587)

Preface — View Full PDF (/pdf.aspx?tid=309772&ptid=290354&ctid=15&t=Preface&isxn=9781668445587)

Alex Noel Joseph Raj, Vijayalakshmi G. V. Mahesh, Ruban Nersisson, Ang Yu, Jennifer Gentry

Acknowledgment — View Full PDF (/pdf.aspx?tid=309773&ptid=290354&ctid=15&t=Acknowledgment&isxn=9781668445587)

Chapter 1

A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework (/chapter/a-review-on-forensic-science-and-criminal-investigation-through-a-deep-learning-framework/309774) (pages 1-72)

Pinaki Pratim Acharjya, Santanu Koley, Subhabrata Barman

Preview Chapter (/viewtitlesample.aspx?id=309774&ptid=290354&t=A Review on Forensic Science and Criminal Investigation Through a Deep Learning Framework&isxn=9781668445587) **$37.50** Add to Cart

Chapter 2

Digital Forensic Techniques and Principles in a Cloud Environment (/chapter/digital-forensic-techniques-and-principles-in-a-cloud-environment/309775) (pages 73-91)

Ayaz Ahmad, Supriya Prashant Diwan, Sudhansu Sekhar Nanda, Lingaraj Sethi, Prashanta Kumar Patra

Preview Chapter (/viewtitlesample.aspx?id=309775&ptid=290354&t=Digital Forensic Techniques and Principles in a Cloud Environment&isxn=9781668445587) **$37.50** Add to Cart

Chapter 3

IoT-Deep Learning-Based Detection of Cyber Security Threats (/chapter/iot-deep-learning-based-detection-of-cyber-security-threats/309776) (pages 92-111)

Ramesh Naidu P., Dankan Gowda V., Kumaraswamy S., Pankaj Dadheech, Ansuman Samal

Preview Chapter (/viewtitlesample.aspx?id=309776&ptid=290354&t=IoT-Deep Learning-Based Detection of Cyber Security Threats&isxn=9781668445587) **$37.50** Add to Cart

Chapter 4

Deep Learning Models for Cyber Security in IoT Networks (/chapter/deep-learning-models-for-cyber-security-in-iot-networks/309777) (pages 112-127)

Dankan Gowda V., Puneeth Kumar B. S., Shekhar R., Pankaj Dadheech, Thangadurai N.

Preview Chapter (/viewtitlesample.aspx?id=309777&ptid=290354&t=Deep Learning Models for Cyber Security in IoT Networks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 5

Deep Learning-Based AI Modeling, Intrusion Detection (/chapter/deep-learning-based-ai-modeling-intrusion-detection/309778) (pages 128-143)

Madhab Paul Choudhury, Madhab Paul Choudhury, Chandrashekhar Azad

Preview Chapter (/viewtitlesample.aspx?id=309778&ptid=290354&t=Deep Learning-Based AI Modeling, Intrusion Detection&isxn=9781668445587) **$37.50** Add to Cart

Chapter 6

Face Detection and Recognition From Distance Based on Deep Learning (/chapter/face-detection-and-recognition-from-distance-based-on-deep-learning/309779) (pages 144-160)

Hui Wang, Wei Qi Yan

Preview Chapter (/viewtitlesample.aspx?id=309779&ptid=290354&t=Face Detection and Recognition From Distance Based on Deep Learning&isxn=9781668445587) **$37.50** Add to Cart

Chapter 7

Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks (/chapter/investigation-of-deep-fake-images-using-pre-trained-cnn-frameworks/309780) (pages 161-173)

Anitha Ruth J., Uma R., Vijayalakshmi G. V. Mahesh, P. Ramkumar

Preview Chapter (/viewtitlesample.aspx?id=309780&ptid=290354&t=Investigation of Deep Fake Images Using Pre-Trained CNN Frameworks&isxn=9781668445587) **$37.50** Add to Cart

Chapter 8

Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics (/chapter/implementation-of-machine-learning-aided-speech-analysis-for-speaker-accent-identification-applied-to-audio-forensics/309781) (pages 174-194)

Vijayalakshmi G. V. Mahesh, Alex Noel Joseph Raj, Ruban Nersisson

Preview Chapter (/viewtitlesample.aspx?id=309781&ptid=290354&t=Implementation of Machine Learning-Aided Speech Analysis for Speaker Accent Identification Applied to Audio Forensics&isxn=9781668445587) **$37.50** Add to Cart

Chapter 9

Medical Image Encryption Using Distance-Based Biokey Generation (/chapter/medical-image-encryption-using-distance-based-biokey-generation/309782) (pages 195-218)

Vijayarajan Rajangam, Shani S. Kumar, Sathiya Narayanan, Sangeetha N., Avudaiammal R.

Preview Chapter (/viewtitlesample.aspx?id=309782&ptid=290354&t=Medical Image Encryption Using Distance-Based Biokey Generation&isxn=9781668445587) **$37.50** Add to Cart

Chapter 10

Deep Learning-Based Automatic Student Authentication (/chapter/deep-learning-based-automatic-student-authentication/309783) (pages 219-236)

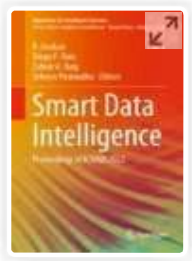
Sheetalrani R Kawale, Swetha K. R., Swathi Pai M., Namitha A. R., Dankan Gowda V.

Preview Chapter (/viewtitlesample.aspx?id=309783&ptid=290354&t=Deep Learning-Based Automatic Student Authentication&isxn=9781668445587) **$37.50**

Add to Cart

About the Contributors — View Full PDF (/pdf.aspx?tid=309785&ptid=290354&ctid=17&t=About the Contributors&isxn=9781668445587)

Index — View Full PDF (/pdf.aspx?tid=309786&ptid=290354&ctid=17&t=Index&isxn=9781668445587)



Smart Data Intelligence pp 343–350

Doppler Radar Technique for Geriatric Fall Detection

N. Rashmi & K. R. Mamatha

Conference paper | First Online: 18 August 2022

Part of the Algorithms for Intelligent Systems book series (AIS)

Abstract

Elderly falls have long been one of the most critical challenges affecting the elderly, as well as an important research topic in the domain of life signs. Their lives have been significantly impacted by the injuries. Furthermore, if the person continues lying for a prolonged period of time after falling, the risk of significant consequences increases. Such issues should be prevented in the first place. As a result, it is

critical in the field of life science to progress a methodology for studying fall detection and prevention. This paper explains how to detect human motion, which causes a frequency change in the produced radar echo signal and then calculates the Doppler fingerprints. Using the characteristic of the micro-Doppler signals, the resultant Doppler measurements are used to discriminate between diverse human movements and gestures. These data will subsequently be utilized to detect and range humans, as well as to differentiate between different human movements at various ranges.

Keywords

Doppler signatures **Fall detection** **Life signs**

Micro-doppler signals **Radar echo signal**

References

1. Alnaeb A, Syamsul R, Abdullah AR, Salah A, Sali A, Rashid NEA, Pasya I (2018) Human activities detection and classification based on micro-Doppler signatures near the baseline of forward scattering radar. In: 2018 international conference on radar (RADAR). IEEE, pp 1–6

2. Javier RJ, Kim Y (2014) Application of linear predictive coding for human activity classification based on micro-Doppler signatures. IEEE Geosci Remote Sens Lett 11(10):1831–1834

3. Adhikari K, Bouchachia H, Nait-Charif H (2017) Activity recognition for indoor fall detection using convolutional neural network. In: 2017 fifteenth IAPR international conference on machine vision applications (MVA). IEEE, pp 81–84

4. Selvabala VSN, Ganesh AB (2012) Implementation of wireless sensor network based human fall detection system. Procedia Eng 30:767–773

5. Tekeli B, Gurbuz SZ, Yuksel M (2016) Information-theoretic feature selection for human micro-Doppler signature classification. IEEE Trans Geosci Remote Sens 54(5):2749–2762

6. Noury N, Fleury A, Rumeau P, Bourke AK, Laighin GO, Rialle V, Lundy JE (2007) Fall detection-principles and methods. In: 2007 29th annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 1663–1666

7. Amin MG, Zhang YD, Ahmad F, Ho KD (2016) Radar signal processing for elderly fall detection: the future for in-home monitoring. IEEE Signal Process Mag 33(2):71–80

8. Mubashir M, Shao L, Seed L (2013) A survey on fall detection: principles and approaches. Neurocomputing 100:144–152

9. Pannurat N, Thiemjarus S, Nantajeewarawat E (2014) Automatic fall monitoring: a review. Sensors 14(7):12900–12936

10. Cleland I, Kikhia B, Nugent C, Boytsov A, Hallberg J, Synnes K et al (2013) Optimal placement of accelerometers for the detection of everyday activities. Sensors 13(7):9183–9200

11. Mathie MJ, Coster AC, Lovell NH, Celler BG (2004) Accelerometry: providing an integrated, practical method for long-term, ambulatory monitoring of human movement. Physiol Meas 25(2):R1

12. Wu M, Dai X, Zhang YD, Davidson B, Amin MG, Zhang J (2013) Fall detection based on sequential modeling of radar signal time-frequency features. In: 2013 IEEE international conference on healthcare informatics. IEEE, pp 169–174

13. Li Y, Ho KC, Popescu M (2012) A microphone array system for automatic fall detection. IEEE Trans Biomed Eng 59(5):1291–1301

14. Kwolek B, Kepski M (2014) Human fall detection on embedded platform using depth maps and wireless accelerometer. Comput Methods Programs Biomed 117(3):489–501

15. De Miguel K, Brunete A, Hernando M, Gambao E (2017) Home camera-based fall detection system for the elderly. Sensors 17(12):2864

16. Yu X (2008) Approaches and principles of fall detection for elderly and patient. In: HealthCom 2008–10th international conference on e-health networking, applications and services. IEEE, pp 42–47

17. Yu X (2008) Approaches and principles of fall detection for elderly and patient. In: HealthCom 2008–10th international conference on e-health networking, applications and services. IEEE, pp 42–47

18. Xiang Y, Savarese S (2013) Object detection by 3d aspectlets and occlusion reasoning. In:

Proceedings of the IEEE international conference on computer vision workshops, pp 530–537

19. Gadde A, Amin MG, Zhang YD, Ahmad F (2014) Fall detection and classifications based on time-scale radar signal characteristics. In: Radar sensor technology XVIII, vol 9077. SPIE, pp 330–338

20. Chen VC, Li F, Ho SS, Wechsler H (2006) Micro-Doppler effect in radar: phenomenon, model, and simulation study. IEEE Trans Aerosp Electron Syst 42(1):2–21

21. Wang F, Skubic M, Rantz M, Cuddihy PE (2014) Quantitative gait measurement with pulse-Doppler radar for passive in-home gait assessment. IEEE Trans Biomed Eng 61(9):2434–2443

22. Thayaparan T, Stanković LJ, Djurović I (2008) Micro-Doppler-based target detection and feature extraction in indoor and outdoor environments. J Franklin Inst 345(6):700–722

23. Liu L, Popescu M, Ho KC, Skubic M, Rantz M (2012) Doppler radar sensor positioning in a fall

detection system. In: 2012 annual international conference of the IEEE engineering in medicine and biology society. IEEE, pp 256–259

24. Jacob IJ, EbbyDarney P (2021) Design of deep learning algorithm for IoT application by image based recognition. J ISMAC 3(3):276–290

25. Balasubramaniam V (2020) IoT based biotelemetry for smart health care monitoring system. J Inf Technol Digital World 2(3):183–190

Author information

Authors and Affiliations

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, India

N. Rashmi & K. R. Mamatha

Corresponding author

Correspondence to N. Rashmi .

Editor information

Editors and Affiliations

Kongunadu College of Engineering and Technology, Tholurpatti, India

Dr. R. Asokan

School of Information Technology and Telecommunications Engineering, University of Granada, Granada, Spain

Dr. Diego P. Ruiz

Deakin University, Waurn Ponds, VIC, Australia

Dr. Zubair A. Baig

Information Systems and Operations Management (ISOM), University of Florida, Gainesville, FL, USA

Dr. Selwyn Piramuthu

Rights and permissions

Copyright information

About this paper

Cite this paper

Rashmi, N., Mamatha, K.R. (2022). Doppler Radar Technique for Geriatric Fall Detection. In: Asokan, R., Ruiz, D.P., Baig, Z.A., Piramuthu, S. (eds) Smart Data Intelligence. Algorithms for Intelligent Systems. Springer, Singapore. https://doi.org/10.1007/978-981-19-3311-0_29

.RIS .ENW .BIB

DOI	Published	Publisher Name
	18 August 2022	

https://doi.org/10.1007/978-981-19-3311-0_29

Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-3310-3	978-981-19-3311-0	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)

Printed Antennas (https://www.taylorfrancis.com/books/mono/10.1201/9781003347057/printed-antennas?refId=469218de-c0d3-4d11-bc24-432b1b615b66&context=ubx)

Show Path

Chapter

SRR-loaded octagonal Sierpinski-based carpet-shaped antenna for multiband application

By K. Yogaprasad (/search?contributorName=K. Yogaprasad&contributorRole=author&redirectFromPDP=true&context=ubx), M. Nanda Kumar (/search?contributorName=M. Nanda Kumar&contributorRole=author&redirectFromPDP=true&context=ubx), V. R. Anitha (/search?contributorName=V. R. Anitha&contributorRole=author&redirectFromPDP=true&context=ubx), Anil Kumar Nayak (/search?contributorName=Anil Kumar Nayak&contributorRole=author&redirectFromPDP=true&context=ubx)

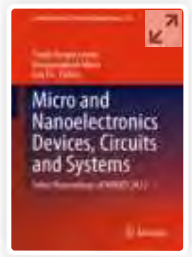
Book Printed Antennas (https://www.taylorfrancis.com/books/mono/10.1201/9781003347057/printed-antennas?refId=469218de-c0d3-4d11-bc24-432b1b615b66&context=ubx)

Edition	1st Edition
First Published	2022
Imprint	CRC Press
Pages	14
eBook ISBN	9781003347057

ABSTRACT

< Previous Chapter (chapters/edit/10.1201/9781003347057-5/microstrip-interconnect-design-modeling-using-reverse-approach-obtain-efficient-wideband-ms-line-rwg-hybrid-transition-varshney-sharma-roshan-kumar?context=ubx)

Next Chapter > (chapters/edit/10.1201/9781003347057-7/antenna-rfid-smart-systems-shailendra-shastri-archana-deshpande?context=ubx)



Micro and Nanoelectronics Devices, Circuits and Systems pp 359–366

Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection

M. Lakshmi Prasanna & V. R. Anitha

Conference paper | First Online: 12 September 2022

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 904)

Abstract

Many studies have shown that most MEMS devices use light beams whose deflection response is the main focus before choosing it for specific applications. This paper deals with the simulation of the MEMS-based cantilever beam for low-pressure detection. COMSOL 5.4 multiphysics FEM model is used to study the behavior of beam and compared

with the conventional cantilever beam. Both are analyzed with two different materials SiO_2 and SU-8. The change in the deflection has been analyzed with various pressure values applied on the cantilever beam. Through simulation and analysis results, it is observed that the deflection of the proposed beam is good at lower pressure values.

Keywords

Microcantilever **Deflection**

Rectangular beam **FEM**

References

1. Guo K, Jiang B, Liu B, Xingeng Li, Wu Y, Tian S, Gao Z, Zong L, Yao S, Zhao M, Mi C, Zhu G (2021) Study on the progress of piezoelectric microcantilever beam micromass sensor. IOP Conf Ser Earth Environ Sci 651:022091. https://doi.org/10.1088/1755-1315/651/2/022091

2. Ali S, Bhuvaneswari H, Kumar B (2020) Design and modeling of a large deflection microcantilever using rectangular SCR. AIP Conf Proc 2281:020037. https://doi.org/10.1063/5.0027942

3. Rotake D, Darji A (2020) Design and reliability testing of microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.13140/RG.2.2.25052.92803

4. Ashok A, Nighot R, Sahu N, Pal P, Pandey A (2019) Design and analysis of microcantilever beams based on arrow shape. Microsyst Technol 25:4379–4390. https://doi.org/10.1007/s00542-019-04555-4

5. Passian A, Thundat T, Thanihaichelvan M (2017) Microcantilever sensors. https://doi.org/10.1016/B978-0-12-803581-8.10525-9

6. Rotake D, Darji A, Singh J (2020) Ultrasensitive multi-arm-microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.1109/VDAT50263.2020.9190249

7. Gharge B, Upadhye V, Bodas D (2015) Design and simulation of microcantilevers for detection of pathogens. https://doi.org/10.1109/ISPTS.2015.7220122

8. Parsediya D (2016) Deflection and stresses of effective micro-cantilever beam designs under low mass loading:111–114. https://doi.org/10.1109/ICEPES.2016.7915915

9. Mathew R, Sankar R (2018) A review on surface stress-based miniaturized piezoresistive SU-8 polymeric cantilever sensors. Nano-Micro Lett 10:1–41. https://doi.org/10.1007/s40820-018-0189-1

10. Gopinath PG (2014) Design and characterization of various shapes of microcantilever for human immunodeficiency virus detection. Int J Res Eng Technol 03:24–29. https://doi.org/10.15623/ijret.2014.0310004

11. Schmid S (2009) Electrostatically actuated all-polymer microbeam resonators—characterization and application

12. Lamba M, Singh K, Chaudhary H (2020) Design analysis of polysilicon piezoresistors PDMS (polydimethylsiloxane) microcantilever based MEMS force sensor. Int J Mod Phys B 34:2050072.https://doi.org/10.1142/S0217979220500721

Acknowledgements

A Special thanks to the National MEMS Design Centre, Sree Vidyanikethan Engineering College, A. Rangampet, Tirupati, for providing the research facility.

Author information

Authors and Affiliations

Department of Electronics and Communication Engineering, JNTUA, Ananthapuramu, Andhra Pradesh, India

M. Lakshmi Prasanna

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Yelahanka, Bangalore, Karnataka, 560064, India

V. R. Anitha

Corresponding author

Correspondence to M. Lakshmi Prasanna .

Editor information

Editors and Affiliations

Electronics and Communication Engineering, National Institute Of Technology Silchar, Silchar, India

Trupti Ranjan Lenka

Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA

Durgamadhab Misra

Electronic Materials Engineering, Australian National University, Canberra, ACT, Australia

Lan Fu

Rights and permissions

Copyright information

About this paper

Cite this paper

Lakshmi Prasanna, M., Anitha, V.R. (2023). Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection. In: Lenka, T.R., Misra, D., Fu, L. (eds) Micro and Nanoelectronics Devices, Circuits and Systems. Lecture Notes in Electrical Engineering, vol 904. Springer, Singapore. https://doi.org/10.1007/978-981-19-2308-1_37

.RIS .ENW .BIB

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-19-2308-1_37	12 September 2022	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-2307-4	978-981-19-2308-1	Engineering Engineering (R0)

International Journal of Electronics Letters ISSN: (Print) (X-band Miniaturized frequency reconfigurable travelling wave antenna X-band Miniaturized frequency reconfigurable trave...

Article *in* International Journal of Electronics Letters · April 2022

CITATIONS
0

READS
80

3 authors, including:

Anitha V R
BMS Institute of Technology
40 PUBLICATIONS **269** CITATIONS

SEE PROFILE

International Journal of Electronics Letters

ISSN: (Print) (Online) Journal homepage: https://www.tandfonline.com/loi/tetl20

X-band Miniaturized frequency reconfigurable travelling wave antenna

Narjala Sri Pravallika, V R Anitha & K. Rama Naidu

To cite this article: Narjala Sri Pravallika, V R Anitha & K. Rama Naidu (2022): X-band Miniaturized frequency reconfigurable travelling wave antenna, International Journal of Electronics Letters, DOI: 10.1080/21681724.2022.2068194

To link to this article: https://doi.org/10.1080/21681724.2022.2068194

INTERNATIONAL JOURNAL OF ELECTRONICS LETTERS
https://doi.org/10.1080/21681724.2022.2068194

Taylor & Francis
Taylor & Francis Group

X-band Miniaturized frequency reconfigurable travelling wave antenna

Narjala Sri Pravallika[a], V R Anitha[b] and K. Rama Naidu[c]

[a]Research Scholar, Department of ECE, JNTUA, Anantapur, India; [b]Department of ECE, BMS Institute of Technology & Management, Bangalore, India; [c]Department of ECE, JNTUA, Anantapur, India

ABSTRACT

The paper presents a novel miniaturized ($1.2\ \lambda_0 \times 0.5\ \lambda_0 \times 0.026\ \lambda_0$) frequency reconfigurable travelling wave antenna (TWA) operating at 8–12 GHz band. The antenna uses PIN diodes arranged in a novel configuration to produce frequency reconfigurability and operates at single frequency bands of 8.5, 9, 9.35, 9.9, 10.45 and 11 GHz with an average bandwidth of 1 GHz based on different switch configurations. The proposed antenna demonstrates an efficiency>80% for the operating frequency in different stations while maintaining a stable average gain of 8.5 dBi. The antenna's compact nature and high gain make it suitable for several applications within modern satellite communication devices and radar systems. The proposed antenna is designed and tested for its conformality using CST Microwave-Studio. The proposed design is further verified with a fabricated prototype measured in an anechoic chamber, agreeing well with the simulated results.

ARTICLE HISTORY
Received 4 October 2021
Accepted 6 March 2022

KEYWORDS
TWA; frequency reconfigurability; leakywave; miniaturization; conformality

1. Introduction

To support the recent technical advancements, miniaturized reconfigurable antennas have been in great demand for remote sensing applications. Frequency reconfigurable antennas allow the user to resonate at different frequency bands, allowing several functions to be operated within the same device. Furthermore, frequency reconfigurable antennas allow the development of compact and flexible systems.

Different frequency reconfigurability techniques have been explored in the literature. One of the most common techniques is changing the antenna's electrical length, resulting in a change in the resonant frequency. Planar antennas such as microstrips (Bansal & Gupta, 2020) are demonstrated to be easily reconfigured using PIN diodes and micro-electro-mechanical system (MEMS) switches. Several antennas have been presented in the literature (Abdulraheem et al., 2017; Fadamiro et al., 2019; Ismail et al., 2021; Sahu & Sharma, 2017; Valkonen et al., 2010) where such RF switches are used in radiating elements and a ground plane to control the surface current and, hence, affect the band of operation. Such RF switches allow fast response time and ease of fabrication, making them very popular for several reconfigurable antennas. Another interesting method of

CONTACT Narjala Sri Pravallika sripravallika37@gmail.com Electronics and Communication Engineering Jawaharlal Nehru Technological University - Anantapur 515002 India

using PIN diodes for frequency reconfigurability was presented in (Nikolaou et al., 2009), where three frequency bands were switched by truncating the antenna using closed stub matching networks. Another standard method for introducing frequency reconfigurability involves varying electrical properties of the resonator using varactor diodes. Varactors introduce an external capacitance that can be varied with the change in voltage, which further introduces frequency reconfigurability (Atallah et al., 2016; Shirazi et al., 2018; Young et al., 2015).

Apart from the reconfigurable characteristics of an antenna, miniaturization is another essential factor. Compact antennas allow smaller form-factor for the overall device (Iftikhar et al., 2020). Several ultra-wideband (UWB) and multi-band antennas with some sort of reconfigurability have been presented in the literature (Arif et al., 2019; Gupta et al., 2020; Luadang et al., 2018; Mohamadzade et al., 2020; Paracha et al., 2018; Simorangkir et al., 2018), but such antennas tend to have limited applications. Modern-day communication systems require compact solutions with multi-band operation and easy switching mechanisms (Ahmad et al., 2018).

This paper presents a compact travelling wave antenna operating at X-band with frequency reconfigurability. Antenna design methodology and the novel switching technique are described in Section 2. Antenna's performance with the help of simulation and further verification with measured results are presented in Section 3. Parametric analysis for the antenna and results for its conformality is described in Section 4, followed by a conclusion in Section 5.

2. Proposed design

This section provides the geometrical description of the proposed antenna and proposed switched array configuration. The antenna is demonstrated to be reconfigurable with the help of PIN diodes used to control the electrical parameters of the antenna.

2.1. Antenna geometry

The proposed antenna geometry is presented in Figure 1. The proposed antenna is designed on an RT-Duroid 5880 substrate (relative permittivity, $\varepsilon_r = 2.2$; loss tangent, $tan\ \delta = 0.0009$) of height, $h = 0.787$ mm. The unit cell is designed to be a mutually coupled S-shaped microstrip line with an open stub fed using a microstrip line. The proposed structure is an array of 8 S-shaped radiating elements connected using PIN-diodes. The S-shaped radiating elements are bent at an angle of 30° to achieve the maximum radiation characteristics for the antenna. The radiating elements are backed with a full conducting ground plane of thickness $t = 0.035$ mm. The geometrical parameters of the proposed antenna are as follows: $w_f = 1.6$ mm, $l_1 = 9$ mm, $l_2 = 4.3$ mm, $d = 0.4$ mm, $g = 1$ mm, $w_u = 0.5$ mm, $l = 35$ mm ($\approx 1.2\lambda_0$) and $w = 15$ mm ($\approx 0.5\lambda_0$).

To further facilitate the design process and understand the unit-cell better, the unit cell was simulated using CST-Microwave Studio. The simulated S-parameters can then be used to calculate the phase constant for the unit cell using the equation (Sarkar et al., 2018),

$$\beta p = \cos^{-1}\left(\frac{1 - S_{11}S_{22} + S_{12}S_{21}}{2S_{21}}\right) \tag{1}$$

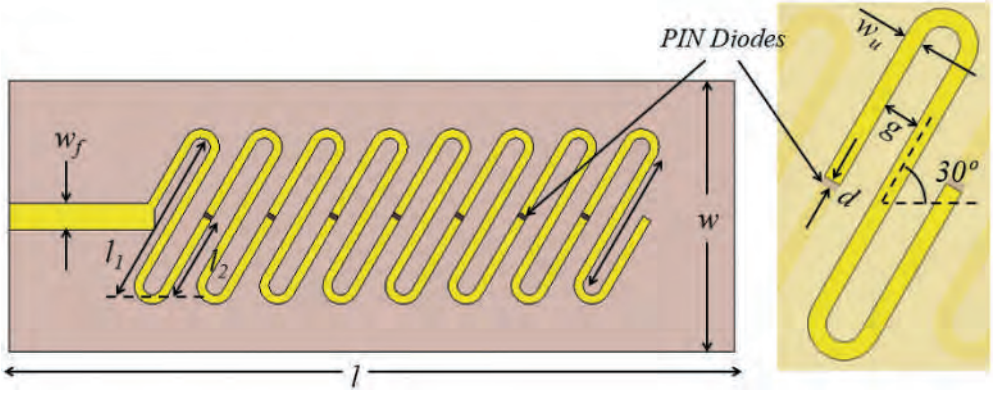


Figure 1. Top view of the schematic for the proposed antenna and zoomed-in view of the radiating element.

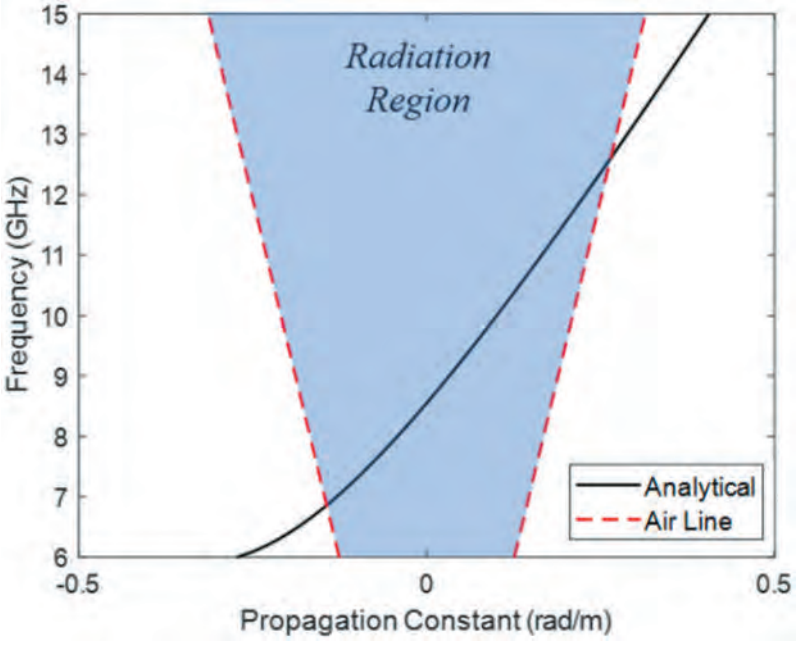


Figure 2. Propagation constant for the proposed antenna.

Here, β is the phase constant for the proposed antenna, p is the periodicity of the unit cell, S_{11}& S_{22} are the simulated reflection coefficients, and S_{12} and S_{21} are the simulated transmission coefficients for the unit cell. The calculated dispersion curve is presented in Figure 2, which demonstrates the frequency band of 7–12 GHz falling under the radiating region.

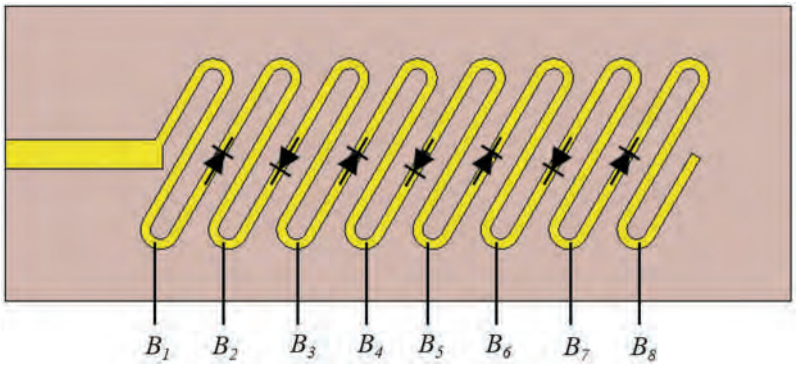


Figure 3. Proposed PIN diodes-based switching configuration and biasing lines.

2.2. Switching technique

PIN diodes behave as RF switches in the radiating elements, and they create an RLC circuit for both ON and OFF states. The proposed design uses Skyworks' SMP1340 PIN diodes. It offers a resistance of 0.81 Ω in the ON state and a very high resistance for the OFF state, making it suitable for switching applications at X-band. The proposed design uses seven PIN diodes, connecting the radiating elements. The switches are arranged alternatively for ease in biasing and make the switching method robust, as shown in Figure 3. The switch diodes being arranged alternatively can be controlled using alternative biasing lines. If the diode is being maintained at the same potential at both arms, it will behave in the OFF state unless the potential difference between the two arms is 3.3 V.

3. Simulated and measured results

3.1. Antenna performance

To evaluate the performance of the proposed leaky-wave antenna, it was designed and simulated using CST Microwave Studio. The design uses a waveguide port to feed the designed model and touchstone files for the PIN diodes to include their effects on the antenna performance within the simulation. The final model was fabricated on an RT-Duroid 5880 substrate (relative permittivity, $\varepsilon_r = 2.2$; loss tangent, $tan\ \delta = 0.0009$) of height, $h = 0.787$ mm, and an SMA connector is used to feed the proposed antenna. The fabricated antenna has been presented in Figure 4 and is measured on an Anritsu MS46522B vector network analyser (VNA).

The resonant frequency for the proposed antenna is reconfigured using different switch configurations. The switches are biased (for ON state) by maintaining a 3.3 V potential difference on the corresponding biasing lines and are kept in OFF state by keeping the potential difference at 0 V (by either grounding or providing the same voltage for both biasing lines of the switch). Table 1 gives a truth table specifying different resonant bands for different switch configurations. Simulated and measured S-parameter results for the proposed antenna shows a perfect agreement and further verify the frequency reconfigurability characteristics of the proposed antenna (see, Figure 5). The antenna offers an average gain of 8.4 dBi with an efficiency greater than 80% for the resonant frequency at different frequency bands, as shown in Figure 6.

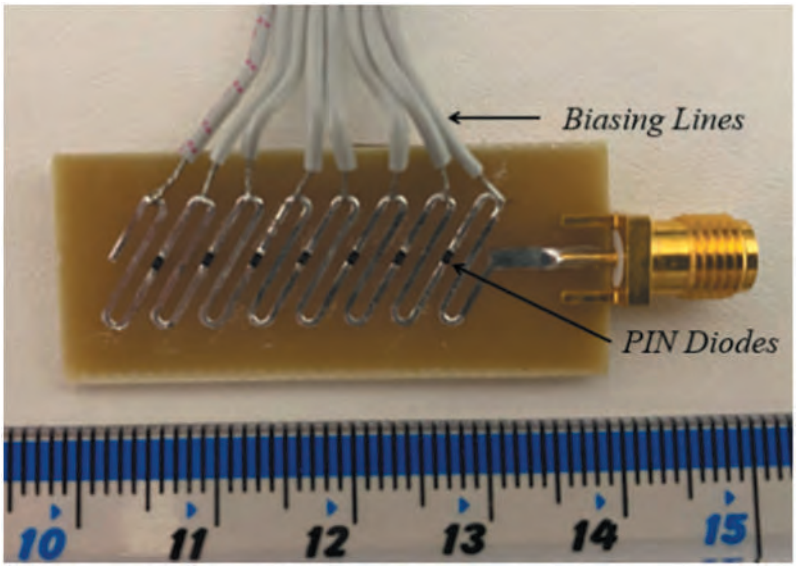


Figure 4. Fabricated TWA design with highlighted PIN diodes and biasing lines.

Table 1. Frequency truth table.

Case	SW1	SW2	SW3	SW4	SW5	SW6	SW7	Freq (GHz)	BW (MHz)
A	0	0	0	0	0	0	0	8.50	1200
B	1	0	0	0	0	0	0	9.05	1000
C	1	0	0	1	0	0	1	9.35	1200
D	1	0	1	0	1	0	1	9.90	1150
E	1	1	0	0	1	1	0	10.45	1200
F	1	1	1	1	1	1	1	11.00	1150

The measured radiation pattern for the proposed antenna is presented in Figure 7. The antenna presents an average gain of 8.52 dBi for all six switch configurations. The slight change in beam-angle is due to a change in phase delay between each radiating element with the change in frequency and can simply be calculated using the general principle of leaky-wave antennas (Balanis, 2005; Elliott, 2003). However, the antenna presents an average half-power beamwidth of 60° in the elevation plane and 90° in the azimuth plane, allowing almost constant radiation in one direction. The side-lobe level (SLL) is maintained below 12 dB for all cases.

3.2. Parametric analysis

To present the robustness of the design and further verify the antenna's performance for its conformality, some of the crucial parameters of the proposed design have been tested in simulation. The S-shaped radiating element is bent at an angle of 30° for a suitable impedance match between the feedline and the radiating element, allowing maximum radiation. This is further verified with the S-parameter plot for the unit cell presented in Figure 8.

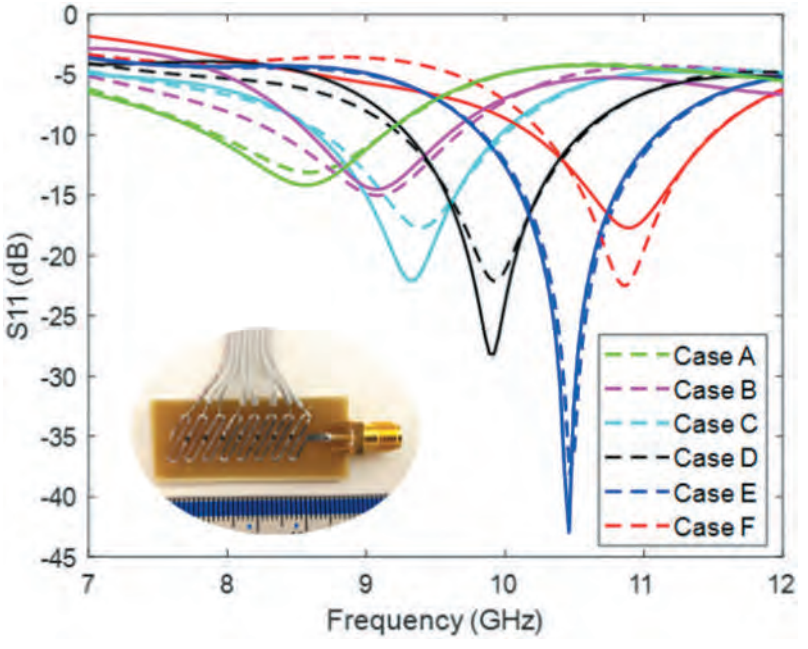


Figure 5. Simulated (dashed) and measured (solid) S11 for the proposed antenna.

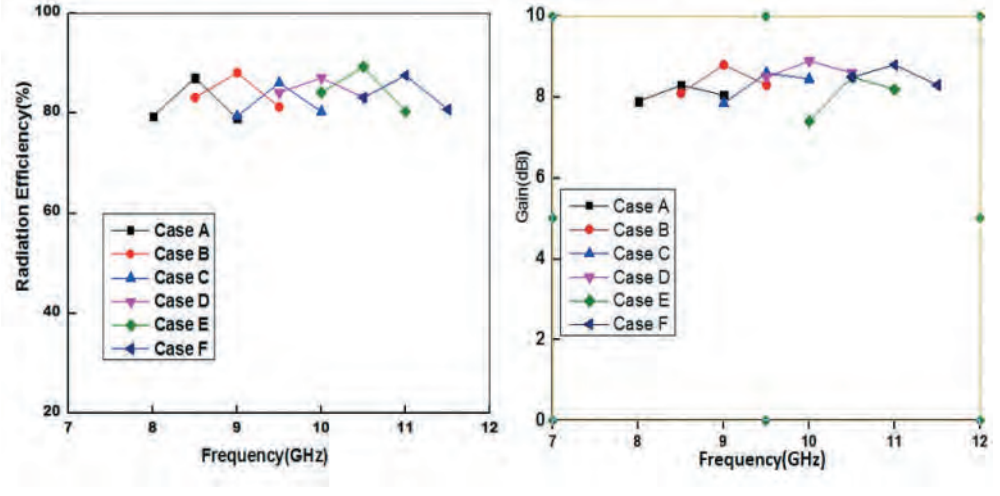


Figure 6. Measured radiation efficiency (dB) measured gain (dBi) for the proposed antenna.

The antenna is a travelling wave structure, and each radiating element receives a reduced power from the last element. The amount of power left after the eight elements have radiated and reflected is negligible, hence the number of radiating elements. This is further verified in simulation by defining a waveguide port on either side of the radiating element and studying the S-parameters. The amount of power reflected can be calculated using S_{11}, and the amount of power sent forward can be calculated using S_{21}. Figure 9

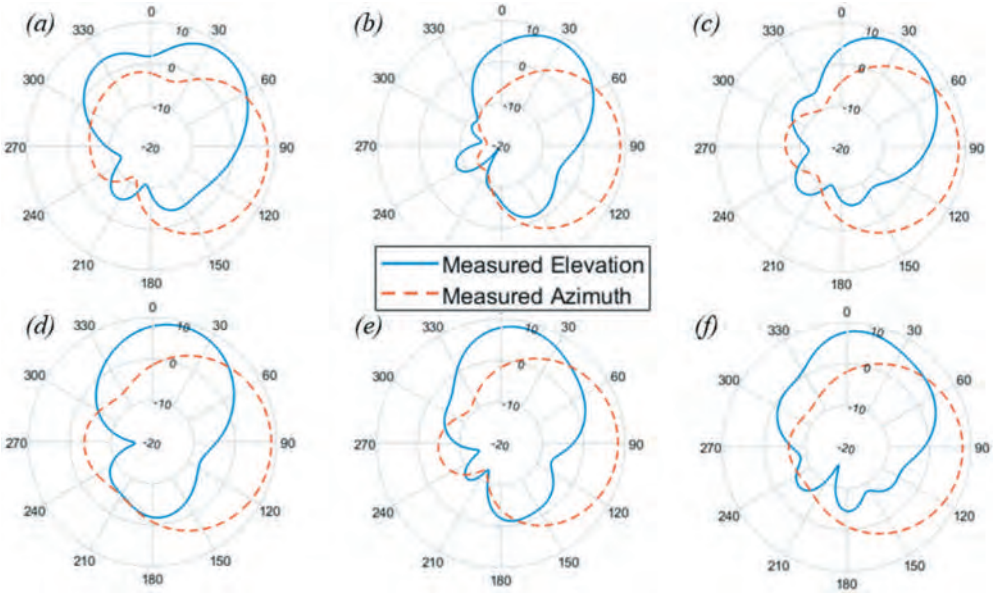


Figure 7. Measured radiation pattern for different cases in Azimuth and Elevation plane at (a) 8.5 GHz, (b) 9 GHz, (c) 9.35 GHz, (d) 9.9 GHz, (e) 10.45 GHz and (f) 11 GHz.

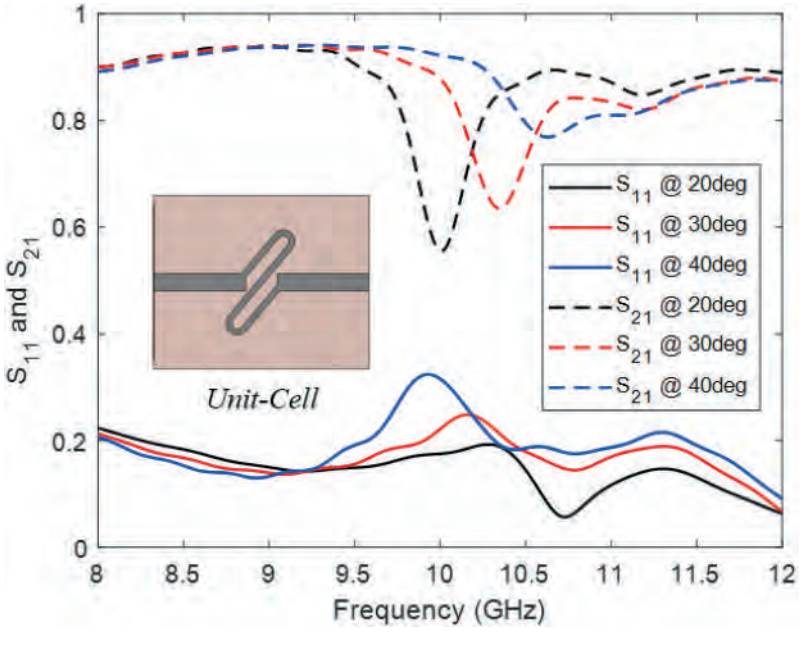


Figure 8. Simulated linear S_{11} and S_{21} for the unit cell at different tilt angles.

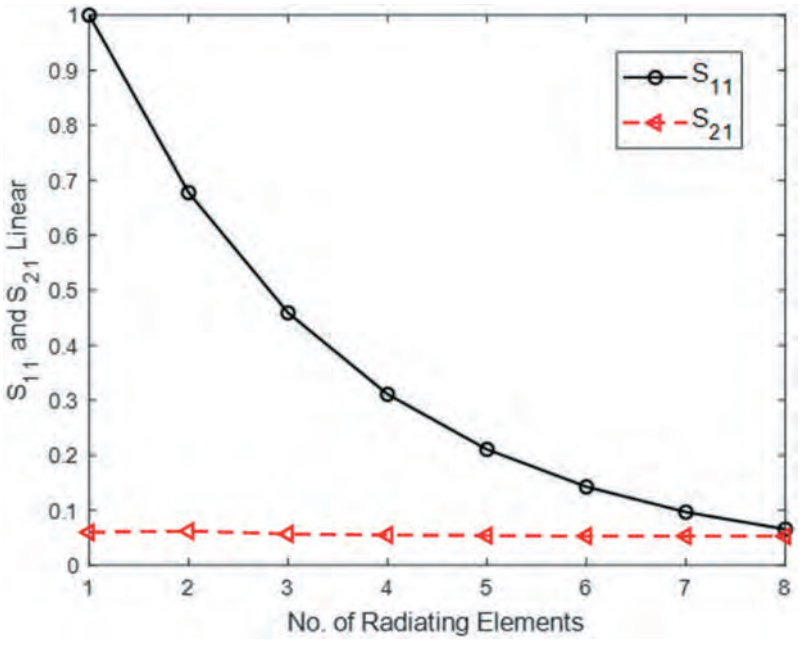


Figure 9. Simulated S11 and S21 demonstrating the amount of power left after 8 elements of the proposed array have radiated out.

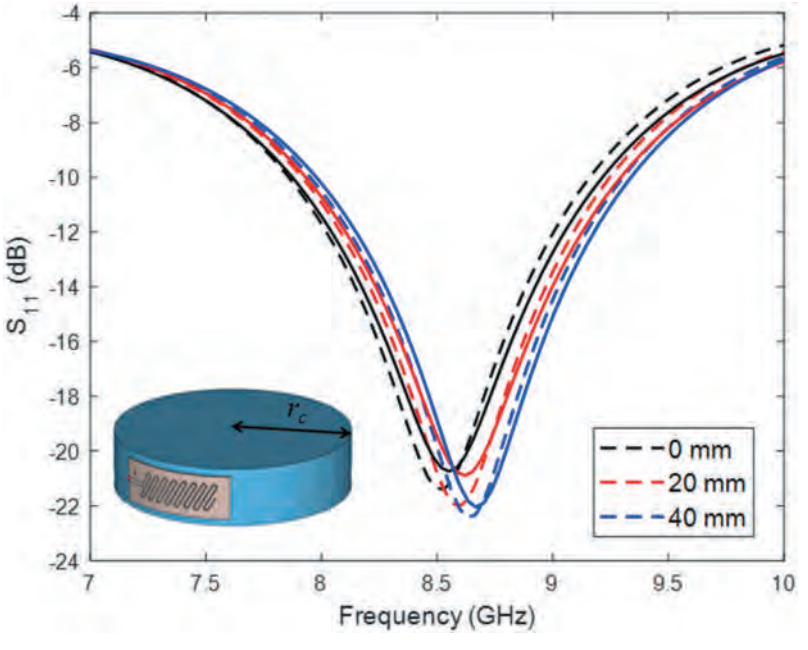


Figure 10. Measured (solid lines) and simulated (dashed lines) S-parameter plot for conformality of the proposed antenna for a radius of 0, 20 and 40 mm.

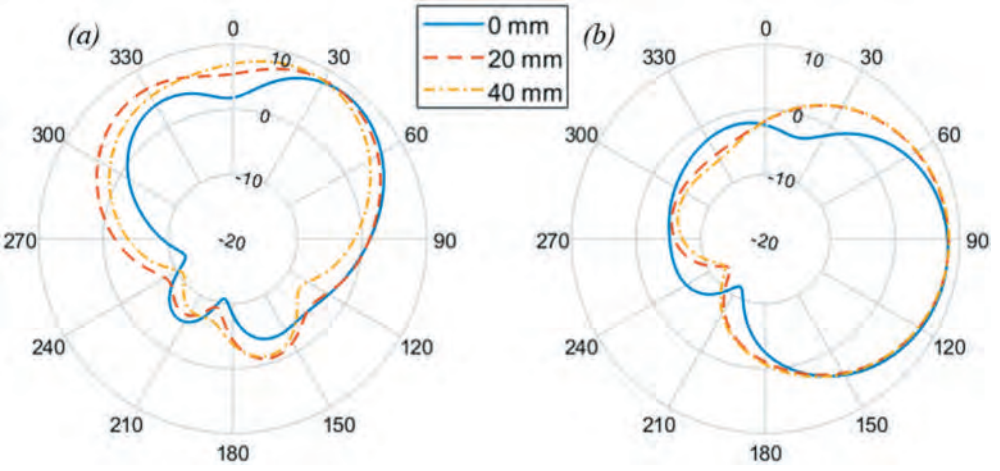


Figure 11. Measured radiation pattern in (a) Elevation and (b) Azimuth plane for the conformed antenna for radius, r_c = 0, 20 and 40 mm.

Table 2. Comparison of the proposed work with the literature.

Ref	Freq. Band	Size	No. of Reconf. Freq. Bands	Avg. Bandwidth	Gain (dBi)
(Abdulraheem et al., 2017)	2.2–6 GHz	50 mm × 45 mm	10	240 MHz	0.8–5.8
(Fadamiro et al., 2019)	1–3 GHz	63 mm × 69.75 mm	4	1 GHz	1.2–4
(Shirazi et al., 2018)	1.8–5.7 GHz	80 mm × 80 mm	2	300 MHz	0.1–4.2
(Young et al., 2015)	0.5–3.5 GHz	47.7 mm × 45 mm	5	500 MHz	2
(Atallah et al., 2016)	1.7–5.7 GHz	40 mm × 40 mm	3	500 MHz	0.3–0.8
(Ahmad et al., 2018)	4.2–7 GHz	24 mm × 20 mm	5	400 MHz	2.5–3.1
This Work	8–12 GHz	35 mm × 15 mm	7	1 GHz	8.1–8.6

presents simulated S_{11} and S_{21} for the antenna when tested for one to ten elements. As seen from the plotted graph, the amount of power left after the eighth element is radiated out is negligible, and hence, verifying the initial theory.

Conformal antennas have been in massive demand for weather radar and similar applications. The proposed antenna is printed on a thin substrate, allowing for minimal conformality. However, the antenna being tiny in the Y-axis (w = 15 mm), it is impossible to bend the antenna in this direction. The antenna was measured for conformality in X-axis for a bend radius of 0 mm, 20 mm, and 40 mm. The variation in S-parameters for the conformed shape is presented in Figure 10.

The measured radiation pattern for the conformality is shown in Figure 11. The antenna was measured in an NSI2000 anechoic chamber and was conformed on a metallic pipe of radius, r_c = 20, and 40 mm. The antenna's gain is found to have a minimal change of ±0.12 dBi which can be considered negligible. However, the direction of radiation had a significant change. The beam direction changed from 31° in elevation

plane for no conformality to 27° for r_c = 20 mm and 22° for r_c = 40 mm. This can be explained by the change in the radiator's direction with conformality. The antenna maintains a similar efficiency with and without conformality, and hence, the antenna can be considered suitable for several weather radar applications and for mounting on curved surfaces.

The performance and miniaturised nature of the proposed antenna are compared with the literature in Table 2 to further appreciate the presented work's importance.

4. Conclusion

The article presented a miniaturized travelling-wave antenna with frequency reconfigurability, operating at X-band (8–12 GHz). The proposed antenna used PIN diodes in a novel configuration to control different radiating elements of the array. The antenna is presented to be operating at 8.5, 9.05, 9.35, 9.9, 10.45 and 11 GHz for different switch configurations. The frequency reconfigurability is attained by connecting different radiating elements of the proposed array using PIN diodes. Each operating mode is shown to have a minimum bandwidth of 1 GHz. The antenna was measured in an NSI2000 anechoic chamber for its radiation characteristics, and it presented an average gain of 8.5 dBi throughout the frequency band for different configurations with an efficiency >80%. The antenna is further tested for its conformality, where it was bent across its length on a metallic pipe of radius 20 and 40 mm, respectively. The antenna presented a stable gain with a slight change in the beam direction. Each antenna configuration demonstrates a slight change in the beam declination, which is believed to be because of the antenna's travelling-wave nature. Each element being fed sequentially introduces a phase delay, resulting in a beam declination. This beam-declination changes with the phase delay, which is further controlled with the operating frequency. The fabricated prototype presents a good agreement with the simulated results. Furthermore, the directional behaviour of the antenna, along with frequency reconfigurability, conformality and high gain, maintains a small form-factor of 1.2 $\lambda_0 \times 0.5\ \lambda_0 \times 0.026\ \lambda_0$, which makes it a suitable candidate for several communications and radar applications.

Disclosure statement

No potential conflict of interest was reported by the authors.

References

Abdulraheem, Y. I., Oguntala, G. A., Abdullah, A. S., Mohammed, H. J., Ali, R. A., Abd-Alhameed, R. A., & Noras, J. M. (2017). Design of frequency reconfigurable multiband compact antenna using two PIN diodes for WLAN/WiMAX applications. *IET Microwaves, Antennas & Propagation*, *11*(8), 1098–1105. https://doi.org/10.1049/iet-map.2016.0814

Ahmad, A., Arshad, F., Naqvi, S. I., Amin, Y., Tenhunen, H., & Loo, J. (2018). Flexible and compact spiral-shaped frequency reconfigurable antenna for wireless applications. *IETE Journal of Research*, *66*(1), 22–29. https://doi.org/10.1080/03772063.2018.1477629

Arif, A., Zubair, M., Ali, M., Khan, M. U., & Mehmood, M. Q. (2019). A compact, low-profile fractal antenna for wearable on-body WBAN applications. *IEEE Antennas Wireless Propagation Letter*, *18* (5), 981–985. https://doi.org/10.1109/LAWP.2019.2906829

Atallah, H. A., Abdel-Rahman, A. B., Yoshitomi, K., & Pokharel, R. K. (2016). Compact frequency reconfigurable filtennas using varactor loaded T-shaped and H-shaped resonators for cognitive radio applications. *IET Microwaves, Antennas & Propagation*, *10*(9), 991–1001. https://doi.org/10.1049/iet-map.2015.0700

Balanis, C. A. (2005). *"antenna theory: analysis and design"* (3rd ed.). Wiley.

Bansal, A., & Gupta, R. A. (2020). Review on microstrip patch antenna and feeding techniques. *International Journal of Information Technology*, *12*(1), 149–154. https://doi.org/10.1007/s41870-018-0121-4

Elliott, R. S. (2003). *Antenna Theory And Design* (Revised ed.). IEEE Press Series on Electromagnetic Wave Theory.

Fadamiro, A. O., Famoriji, O. J., Zakariyya, R. S., Zhang, Z., & Lin, F. (2019). Design of H-Tree fractal slots frequency reconfigurable hexagonal patch antenna using PIN diodes. *Journal of Electromagnetic Waves and Applications*, *33*(12), 1591–1604. https://doi.org/10.1080/09205071.2019.1618740

Gupta, R., Bakshi, G., & Bansal, A. (2020). Dual-band circularly polarized stacked sapphire and TMM13i rectangular DRA. *Progress In Electromagnetics Research*, *91*, 143–153. https://doi.org/10.2528/PIERM20012701

Iftikhar, A., Parrow, J., Asif, S., Fida, A., Allen, J., Allen, M., Braaten, B., & Anagnostou, D. (2020). Characterization of novel structures consisting of micron-sized conductive particles that respond to static magnetic field lines for 4G/5G (Sub-6 GHz) reconfigurable antennas. *Electronics*, *9*(6), 903–910. https://doi.org/10.3390/electronics9060903

Ismail, M. F., R, M. K. A., & Majid, H. A. (2021). Wideband Frequency Reconfiguration using PIN diode. *Microwave Optical Technology Letter*, *54*(6), 1407–1412. https://doi.org/10.1002/mop.26826

Luadang, B., Sakonkanapong, A., Dentri, S., Pansomboon, R., & Phongcharoenpanich, C. (2018). NFC-enabled far-field antenna on PET flexible substrate for 3G/4G/LTE mobile devices. *IEEE Access*, *7*, 71966–171973.

Mohamadzade, B., Simorangkir, R. B. V. B., Hashmi, R. M., Chao-Oger, Y., Zhadobov, M., & Sauleau, R. (2020). A conformal band-notched ultrawideband antenna with monopole-like radiation characteristics. *IEEE Antennas Wireless Propagation Letter*, *19*(1), 203–207. https://doi.org/10.1109/LAWP.2019.2958036

Nikolaou, S., Kim, B. K. B., & Vryonides, P.(2009). "Reconfiguring antenna characteristics using PIN diodes," *Proceedings of EuCAP*, pp 3748–3752.

Paracha, K. N., Rahim, S. K. A., Chattha, H. T., Aljaafreh, S. S., Rehman, S. U., & Lo, Y. C. (2018). Low-cost printed flexible antenna by using an office printer for conformal applications. *International Journal of Antennas and Propagation*, *2018*, 1–7. https://doi.org/10.1155/2018/3241581

Sahu, N. K., & Sharma, A. K. (2017). "An investigation of pattern and frequency reconfigurable microstrip slot antenna using PIN diodes", *Progress in electromagnetics research symposium-spring (PIERS)*, pp. 971–976.

Sarkar, A., Adhikary, M., Sharma, A., Biswas, A., Akhtar, M. J., & Hu, Z. (2018). "Composite right/left-handed based compact and high gain leaky-wave antenna using complementary spiral resonator on HMSIW for Ku band applications. *IET Microw Antennas Propag*, *12*(8), 1310–1315. https://doi.org/10.1049/iet-map.2017.0478

Shirazi, M., Huang, J., Li, T., & Gong, X. (2018). A switchable-frequency slot-ring antenna element for designing a reconfigurable array. *IEEE Antennas and Wireless Propagation Letters*, *17*(2), 229–233. https://doi.org/10.1109/LAWP.2017.2781463

Simorangkir, R. B., Kiourti, A., & Esselle, K. P. (2018). UWB wearable antenna with a full ground plane based on PDMS-embedded conductive fabric. *IEEE Antennas and Wireless Propagation Letters*, *17*(3), 493–496. https://doi.org/10.1109/LAWP.2018.2797251

Valkonen, R., Luxey, C., Holopainen, J., Icheln, C., & Vainikainen, P. (2010). "Frequency-reconfigurable mobile terminal antenna with MEMS switches", *Proceedings of the fourth european conference on antennas and propagation* (pp. 1–5).

Young, M. W., Yong, S., & Bernhard, J. T. (2015). A miniaturized frequency reconfigurable antenna with single bias, dual varactor tuning. *IEEE Transactions on Antennas and Propagation*, *63*(3), 946–951. https://doi.org/10.1109/TAP.2015.2388776

Wireless Personal Communications (2022) 123:2209–2225
https://doi.org/10.1007/s11277-021-09235-z

Parametric Analysis of Compact UWB-MIMO Antenna with Improved Isolation Using Parasitic Reflectors and Protruded Ground Strips

Tathababu Addepalli[1] · V. R. Anitha[2]

Accepted: 14 October 2021 / Published online: 27 October 2021

Abstract
A Compact Multiple-Input Multiple-Output (MIMO) antenna of size 16×26 mm^2 ($0.144\ \lambda_0 \times 0.234\ \lambda_0$) with improved isolation is presented in this paper. It covers the bandwidth of 2.7–14.9 GHz (FBW is 138.2%), which includes the entire UWB, X band, and K_u band applications. The elements in the structure are placed at a distance of $0.018\ \lambda_0$ (2 mm). Optimized values of two parallel strips are placed between the elements for isolation, giving $|S_{21}| > 15$ dB in most of the band. For further enhancement, four small rectangular strips with different dimensions are protruded to the ground plane, resulting in improved isolation $|S_{21}| > 20$ dB in most of the band. The radiation efficiency and peak gain values are 86–96% and 0.8–6.6 dBi respectively. Diversity performance is evaluated with the Envelope Correlation Coefficient (ECC < 0.06), Diversity Gain (DG > 9.975 dB) and Total Active Reflection Coefficient (TARC < − 10 dB). The Parametric study has been done for optimization values of the proposed model. The proposed model is simulated & fabricated on the FR4 substrate, tested and verified practically. The measured results almost follow the simulated results.

Keywords Elliptical patch · MIMO · Isolation · ECC · DG and TARC

1 Introduction

Multiple Input Multiple Output (MIMO) technology plays a key role in present wireless communication owing to the high data rate and reliability. High data rates are achieved through spatial multiplexing and reliability with space–time coding. But the major limitation of MIMO technology is coupling between antenna elements when they are arranged in a closed manner. If the separation is high, less coupling occurs but the size of the antenna housing will increase. The number of antennas connected tightly together for compactness cause more coupling between them. Coupling is introduced mainly due to currents flowing

✉ Tathababu Addepalli
babu.478@gmail.com

[1] Department of ECE, JNTUA, Anantapur, Andra Pradesh, India

[2] Department of ECE, BMS Institute of Technology & Management, Bangalore, Karnataka, India

on the metal surface, surface waves in common substrate and collision of radiation in free space. Coupling due to free space radiation can be overcome by pattern diversity technique. Effect of surface waves and currents on element to element can be reduced by placing parasitic patches/stubs, EBG structures, neutralization lines and DGS [1–4]. Ultra-wideband technology is being developed rapidly due to low power consumption and less interference. But it is facing a problem with multipath fading. This problem is overcome by MIMO technology using diversity techniques. The combination of UWB with MIMO technology is leading efficient role in wireless systems. Isolation is improved between rectangular shaped MIMO antennas of size 70×120 mm^2 and separated with a distance of 4 mm using printed stubs [5]. Circular shaped MIMO antenna of size 47×93 mm^2.

($0.49\ \lambda_0\times0.95\ \lambda_0$) with improved isolation using T- shaped stub is mentioned in [6]. Two elements hexagonal-shaped MIMO antenna of dimension 50×60 mm^2 with isolation enhancement using double-layer EBG structures is discussed in [7]. Orthogonal arrangement of four-element fractal antenna for UWB applications with 45×45 mm^2 substrate dimensions is discussed in [8]. 35×50 mm^2 ($0.333\lambda_0\times0.233\ \lambda_0$) size UWB MIMO antenna with correlation coefficient (ECC < 0.0036) is studied in [9]. Reconfigurable MIMO antenna for ultra-wideband applications with the port-to-port isolation above 20 dB is presented in [10]. Offset fed microstrip patch antenna operating from 3.1 to 12 GHz for MIMO systems is discussed in [11].

Circular shaped MIMO antenna with improved isolation using a resistive loading technique is discussed in [12]. Compact MIMO antennas structure of size of 35×35 mm^2 with improved isolation using diversity technique for UWB applications [13]. Quasi circular-shaped MIMO antenna of size 24×42mm^2 is arranged in an orthogonal manner for isolation improvement [14]. Ultra-wideband MIMO antenna, achieve high isolation using L shaped stubs as discussed in [15]. UWB-MIMO antenna of size 30×30 mm^2 with isolation enhancement using Y shaped DGS structure is discussed in [16]. Rectangular shaped UWB antenna with isolation is $S_{21} > 15$ dB for MIMO systems is discussed in [17]. Decoupling structures are used for the reduction of coupling between tightly coupled antennas [18]. Isolation is improved between antenna elements using the neutralization line technique, the size of PCB is 80×135 mm^2 as discussed in [19]. Owl shaped MIMO antenna with isolation improvement using protruded strip and the dimensions of the antenna are 26×31 mm^2 as discussed in [20]. The overall size of MIMO antennas is 22×36 mm^2 and an isolation enhancement of $S_{21} > 15$ dB for UWB-MIMO applications [21]. A compact half hexagonal-shaped UWB-MIMO antenna of size 20×34 mm^2 with high isolation ($S_{21} > 20$ dB) in most of the band using a double circular ring resonator is discussed in [22]. Eye shaped MIMO antenna with low mutual coupling and ECC values less than 0.02 for ultra-wideband applications is discussed in [23]. 18×34 mm^2 sized MIMO antenna with high isolation ($S_{21} > 22$ dB) using stubs and operating from 2 to 20 GHz is presented in [24]. Two nine shaped antenna fed with an orthogonal arrangement of microstrip lines, whose isolation is greater than 22.5 dB is discussed in [25]. Coupling is reduced between antennas to -22 dB when connected side by side with spacing of 2 mm using neutralization line and the size of the substrate is 16×35 mm^2 as mentioned in [26]. The compact elliptical-shaped MIMO antenna of dimensions 18×23 mm^2 with isolation 15 dB from 3 to 4 GHz and 20 dB from 4 to 12.4 GHz using modified T shaped stub is presented in [27].

In this communication, a compact MIMO antenna is designed, fabricated on low cost & widely available FR-4 substrate of size 16×26 mm^2 ($0.144\ \lambda_0\times0.234\ \lambda_0$), whose dielectric constant is 4.4. The antennas are separated with a distance of 2 mm ($0.018\lambda_0$), which is very small and automatically leads to high coupling ($|S_{21}| < 15$ dB) between them. To enhance the isolation, optimized values of two parallel strips are placed between the

antennas, however not sufficient for MIMO antennas. Four rectangular strips of varying dimensions are protruded to the ground plane, which further improves the isolation of $|S_{21}| > 20$ dB in the entire band. The proposed model simulated using ANSYS Electronics desktop 16.2 and checked the S-parameter values with an Agilent N5230A vector network analyzer. It was tested and verified practically in an anechoic chamber.

2 Antenna Theory

2.1 Design Process

The schematic of the proposed model with optimized values is shown in Fig. 1. It consists of two elliptical-shaped antennas, two parallel strips and DGS with protruded strips. The elements in the model are fed by a microstrip line, whose impedance is 50 Ω. The evolution of proposed model is in seven stages and it is depicted in Fig. 2. The basic single radiating antenna is designed from the basic elliptical shaped radiator equations [28] and the equations are represented in (1). To get the better impedance matching, truncation ground has been taken to the basic antenna. The parametric study has been done on various parameters like ground length, truncated ground and major axis and axial ratios of elliptical shaped patch antenna for getting optimized values. After optimizing the parameters, it provides a wide impedance bandwidth which is from 3.9 to 12.5 GHz. Later, the same radiator placed at a distance of 2 mm, 4 mm, 6 mm and 8 mm is shown in Fig. 3.

If the two radiators are closer, then more coupling is introduced. The second radiator placed at a distance of 2 mm ($0.018\lambda_0$, λ_0 measured at the lower frequency of 2.7 GHz) the obtained isolation is below 15 dB in most of the band and improvement in lower frequency is observed. To improve the isolation between elements two parallel strips with optimized dimensions are placed between them. Due to strips, isolation is improved to 15 dB in most of the band from 10 dB, except in the band 2.7–4.1 GHz and lower frequency is also improved to 2.8 GHz from 3.9 GHz. Coupling is reduced at the lower band by adding rectangular strips to the ground plane. After adding three strips to the ground plane one after the other, the isolation is improved at the lower band by 3 dB. Finally, the fourth strip of

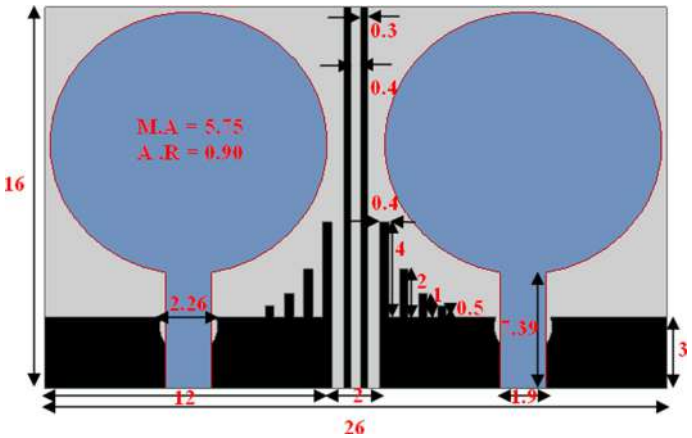


Fig. 1 Proposed model with geometry

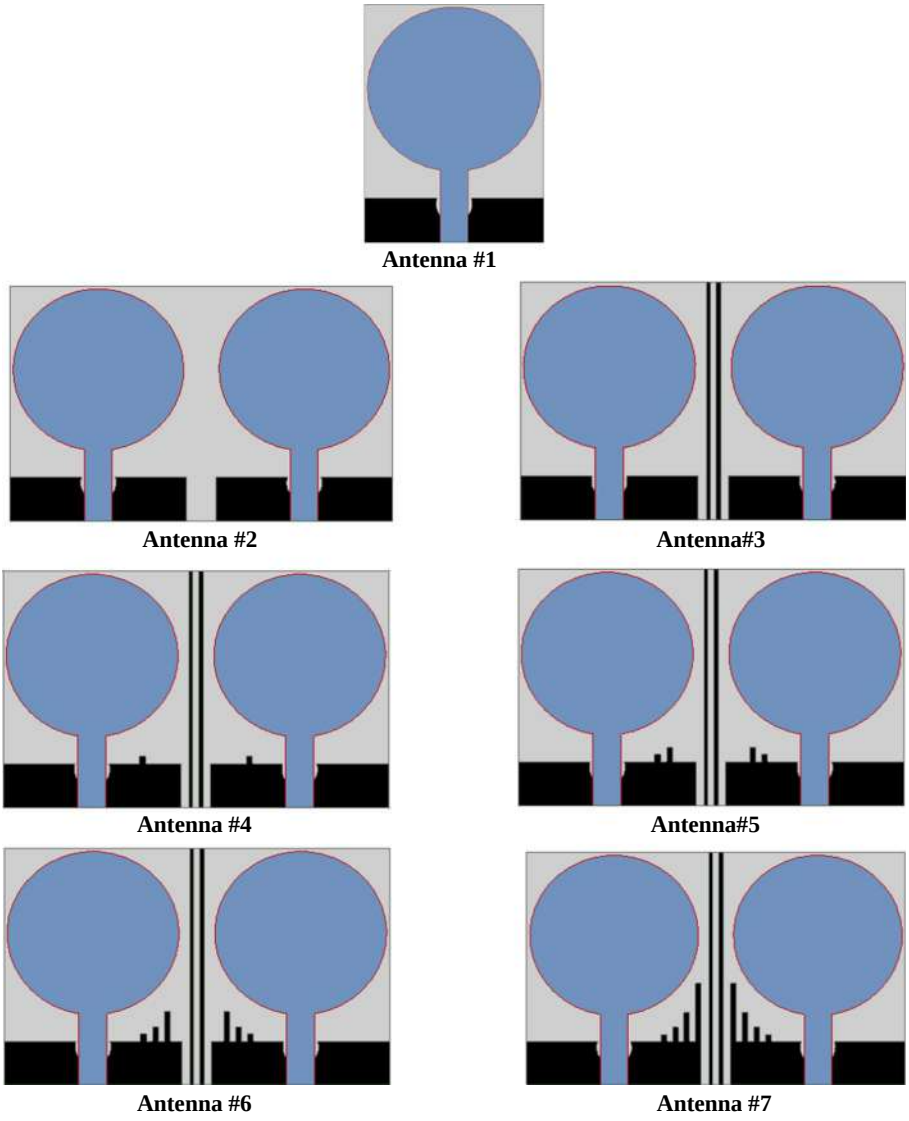


Fig. 2 Evolution stages of proposed model

size 0.4×4 mm^2 is added to the edge of the ground plane, and the isolation is improved to 20 dB in most of the band. The resultant simulated S-parameters are shown in Fig. 4.

$$f_l = \frac{7.2}{\{(L + r + p) \times k\}} \text{ GHz} \tag{1}$$

where, 'L' is the height of the planar monopole antenna in cm, 'r' is the effective radius of the equivalent monopole antenna in cm. L = 2B, r = A/4; B = semiminor axis, A = semimajor axis, 'P' = feed gap with respect to the partial ground in cm, 'K' = 1.15; (Proportionality

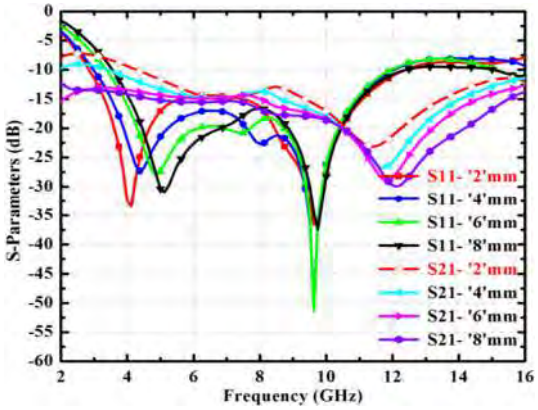


Fig. 3 Simulated S-parameters of antenna #2 with various distance separations

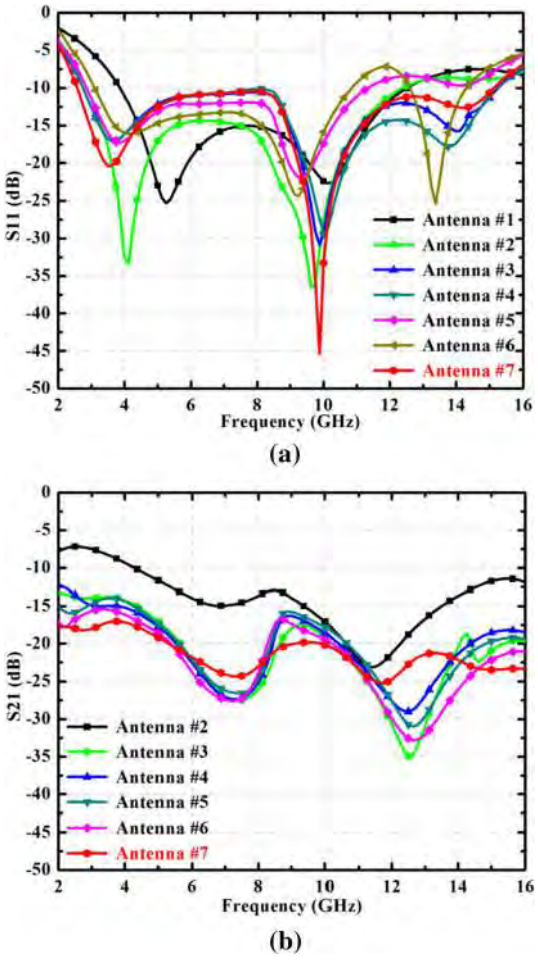


Fig. 4 **a** Simulated S_{11} values; **b** Simulated S_{21} values

constant for the 4.4 dielectric constant of FR4 substrate) estimates lower band-edge frequency within 10%.

From the basic design (antenna I):
A = 0.575 cm, B = 0.517 cm, L = 1.034 cm,
r = 0.143 cm, P = 0.52 cm, k = 1.15.
`According to the equation:
$f_l = 7.2/((1.034 + 0.143 + 0.36) \times 1.15)$
= 4.07 GHz.

2.2 Parametric Analysis

Parametric analysis of major & minor axis of the radiator, the width of the feed element, partial ground plane and the truncated ground plane is performed to determine the impedance performance of antenna. The major axis (M.A = 5.75 mm) & axial ratio (A.R = 0.90) and feed width 1.9 mm of the radiator are observed to give good impedance performance. The resultant graphs are shown in Fig. 5a and b and the overall impedance bandwidth is from 2.7 to 14.9 GHz. Partial ground plane and the truncated ground plane will always give wider bandwidth. The parametric study is carried with optimized values of 3 mm ground length and the axial ratio of 1 & a major axis of 1.2 mm. Figure 5c and d are represents the respective graphs. To enhance the isolation between radiators, two parallel strips are placed between them. Due to these strips of optimized values of 16 mm length give isolation above 15 dB in most of the band except at small band. Figure 5e Shows the parametric study of the length of the two strips. The effect of truncated ground on impedance matching is shown in Fig. 5f, enhancing isolation to some extent.

2.3 Surface Current Distribution

To understand the effect of one radiator on the other, the plots of surface current distributions are required. Figure 6a–h shows the surface current distributions with 3-D polar plots at 3.5 GHz, 5.5 GHz, 9.8 GHz, and 14.2 GHz. Here, parallel strips and protruded strips to the ground act as reflectors and storage elements. Due to reflections from those elements, a little amount of directivity of the proposed model increases thus increasing the gain.

3 Simulated and Measured Results

The comparison of simulated and measured S- Parameter values is shown in Fig. 7. The impedance bandwidths of simulated and measured results are obtained in the range 2.7–14.9 GHz and 2.6–14.7 GHz respectively. The simulated $|S_{21}|$ values are above 20 dB in most of the band except at the small band. The measured $|S_{21}|$ values are above 25 dB in most of the band. The small discrepancy is observed between simulated and measured values due to cable losses and tolerances in fabrication & soldering. Figure 7b and c are the front and rear view photographs of the fabricated model.

The radiation efficiency and peak gain values of the proposed model are shown in Fig. 8 and is above 95% in most of the band. The range of peak gain values of the proposed model is 0.8–6.6 dBi and the maximum value is 6.6 dB at 11.2 GHz. The radiation performance is studied at 3.5 GHz, 5.5 GHz, 9.8 GHz, and 14.2 GHz. The measured and simulated E and H filed patterns are shown in Fig. 9. The shapes of E and H fields are dumbbell and

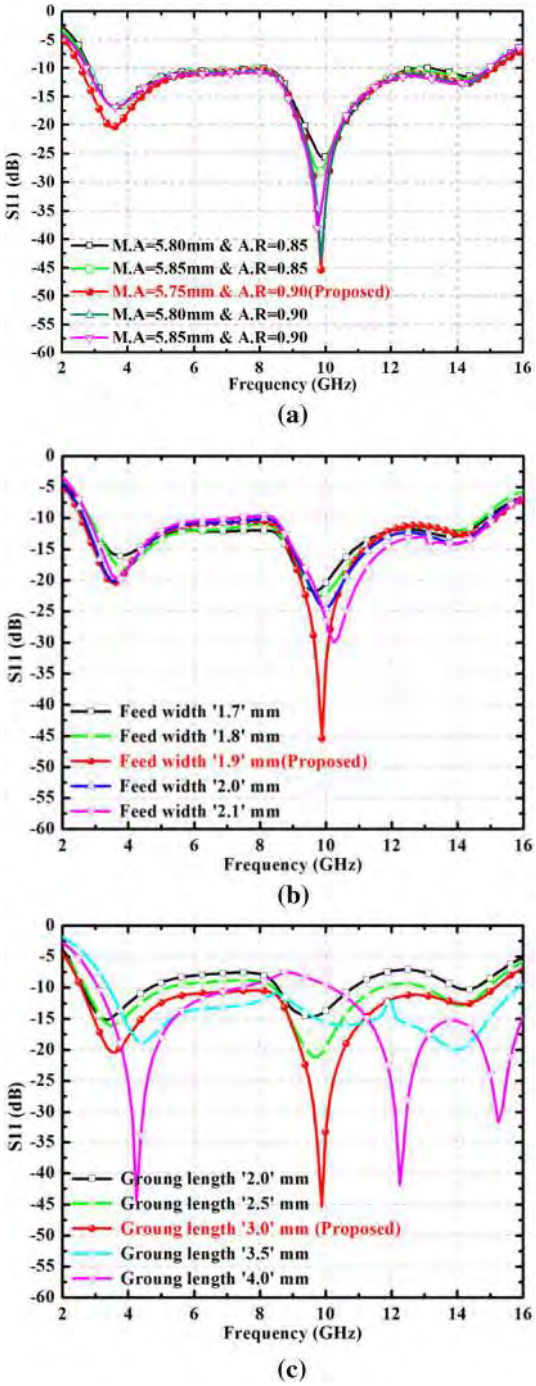
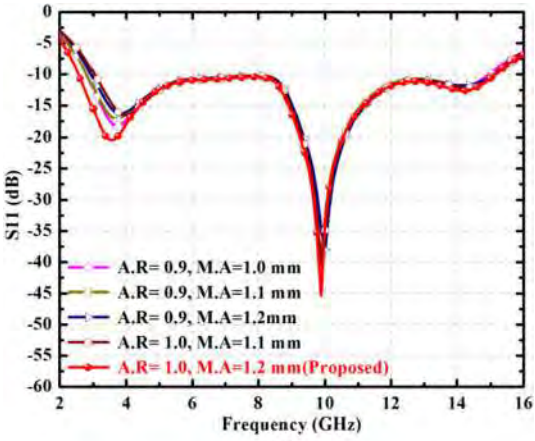
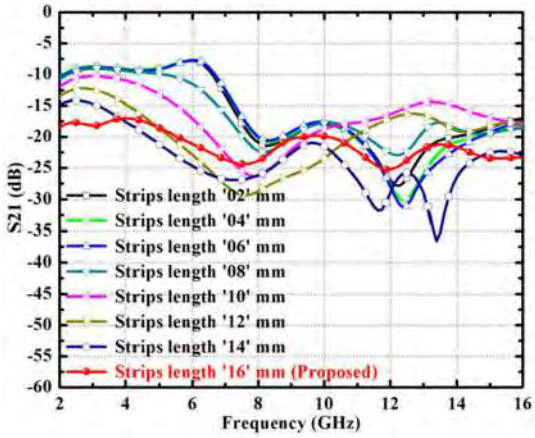


Fig. 5 Parametric analysis of **a** Major axis and Axial ratio of Radiator; **b** Feed Width of Radiator; **c** Ground Length; **d** Truncated Ground Plane; **e** Length of Parallel Strip Lines and **f** Simulated S—parameter values of partial ground with & without elliptical cut

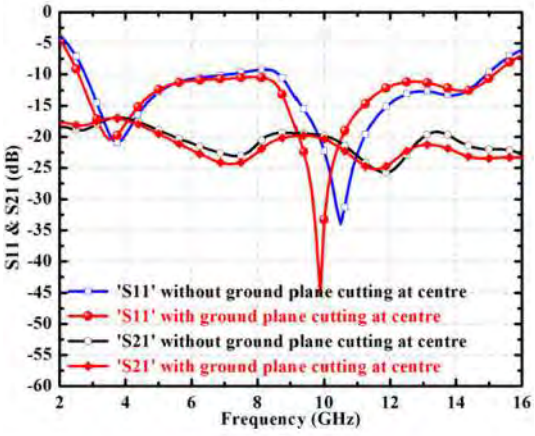
Fig. 5 (continued)



(d)



(e)



(f)

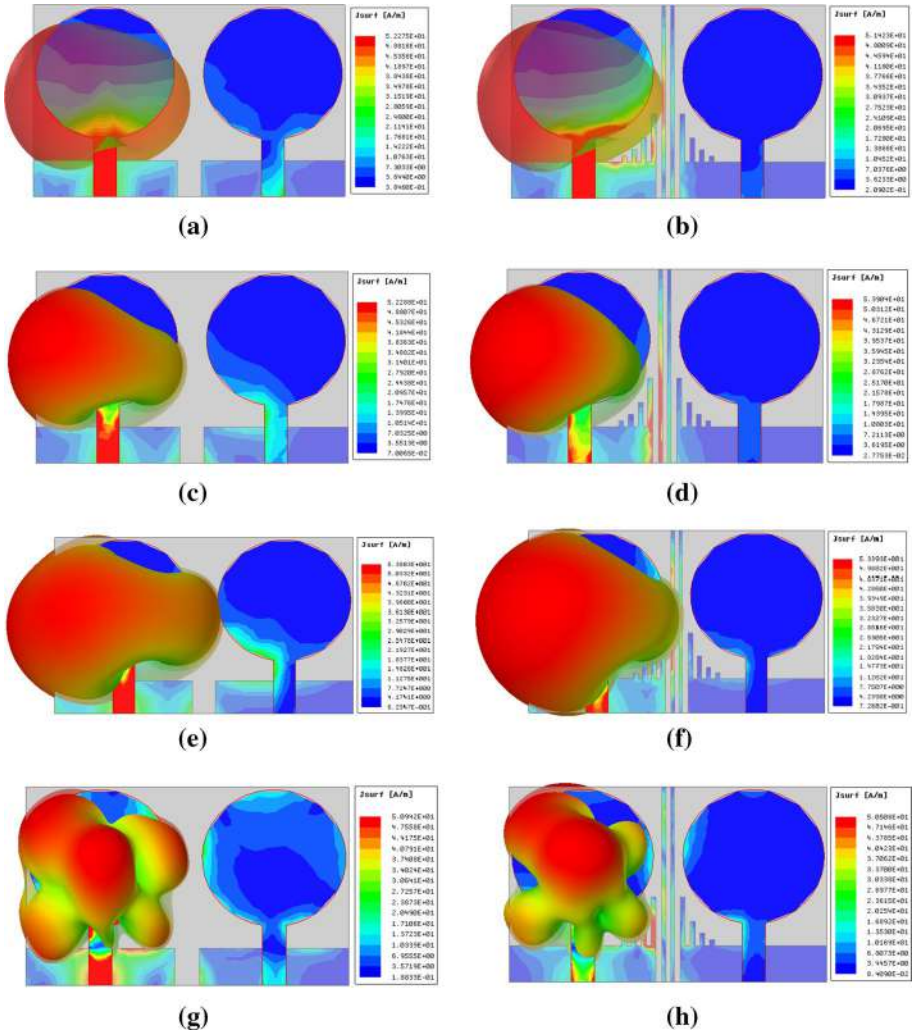


Fig. 6 Surface Current Distributions: (**a**) and (**b**) at 3.5 GHz; (**c**) and (**d**) at 5.5 GHz; (**e**) and (**f**) at 9.8 GHz and (**g**) and (**h**) at 14.2 GHz

Omnidirectional patterns at low frequencies. But, due to higher modes at higher frequencies, the patterns are disturbed. At these frequencies, the energy will be distributed to all the modes, hence patterns are not in shape of dumbbell and Omnidirectional.

4 Diversity Performance

The diversity performance of MIMO antennas is studied through three metrics namely the envelope correlation coefficient (ECC), diversity gain (DG) and total active reflection coefficient (TARC). The isolation between the two radiators can be studied easily

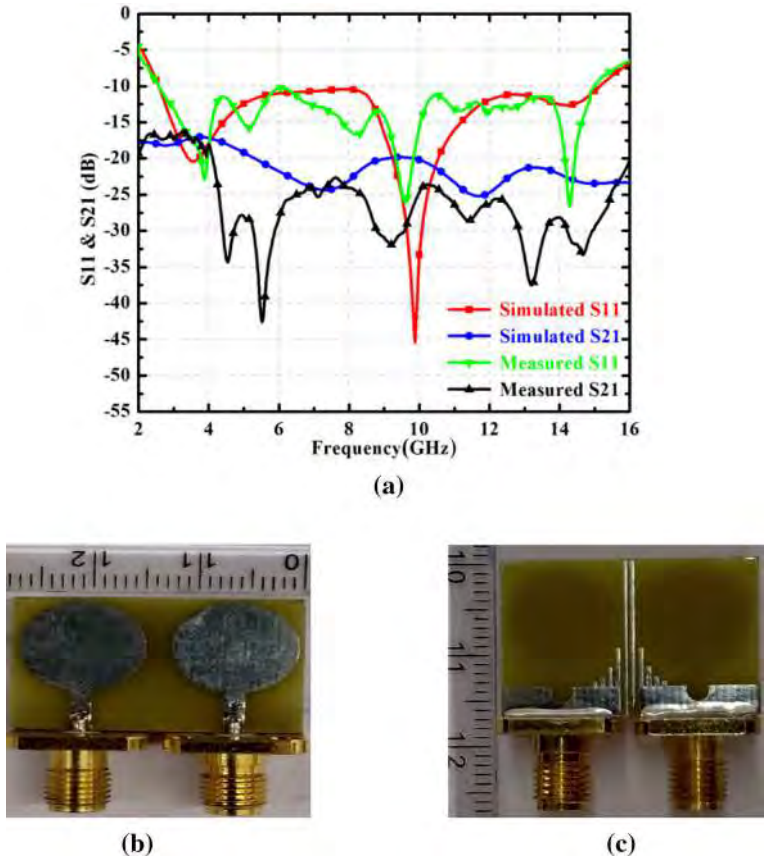
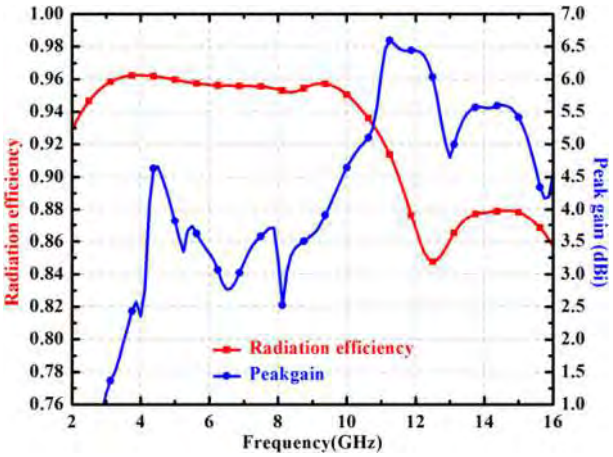


Fig. 7 (**a**) Simulated and Measured S- parameters of proposed model; (**b**) & (**c**) are photographs of front and rear views of proposed model

Fig. 8 Simulated radiation efficiency and peak gain values of proposed model



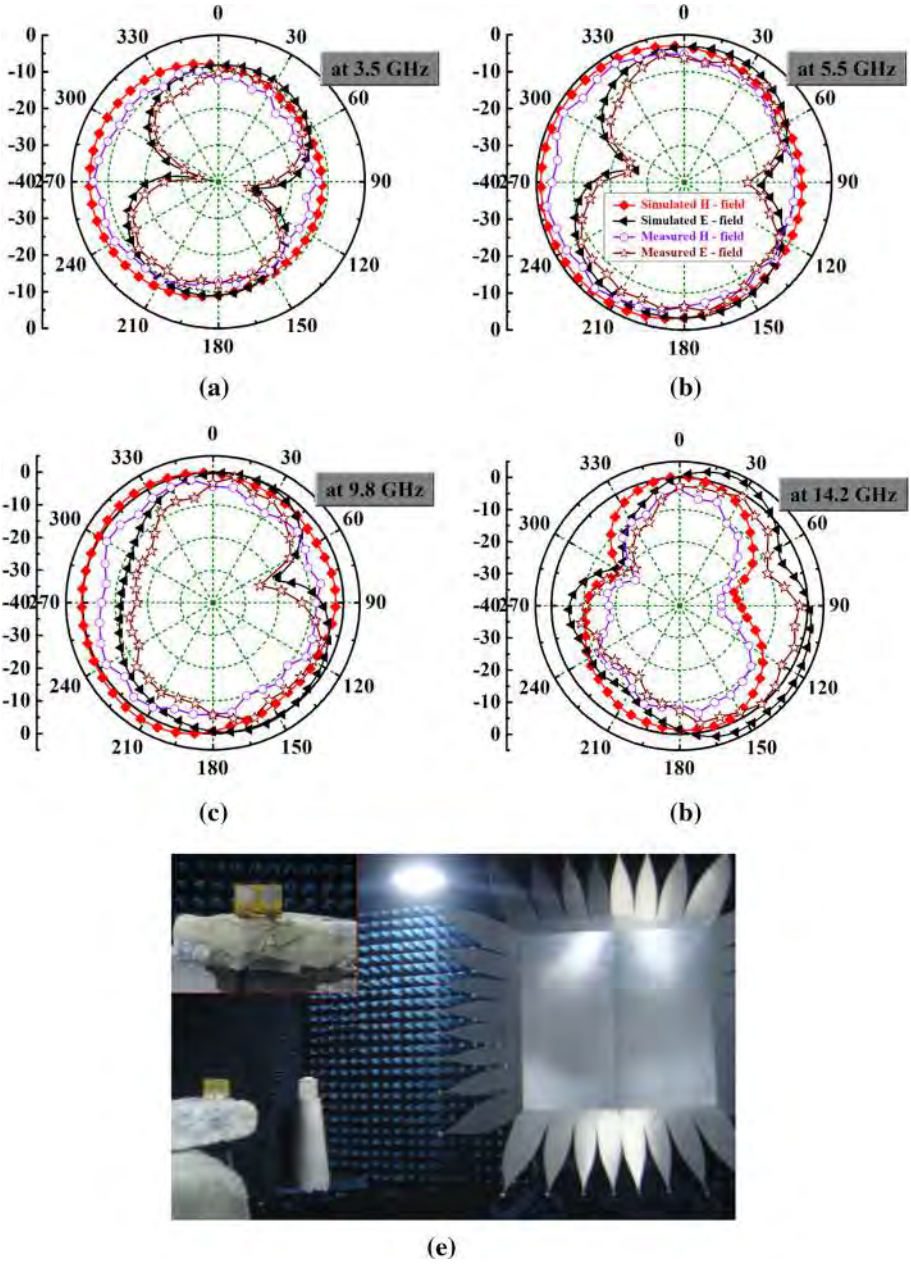


Fig. 9 E and H field radiation patterns: **a** at 3.5 GHz; **b** at 5.5 GHz; **c** at 9.8 GHz; **d** at 14.2 GHz and **e** Photograph of fabricated model in an anechoic chamber while measuring radiation patterns

through ECC value. If two radiators are more isolated, the ECC value is low (Ideally zero) and if more coupling is present between the antennas, then ECC value is high (Ideally one). But, practical acceptable value is 0.5 and it is a real number as it is a

magnitude square of signal correlation [29, 30]. The ECC value can be measured easily with S- parameter values and it is represented in Eq. (2). Diversity gain is also an important metric to evaluate the diversity performance of MIMO antenna. The relation between ECC and DG is represented in Eq. (3). Figure 10a, shows simulated values of ECC and DG parameters. The proposed model's ECC and DG values are lower than 0.06 and higher than 9.975 dB respectively.

$$ECC = \frac{\left|S_{11}^{*}S_{12} + S_{21}^{*}S_{22}\right|^{2}}{\left(1 - |S_{11}|^{2} - |S_{21}|^{2}\right)\left(1 - |S_{22}|^{2} - |S_{12}|^{2}\right)} \tag{2}$$

$$DG = 10\sqrt{1 - ECC^{2}} \tag{3}$$

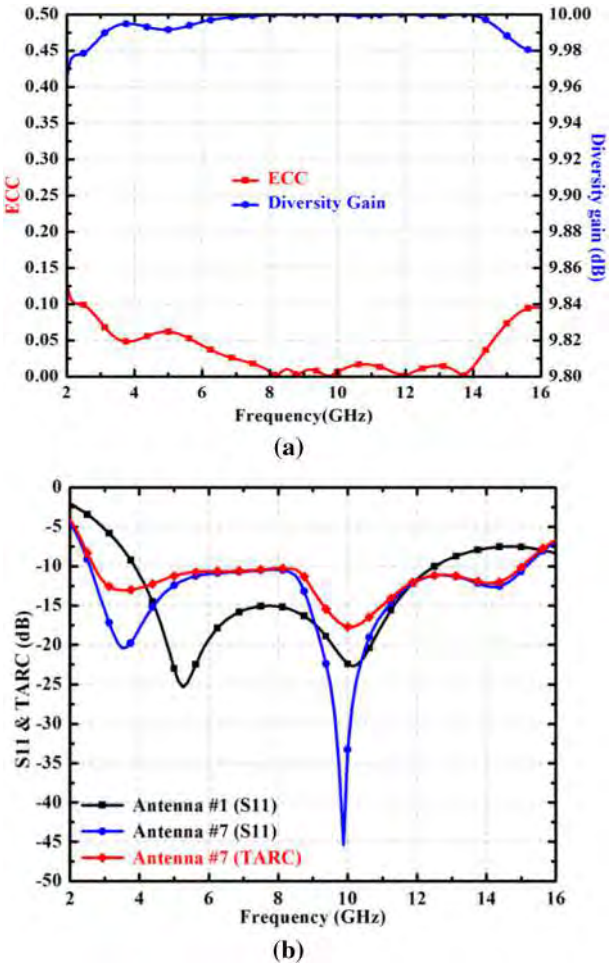


Fig. 10 **a** Simulated ECC and DG Values; **b** Simulated S_{11} and TARC values

In MIMO systems, when two radiators are operated at a time, the radiation of one radiator will affect the performance of the other radiator. This is due to changes in self impedances and mutual impedances. So, the S-parameter values of the individual radiator are not sufficient to evaluate the system performance. The entire system performance can be evaluated by considering changes in self and mutual impedances one term introduced named as total active reflection coefficient (TARC). It is defined as the square root of the difference between all the available power at ports and applied power, divide by the applied power [31] and it is represented in Eq. (4). TARC in terms of S-parameters for the two-element MIMO antenna is represented in Eq. (5). Figure 10b, describes the comparison between S_{11} values of single radiator and S_{11} & TARC values of the proposed model. The acceptable TARC values for MIMO systems is below 0 dB (TARC < 0 dB). The proposed model gives good TARC values for the entire impedance bandwidth, which is below − 10 dB.

$$TARC = \sqrt{\frac{\text{Available power} - \text{Radiated power}}{\text{Available power}}} \tag{4}$$

$$TARC = \sqrt{\frac{\left|\left(S_{11} + S_{12}e^{j\theta}\right)^2\right| + \left|\left(S_{21} + S_{22}e^{j\theta}\right)^2\right|}{2}} \tag{5}$$

The performance comparison of the proposed model with other models is listed in Table 1 in terms of size of the antenna, dielectric constant (εr), the spacing between antenna elements, impedance bandwidth, isolation, peak gain, radiation efficiency, ECC, and TARC with references [5–8, 10–15, 17, 20–23, 25 and 27] for performance comparison. The proposed model has more advantages like compactness, wider impedance bandwidth, high isolation, high-efficiency, low ECC and high DG compared with the latest literature. In most of the references, TARC values are not represented. But, TARC is an important parameter to evaluate the overall MIMO system performance. It considers the changes in individual antenna's impedances due to adjacent elements and changes in mutual impedances. So, S- parameters of individual antennas are not sufficient for analyze the overall system performance. The proposed model gives good TARC values, which are below − 10 dB (Acceptable TARC values are below 0 dB). Hence, the proposed model is suitable choice for the portable wireless device applications.

5 Conclusion

A Compact MIMO antenna of size 16 × 26 mm2 (0.144 λ0 × 0.234 λ0) is presented in this paper. The elements in the model are separated with a distance of 2 mm (0.018λ0). The proposed model gives an impedance bandwidth ($|S_{11}|$ >10 dB) of 2.7 to 14.9 GHz. Two parallel and four strips are protruded to the ground plane for the isolation enhancement. Due to these elements, the isolation ($|S_{21}|$) is improved to above 20 dB. The radiation efficiency and peak gain values are above 86% for the entire band and 0.8 to 6.6 dBi respectively. The diversity performance is checked with metrics ECC, DG and TARC values. These are below 0.06, above 9.975 dB and below − 10 dB respectively. The proposed model covers the entire UWB, X- band and Ku band applications.

Table 1 Performance comparison of proposed model with other models

Ref. no	Antenna size (mm^2)	εr	Elements separation (λ_0)	Impedance bandwidth (GHz)	Isolation $\lvert S_{21} \rvert$ (dB)	Peak gain (dBi)	R.E (%)	ECC	TARC (dB)
#05	70 × 120	3.02	0.02	1.6–4	≥ 20	3.8	–	–	–
#06	47 × 93	4.4	0.35	3.1–10.6	≥ 31	3.5	≥ 75	− 29 dB	–
#07	50 × 60	4.5	0.50	3–6	≥ 21	–	–	–	–
#08	45 × 45	4.4	–	2–10.6	≥ 17	3	–	0.005	–
#10	37.5 × 40	4.4		2–11	≥ 20	–	–	0.1	≤ − 4
#11	38.5 × 38.5	4.4	–	3.08–11.8	≥ 15	1.4–3.6	≥ 75	-	–
#12	35 × 36	2.2	0.12	3–9	≥17	1–5.5	≥ 87	0.02	–
#13	35 × 35	4.4	–	3–12	≥ 20	–	–	0.40	–
#14	30 × 40	4.4	–	3.1–10.6	≥ 16	–	≥ 70	0.15	–
#15	24 × 42	4.3	–	3.1–10.9	≥ 17	1.8–3.5	≥ 75	0.10	–
#17	20 × 40	4.4	–	3–11	≥ 15	1.5–5	≥ 80	0.8	–
#20	26 × 31	4.3	–	3.1–11.12	≥ 20	2.5–5.54	–	0.002	–
#21	22 × 36	3.5	–	3.1–11	≥ 15	1–5	≥ 70	0.1	–
#22	20 × 34	4.4	–	3.1–11	≥ 20	3–4	≥ 75	0.2	–
#23	18 × 36	4.4	–	2.8–20	≥ 20	1.6–6	–	0.02	≤ − 10
#25	21× 27	4.4	0.15	5.1–5.4 & 7.3–7.6	≥ 22.5	> 9.38	≥ 79.8	0.1	≤ − 10
#27	18 × 23	2.65	–	3–12.4	≥ 20	4	≥ 70	0.1	–
Proposed	16 × 26	4.4	0.018	2.7–14.9	≥ 20	0.8–6.6	≥ 86	0.06	≤ − 10

Acknowledgements Authors feel a need to convey heartfelt thanks to Rakesh Kumar Singh, Scientist 'F' and Head, Stealth technologies and test ranges, Directorate of radar seekers and systems (DRSS), RCI, DRDO, Hyderabad, India for providing lab facilities for testing and measure the radiation patterns of the fabricated antenna as a part of research work.

Authors' contributions All authors read and approved the final manuscript.

Funding For this manuscript there is no funding availability.

Availability of data and material In the manuscript must include an 'Availability of data and materials' statement.

Code availability For this manuscript there is no Code availability.

Declarations

Conflicts of interest The authors declare that they have no conflicts of interest /competing interests" in this section.

References

1. Jayanthi, K., & Kalpana, A. M. (2017). Mutual coupling reduction techniques between MIMO antennas for UWB applications. *International Journal of Recent Innovation Trends Computer Communications* 5(9).
2. Addepalli, T., & Anitha, V. R. (2020). Compact two-port MIMO antenna with high isolation using parasitic reflectors for UWB, X and Ku band applications. *Progress In Electromagnetics Research C, 102*, 63–77.
3. Nadeem, I., & Choi, D.-Y. (2018). Study on mutual coupling reduction technique for MIMO antennas. *IEEE Access, 7*, 563–586.
4. Addepalli, T., & Anitha, V. R. (2020). A very compact and closely spaced circular shaped UWB MIMO antenna with improved isolation. *AEU-International Journal of Electronics and Communications, 114*, 153016.
5. Isaac, A. A., Al-Rizzo, H., Yahya, S., Al-Wahhamy, A., & Abushamleh, S. (2018). Decoupling of two closely-spaced planar monopole antennas using two novel printed-circuit structures. *Microwave and Optical Technology Letters, 60*(12), 2954–2963.
6. Radhi, A. H., Nilavalan, R., Wang, Y., Al-Raweshidy, H. S., Eltokhy, A. A., & Ab Aziz, N. (2018). Mutual coupling reduction with a wideband planar decoupling structure for UWB–MIMO antennas. *International Journal of Microwave and Wireless Technologies, 10*(10), 1143–1154.
7. Li, Q., Feresidis, A. P., Mavridou, M., & Hall, P. S. (2015). Miniaturized double-layer EBG structures for broadband mutual coupling reduction between UWB monopoles. *IEEE Transactions on Antennas and Propagation, 63*(3), 1168–1171.
8. Tripathi, S., Mohan, A., & Yadav, S. (2015). A compact Koch fractal UWB MIMO antenna with WLAN band-rejection. *IEEE Antennas and wireless propagation letters, 14*, 1565–1568.
9. Chithradevi, R., & Sreeja, B. S. (2017). A compact UWB MIMO antenna with high isolation and low correlation for wireless applications. In: *2017 IEEE International Conference on Antenna Innovations & Modern Technologies for Ground, Aircraft and Satellite Applications (IAIM)*. IEEE.
10 Jafri, S. I., Saleem, R., Shafique, M. F., & Brown, A. K. (2016). Compact reconfigurable multiple-input-multiple-output antenna for ultra-wideband applications. *IET Microwaves, Antennas & Propagation, 10*(4), 413–419.
11. Kang, L., Li, H., Wang, X., & Shi, X. (2015). Compact offset microstrip-fed MIMO antenna for band-notched UWB applications. *IEEE Antennas and Wireless Propagation Letters, 14*, 1754–1757.
12. Park, J.-D., Rahman, M., & Chen, H. N. (2019). Isolation enhancement of wide-band MIMO array antennas utilizing resistive loading. *IEEE Access, 7*, 81020–81026.
13. Zhu, J., Li, S., Feng, B., Deng, L., & Yin, S. (2015). Compact dual-polarized UWB quasi-self-complementary MIMO/diversity antenna with band-rejection capability. *IEEE Antennas and Wireless Propagation Letters, 15*, 905–908.

14. Deng, J.-Y., Guo, L.-X., & Liu, X.-L. (2015). An ultra-wideband MIMO antenna with a high isolation. *IEEE antennas and wireless propagation letters, 15*, 182–185.
15 Alsath, M. G. N., & Kanagasabai, M. (2015). Compact UWB monopole antenna for automotive communications. *IEEE Transactions on Antennas and Propagation, 63*(9), 4204–4208.
16. Zhu, J., Feng, B., Peng, B., Deng, L., & Li, S. (2016). A dual notched band MIMO slot antenna system with Y-shaped defected ground structure for UWB applications. *Microwave and Optical Technology Letters, 58*(3), 626–630.
17. Mathur, R., & Dwari, S. (2019). Compact planar reconfigurable UWB-MIMO antenna with on-demand worldwide interoperability for microwave access/wireless local area network rejection. *IET Microwaves, Antennas & Propagation, 13*(10), 1684–1689.
18. Zhao, L., & Ke-Li, Wu. (2015). A dual-band coupled resonator decoupling network for two coupled antennas. *IEEE Transactions on Antennas and Propagation, 63*(7), 2843–2850.
19. Wang, S., & Zhengwei, Du. (2014). Decoupled dual-antenna system using crossed neutralization lines for LTE/WWAN smartphone applications. *IEEE Antennas and Wireless Propagation Letters, 14*, 523–526.
20. Mchbal, A., Amar Touhami, N., Elftouh, H., & Dkiouak, A. (2018). Mutual coupling reduction using a protruded ground branch structure in a compact UWB owl-shaped MIMO antenna. *International Journal of Antennas and Propagation, 2018*.
21 Liu, Li., Cheung, S. W., & Yuk, T. I. (2015). Compact MIMO antenna for portable UWB applications with band-notched characteristic. *IEEE Transactions on Antennas and Propagation, 63*(5), 1917–1924.
22. Mathur, R., & Dwari, S. (2019). A compact UWB-MIMO with dual grounded CRR for isolation improvement. *International Journal of RF and Microwave Computer-Aided Engineering, 29*(1), e21500.
23. Chandel, R., Gautam, A. K., & Rambabu, K. (2018). Design and packaging of an eye-shaped multiple-input–multiple-output antenna with high isolation for wireless UWB applications. *IEEE Transactions on Components, Packaging and Manufacturing Technology, 8*(4), 635–642.
24. Chandel, R., Gautam, A. K., & Rambabu, K. (2018). Tapered fed compact UWB MIMO-diversity antenna with dual band-notched characteristics. *IEEE Transactions on Antennas and Propagation, 66*(4), 1677–1684.
25. Dkiouak, A., Zakriti, A., & El Ouahabi, M. (2020). Design of a compact dual-band MIMO antenna with high isolation for WLAN and X-band satellite by using orthogonal polarization. *Journal of Electromagnetic Waves and Applications, 34*(9), 1254–1267.
26. Zhang, S., & Pedersen, G. F. (2015). Mutual coupling reduction for UWB MIMO antennas with a wideband neutralization line. *IEEE antennas and wireless propagation letters, 15*, 166–169.
27. Tao, J., & Feng, Q. (2016). Compact ultra-wideband MIMO antenna with half-slot structure. *IEEE Antennas and wireless Propagation letters, 16*, 792–795.
28 Ray, K. P. (2008). Design aspects of printed monopole antennas for ultra-wide band applications. *International journal of antennas and propagation, 2008*, 1–8.
29. Sharawi, M. S., Hassan, A. T., & Khan, M. U. (2017). Correlation coefficient calculations for MIMO antenna systems: A comparative study. *International Journal of Microwave and Wireless Technologies, 9*(10), 1991–2004.
30. Addepalli, T., & Anitha, V. R. (2019). Design and parametric analysis of hexagonal shaped MIMO patch antenna for S-band, WLAN, UWB and X-band applications. *Progress In Electromagnetics Research C, 97*, 227–240.
31. Fritz-Andrade, E., Jardon-Aguilar, H., & Tirado-Mendez, J. A. (2020). The correct application of total active reflection coefficient to evaluate MIMO antenna systems and its generalization to N ports. *International Journal of RF and Microwave Computer-Aided Engineering, 30*(4), e22113.

Tathababu Addepalli was born in 1986. He received his B. Tech degree from JNTUH, Hyderabad in 2007 and M. Tech degree from JNTUK, Kakinada in 2010. Currently, he is a full-time research scholar in JNTUA, Anantapuramu and registered in 2017. He has published more than five research papers in various reputed international journals and conferences like ELSEVIER (AEU), PIER, Wiley (IJCS), MDPI (Electronics), IEEE and SPRINGER conferences, etc. His areas of interest are microstrip patch antennas, MIMO antennas, UWB antennas, Characteristic Mode Analysis(CMA), THz antennas and metamaterial antennas.

V.R. Anitha has been working as a Professor, Department of ECE in BMS Institute of Technology & Management, Bangalore, Karnataka, INDIA. She obtained Bachelor of Technology in the Department of Electronics and Telecommunication Engineering from Institute of Electronics and Tele Communications Engineering (IETE), New Delhi and Post Graduate in Microwave Engineering from Nagarjuna University Guntur, and Ph.D. in Microwave Engineering from Sri Venkateswara University, Tirupati. She is a member in IEEE, IETE and ISSS and published 38 papers in national and international journals and conferences and presented many technical lectures in various scientific forums (International & National). She was invited to deliver lectures in various Institutions. She received Travel Grant award from URSI General Assembly-2011, Istanbul, Turkey. She was session chair in International Conference on ICEAA-IEEE APWC-EMS'13, Torino, Italy. She received "Young Scientist" grant from DST for her presentation held at Kuala Lumpur, Malaysia in October 2013. Presently she completed couple of research projects under INUP and NPMASS and she is holding research projects from various funding agencies. She is doing collaborative research with Yeungnam University, Republic of Korea and CSIR-NAL. Dr. Anitha is continuously upgrading her expertise in the areas of Electrically Small Antennas (ESA), Microstrip Antennas, Fractal Antennas and MEMS. Dr.Anitha is instrumental in establishing a National MEMS Design Centre at Sree Vidyanikethan Engineering College under the aegis of NPMASS. Competence in handling COMSOL, CoventorWare and Intellisuite will provide an ideal environment for conducting simulation studies in the area of MEMS, HFSS, ADS & IE3D will provide platform for conducting simulation studies in the area of Antennas & Microwaves. She played a key role to obtain Fund of Rs. 20 laks for Improvement of S&T Infrastructure (FIST) from DST during 2012–13. Also, she is holding few funded projects from government agencies. She is serving as Ph.D guide, BOS member, reviewer, editor and DC member.

Terms and Conditions

Springer Nature journal content, brought to you courtesy of Springer Nature Customer Service Center GmbH (“Springer Nature”).

Springer Nature supports a reasonable amount of sharing of research papers by authors, subscribers and authorised users (“Users”), for small-scale personal, non-commercial use provided that all copyright, trade and service marks and other proprietary notices are maintained. By accessing, sharing, receiving or otherwise using the Springer Nature journal content you agree to these terms of use (“Terms”). For these purposes, Springer Nature considers academic use (by researchers and students) to be non-commercial.

These Terms are supplementary and will apply in addition to any applicable website terms and conditions, a relevant site licence or a personal subscription. These Terms will prevail over any conflict or ambiguity with regards to the relevant terms, a site licence or a personal subscription (to the extent of the conflict or ambiguity only). For Creative Commons-licensed articles, the terms of the Creative Commons license used will apply.

We collect and use personal data to provide access to the Springer Nature journal content. We may also use these personal data internally within ResearchGate and Springer Nature and as agreed share it, in an anonymised way, for purposes of tracking, analysis and reporting. We will not otherwise disclose your personal data outside the ResearchGate or the Springer Nature group of companies unless we have your permission as detailed in the Privacy Policy.

While Users may use the Springer Nature journal content for small scale, personal non-commercial use, it is important to note that Users may not:

1. use such content for the purpose of providing other users with access on a regular or large scale basis or as a means to circumvent access control;
2. use such content where to do so would be considered a criminal or statutory offence in any jurisdiction, or gives rise to civil liability, or is otherwise unlawful;
3. falsely or misleadingly imply or suggest endorsement, approval , sponsorship, or association unless explicitly agreed to by Springer Nature in writing;
4. use bots or other automated methods to access the content or redirect messages
5. override any security feature or exclusionary protocol; or
6. share the content in order to create substitute for Springer Nature products or services or a systematic database of Springer Nature journal content.

In line with the restriction against commercial use, Springer Nature does not permit the creation of a product or service that creates revenue, royalties, rent or income from our content or its inclusion as part of a paid for service or for other commercial gain. Springer Nature journal content cannot be used for inter-library loans and librarians may not upload Springer Nature journal content on a large scale into their, or any other, institutional repository.

These terms of use are reviewed regularly and may be amended at any time. Springer Nature is not obligated to publish any information or content on this website and may remove it or features or functionality at our sole discretion, at any time with or without notice. Springer Nature may revoke this licence to you at any time and remove access to any copies of the Springer Nature journal content which have been saved.

To the fullest extent permitted by law, Springer Nature makes no warranties, representations or guarantees to Users, either express or implied with respect to the Springer nature journal content and all parties disclaim and waive any implied warranties or warranties imposed by law, including merchantability or fitness for any particular purpose.

Please note that these rights do not automatically extend to content, data or other material published by Springer Nature that may be licensed from third parties.

If you would like to use or distribute our Springer Nature journal content to a wider audience or on a regular basis or in any other manner not expressly permitted by these Terms, please contact Springer Nature at

onlineservice@springernature.com

Received: 18 February 2022 | Revised: 15 May 2022 | Accepted: 10 August 2022

DOI: 10.1002/dac.5325

RESEARCH ARTICLE

WILEY

High-gain broadband ENZ metasurface-based RF lens with gradient refractive index for microwave and millimetre-wave applications

Sri Pravallika Narjala[1] | **Anitha V R**[2] | **RamaNaidu K**[1]

[1]Department of ECE, JNTUA, Anantapur, India

[2]Department of ECE, BMS Institute of Technology & Management, Bangalore, India

Correspondence
Sri Pravallika Narjala, Department of ECE, JNTUA, Anantapur, AP, India.
Email: sripravallika37@gmail.com

Summary

Design, simulation and measurement of a broadband epsilon-near-zero (ENZ) metasurface-based gradient refractive index (GRIN) Luneberg lens has been presented with different antennas demonstrating gain enhancement and producing highly directive beams. The unit-cell meta-structure consists of a conducting ring with monopoles at 45°. The horizontal radius of the conducting ring is varied to vary the refractive index of the structure. Unit cells with varying refractive index are arranged in a 2D plane to produce a planar Luneberg lens structure. The lens is demonstrated to produce a gain improvement of 6–13.2 dBi for different configurations and different antennas including a microstrip patch antenna and horn antenna both in simulation and measurement while maintaining the total efficiency of the antenna being used with the lens.

KEYWORDS

metamaterial, metasurface antenna, microwave lens, planar lens

1 | INTRODUCTION

Lens antennas have been heavily discussed in the past two decades for their applications in microwave and millimetre-wave communication. Lens antennas have been shown in the literature to produce highly directive beams in the desired direction. Similar to the optical lenses, RF lenses work on the convergence and divergence properties of the beam passing through them. As the beam passes through the lens, the change in medium changes the velocity of the beam, and hence, the radiated wavelet can be manipulated to reshape the beam and focus energy in one direction.[1–3]

A different form of lenses has been proposed in the literature. Dielectric lenses are one of the most used because of their cost factor and ease in fabrication.[4,5] A common example of integrated lens antennas is dielectric resonator antennas (DRAs) use dielectric shapes on top of a radiator to reshape the beam.[6,7] They, however, tend to be bulky and hard to maintain for long-term deployments, which has led to exploring the planar RF lens structures.[8] Planar lenses are built out of an array of microstrip planar structures on a dielectric slab, which is then illuminated with a wavelet from a radiating structure, creating a full antenna system. These planar lenses are easy to fabricate, cost-friendly and a good fit to produce highly directive beams.

Several approaches for a planar lens antenna are seen in the literature. A lot of work is done using metamaterials (MTMs) and frequency selective surfaces (FSSs) where the lens is comprised of several unit cells arranged to manipulate the wavelet passing through the structure. Multiple layers are then stacked together separated with air gaps (or another dielectric material) to increase the gain enhancement properties of the lens. Several different versions of FSSs and MTMs have been explored recently, including dipole-shaped,[9–11] slot-shaped[12,13] and circular-shaped[14–17] microstrip

structures. However, the FSS-based structures tend to behave as a bandpass filter and hence limit the operation of the lens over a narrow bandwidth.

Metasurface-based lenses recently have found a lot of attraction because of their small size and their ability to manipulate the electromagnetic (EM) waves. Several metasurface-based lens designs[18–22] have been explored for their reflectionless characteristic, allowing to reduce the insertion loss introduced in FSS-based structures significantly. Metamaterial unit cells are used to create a gradient refractive index (GRIN) surface, which can be used to manipulate the wavelet shape illuminated to it. GRIN-based arrangement allows to overcome the narrow bandwidth and high insertion loss issues that are generally faced in FSS-based structures. Several reflector-based AMC antennas are also well known for gain enhancement of antennas; however, they are a different class and tend to be much larger in size.[23,24]

Adding to the research contributions in GRIN-based metasurface lenses for antennas, this paper demonstrates a new epsilon-near-zero (ENZ) metasurface-based multilayer planar Luneberg lens. The proposed lens structure is demonstrated to be extremely compact and easy to fabricate. The lens has been tested in simulation and measurement with different antennas to demonstrate a significant increase in the antenna's directivity and gain while maintaining the efficiency of the antenna system.

The rest of the paper is organised as follows: Section 2 describes the geometry of the lens and its working principle as a GRIN lens, followed by a one-dimensional (1D) and two-dimensional (2D) metasurface-based lens tested with a microstrip patch antenna. The lens performance is explained with the simulated and measured results. Section 3 then describes the proposed 2D metasurface lens with a horn antenna to present the effect of lens with change in frequency bands, followed by parametric analysis of the lens in Section 4. Finally, a conclusion is presented in Section 5.

2 | GRIN METAMATERIAL LENS DESIGN

Lens antennas have been heavily discussed in the past. GRIN lens being derived from Luneberg lens, they generally have high permittivity material in the centre and low permittivity material in the corners. A similar technique has been used in the proposed lens, where a unit-cell MTM has been designed and demonstrated to have a near-zero permittivity. The proposed ENZ metamaterial (MTM) is shown to have varying permittivity with the change in its dimensions. This section describes the proposed unit-cell MTM and then presents the MTM lens process using the proposed unit cell.

2.1 | Unit-cell design

The proposed unit-cell MTM design is shown in Figure 1. It is designed on a 6 × 6 mm FR-4 dielectric substrate of relative permittivity, $\varepsilon_r = 4.3$; loss tangent $tan\ \delta = 0.001$; and thickness, $h = 0.8$ mm. The design is built of a single-ring

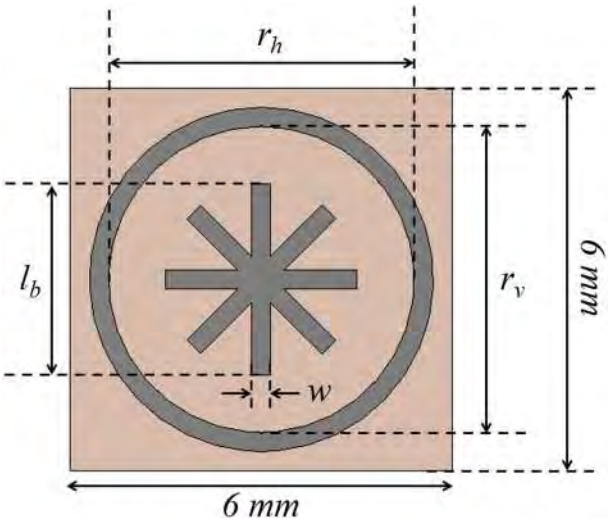


FIGURE 1 Geometry of the proposed unit-cell metamaterial

resonator covering four dipole-like bars of length, $l_b = 3$ mm, at $\theta = 45°$. The split-ring resonator has a horizontal and vertical radius, $r_h = r_v = 5$ mm, with the width of ring being $w = 0.3$ mm.

The ring resonator on the dielectric sheet helps in creating a magnetic resonance, while the bars help to create electrical resonance from different direction. The two magnetic and electric resonance can be controlled with the change in the dimensions of conducting bar and ring. This helps in creating a very relative permittivity in the material, which, when arranged in specific order, can be used to manipulate the current density and hence the E-field distribution as the wavelet passes through a sheet made from an array of the proposed unit cell.

The simulated tangential S parameter and phase plot for the proposed unit cell are shown in Figure 2. The S parameters can be used to calculate the relative permittivity of the artificial substrate created with the MTM sheet using the Nicolson–Ross–Weir (NRW) method.

NRW method provides a direct process to calculate the relative permittivity and permeability of the material under test by using S_{11} and S_{21} of the material, simulated with ports defined on either side or by placing horn antennas on either side and measuring the same using a vector network analyser (VNA). Using the S_{11} and S_{21}, one can calculate the reflection coefficient (Γ), and transmission coefficient (T) can be calculated using Equations (1) and (3), respectively.[25]

$$\Gamma = X \pm \sqrt{X^2 - 1} \tag{1}$$

where X can be defined as

$$X = \frac{S_{11}^2 - S_{21}^2 + 1}{2S_{11}} \tag{2}$$

$$T = \frac{S_{11} + S_{21} - \Gamma}{1 - (S_{11} + S_{21})\Gamma} \tag{3}$$

The reflection and transmission coefficients defined by Equations (1) and (3) can be used to calculate the permittivity and permeability of the material as follows:

$$\mu_r = \frac{1 + \Gamma_1}{\Lambda(1 - \Gamma)\sqrt{\frac{1}{\lambda_0^2} - \frac{1}{\lambda_c^2}}} \tag{4}$$

where λ_0 is the free space wavelength, λ_c is the cut-off wavelength and

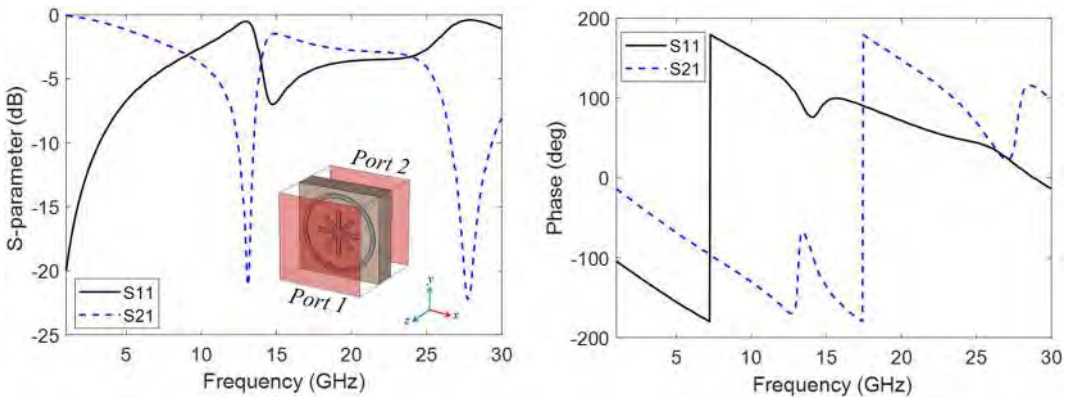


FIGURE 2 Simulated tangential S parameters and phase for the signal travelling tangentially from Port 1 to Port 2

$$\frac{1}{\Lambda^2} = -\left(\frac{1}{2\pi L}\ln\left(\frac{1}{T}\right)\right)^2 \tag{5}$$

The relative permittivity of the material is defined as

$$\varepsilon_r = \frac{\lambda_0^2}{\mu_r}\left(\frac{1}{\lambda_c^2} + \frac{1}{\Lambda^2}\right) \tag{6}$$

The relative permittivity and permeability calculated using Equation (6) are presented in Figure 3. And as observed, the relative permittivity is found to be very close to zero, for a wideband of 10–25 GHz, with maximum value being $\varepsilon_{max} = 0.064$, and hence, the material can be titled as an ENZ MTM substrate. The material is also tested for different configurations and is found to be reducing relative permittivity with increase in the horizontal radius of the ring. The same change can be observed with the change in vertical radius because of the symmetry.

The upper and lower cut-off frequencies for the proposed ENZ MTM unit cell can be seen through the S_{11} shown in Figure 2. The 3 dB wideband characteristic matches to the near-zero property of the MTM substrate shown in Figure 3. This wideband characteristic can be controlled with the length of dipole bars and hence can be reconfigured for any required frequency band.

Further, the relative permittivity of the material is found to decrease with the change in its horizontal (or vertical because of symmetry) radius. The same has been shown in Figure 4. The horizontal radius has been varied from 2 to

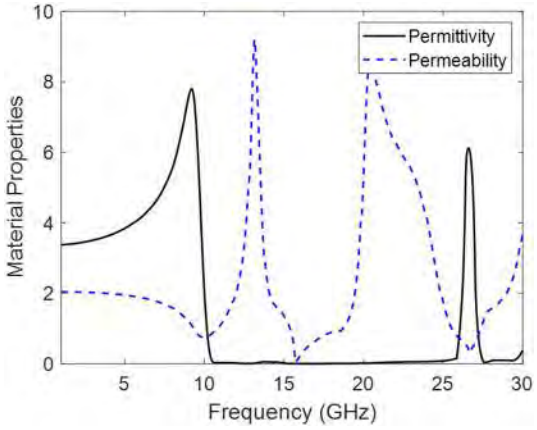


FIGURE 3 Calculated permittivity (using Equation 6) and permeability (using Equation 4) of the proposed metamaterial sheet

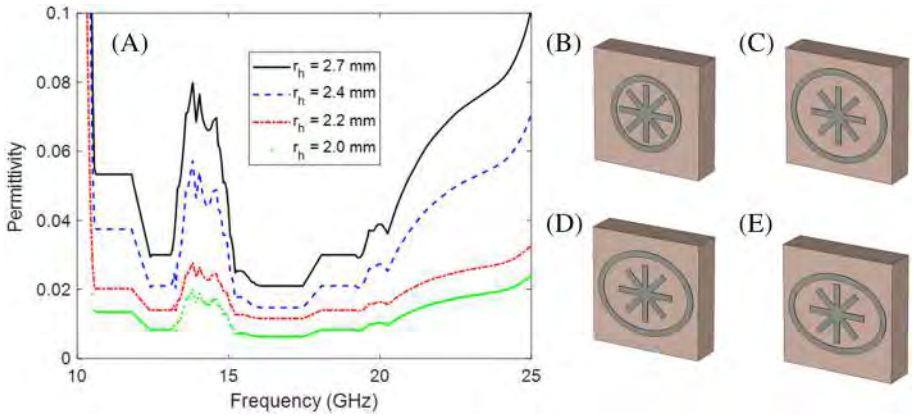


FIGURE 4 (A) Variation in relative permittivity with the change in horizontal radius of the ring in the proposed unit cell. Proposed unit cell with (B) $r_h = r_v = 2.0$ mm, (C) $r_h = 2.7$ mm and $r_v = 2.4$ mm, (D) $r_h = 2.7$ mm and $r_v = 2.2$ mm, and (E) $r_h = 2.7$ mm and $r_v = 2.0$ mm

2.7 mm, parallel to which the relative permittivity shows a proportional decrease in the relative permittivity of the metasurface. This decreasing relative permittivity of the material can be used to create a dynamic metasurface with varying relative permittivity, creating a virtual planar Luneberg structure, which can behave as a lens.

The optical convex lens works on the principle of slowing down the optical wave at different regions to make the spherical wave planar as it exits the lens structure. The same principle is used here, where the lens is designed to have a high permittivity material at the centre and lower permittivity material on the sides. The higher the permittivity, the slower the wavelet gets as it passes through the material and hence allows us to receive a much planar beam. Hence, several planar architectures were built with the four unit-cell MTMs shown in Figure 4B–D, to achieve the optimal lens design. These different configurations have been discussed in detail in the sections following.

2.2 | Single-layer 1D planar lens design

Planar lenses are designed irrespective of the frequency band but based on the focal length of the lens and the size of the antenna. However, MTM-based lenses can be said to be frequency based because they are artificially created to work on a selected bandwidth or set of frequency bands. These lenses are designed on the same principle described in Giddens and Hao,[26] except the varying width of the lens has been substituted with the varying permittivity of the metasurface.[27]

Similar to a GRIN lens, the 1D lens is simply made with variation in permittivity of the material in only one direction. This creates a gradient variation of dielectric constant in one direction, which allows the beam to be made more focused in the direction of the variation of material. The proposed simulated and fabricated 1D lens is shown in Figure 5 where the lens has high permittivity MTM at the centre, and it keeps reducing as we move further in $\pm y$ axis. This creates a virtual 1D convex structure for the EM wave. The lens is designed for a focal length of $f_L = 20$ mm.

The proposed lens was tested with a patch antenna operating at 15 GHz within simulation. The antenna under test (AUT) was designed based on Bansal and Gupta[28]; is a rectangular patch of length, $l_p = 4.5$ mm, and width, $w_p = 7$ mm, with a 50 Ω inset microstrip feed; and fabricated on a FR-4 substrate of height, $h = 1.6$ mm. The patch antenna and its S parameters and efficiency are shown in Figure 6. The patch antenna has a 10 dB bandwidth of 2500 MHz and has an efficiency greater than 3 dB throughout the operating frequency band. The lens was placed at a distance of $f_L = 20$ mm from the microstrip patch antenna. And the change in radiation pattern was observed. The antenna was measured in an anechoic chamber with the single layer of the proposed lens. The measured radiation pattern in the elevation and azimuth planes with and without the single layer of the proposed lens is shown in Figure 7.

The single layer of the proposed 1D lens shows a gain enhancement of 2.2 dBi. It is important to note that the half-power beamwidth (HPBW) is reduced from 100° to 60° only in the elevation plane, and there is no significant change in the HPBW in azimuth plane. This is because the variation in relative permittivity is only in the elevation plane. This gain improvement and the beamwidth is proposed to improve further either with introduction of more layers of the similar lens, creating a thicker but more effective lens, or creating a 2D lens, allowing to create a more focused beam in

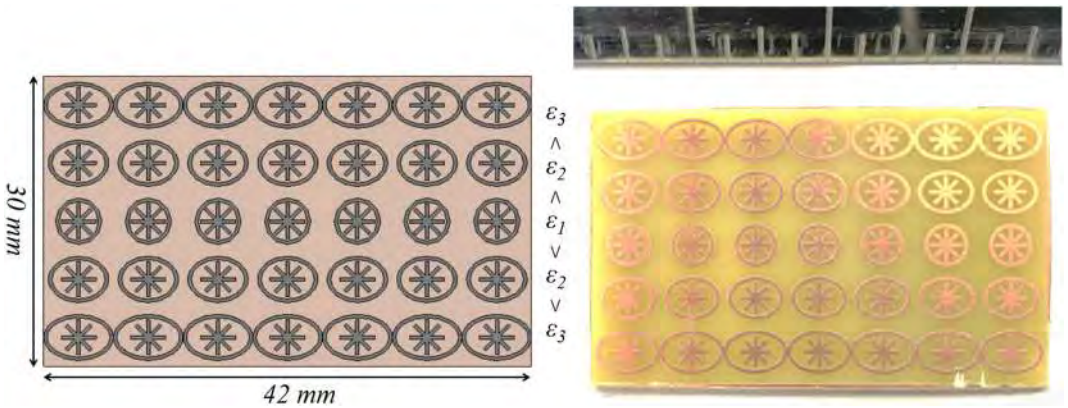


FIGURE 5 Proposed simulated and fabricated one-dimensional metasurface lens, where $\varepsilon_1 = 0.025$, $\varepsilon_2 = 0.018$ and $\varepsilon_3 = 0.012$ at 18 GHz (see Figure 4)

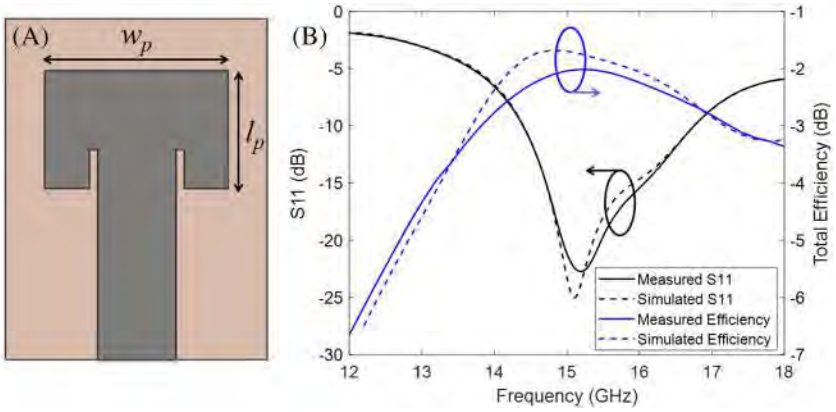


FIGURE 6 (A) Microstrip patch antenna under test. (B) Simulated and measured S_{11} and total efficiency for the microstrip patch antenna

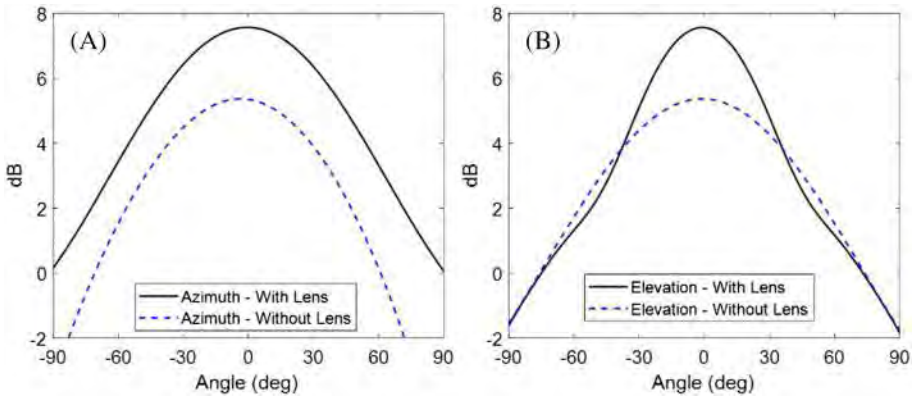


FIGURE 7 2D measured radiation pattern for the microstrip patch antenna at 15 GHz with and without the lens in (A) azimuth and (B) elevation planes

both azimuth and elevation planes, as well as increasing the gain further. The multilayer 1D lens is discussed in detail in Section 2.3 followed by a 2D single-layered and multilayered MTM lens in Section 2.4, respectively.

2.3 | Multilayer 1D planar lens design

The arrangement of the proposed multilayer 1D MTM lens, made up of five layers along with the patch antenna, is shown in Figure 8. The multilayer lens structure behaves as a thick lens, which converts the spherical wavefront to planar front. The layers are kept on top of each other with a spacing of 0.1 mm to ensure easy fabrication and measurement process. However, this spacing has been tested in simulation, which confirmed that there is no significant change in lens performance, until the spacing is greater than 0.25 mm. It allows the antenna to be more focused and hence increases the gain further. This change in the wavefront with the lens can be seen with the electric field plot shown in Figure 9.

This planar wavefront shown in Figure 9 confirms a more directive beam in the elevation plane, which can be confirmed further with the measured radiation pattern in the azimuth and elevation planes shown in Figure 10. The lens shows a gain enhancement of about 6.4 dBi at the resonant frequency and is found to maintain a similar gain enhancement of 6.4 ± 0.2 dBi throughout the operating bandwidth. As described in Section 2.2, the lens being 1D in the elevation plane, the change in HPBW is also only seen in the elevation plane, where it reduced from 100° to 60° for single layer to 27° for multilayer antenna. This improvement in gain has minimalistic effect on the total efficiency of the antenna system and is shown in Figure 11. Hence, it confirms that the proposed lens is highly effective.

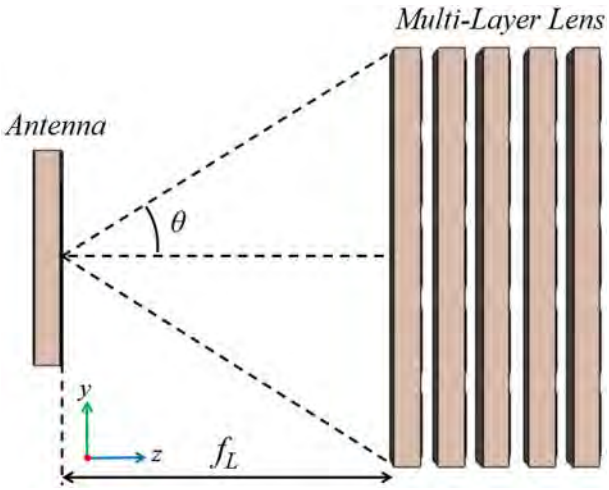


FIGURE 8 Architecture of the multilayer lens with the microstrip patch antenna

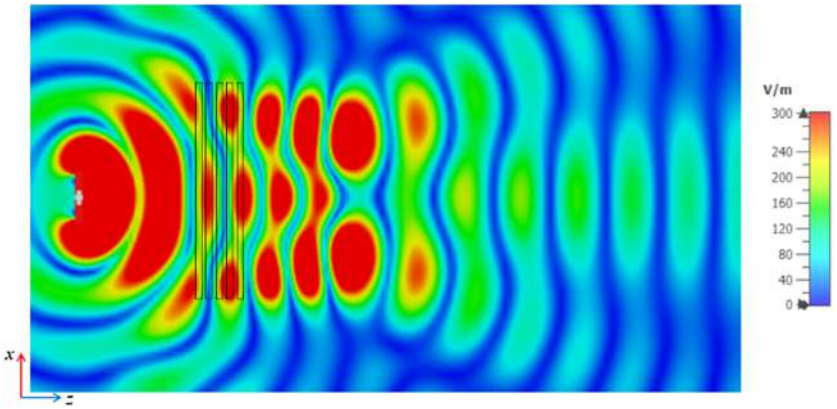


FIGURE 9 Simulated E-field distribution of the antenna system with multilayer metasurface lens producing a planar wavefront

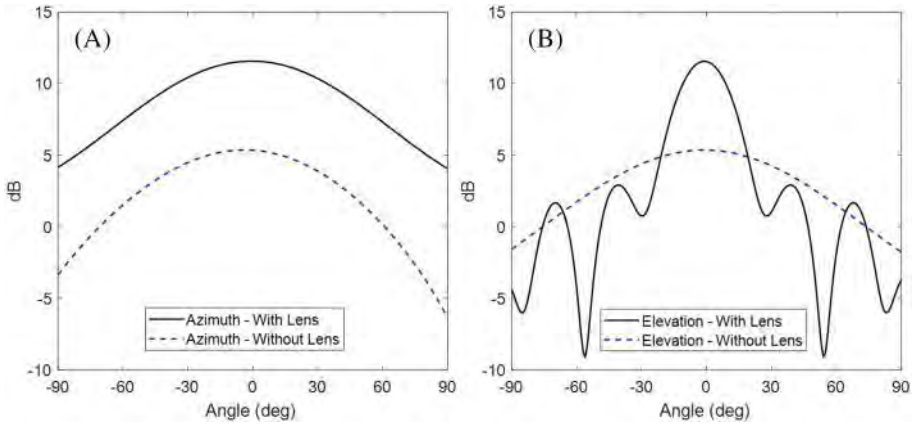


FIGURE 10 Measured radiation pattern for the microstrip patch antenna with the proposed multilayer metasurface lens at 15 GHz in (A) azimuth and (B) elevation planes

Despite several benefits, the antenna lacks in the idea that the beam is highly directional in only one direction, which needs to be improved further. This is proposed to be improved with a multilayer 2D lens, which is described in Section 2.4.

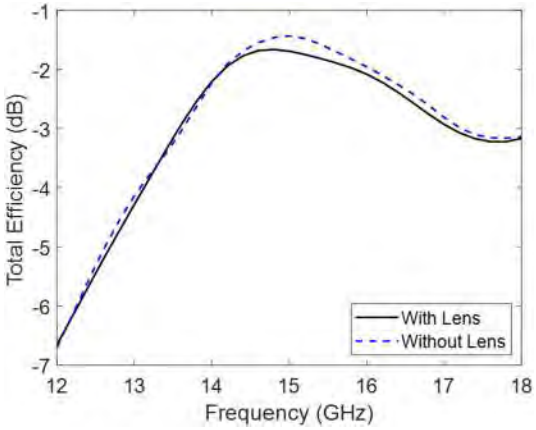


FIGURE 11 Measured total efficiency of the antenna system with and without the proposed multilayered lens

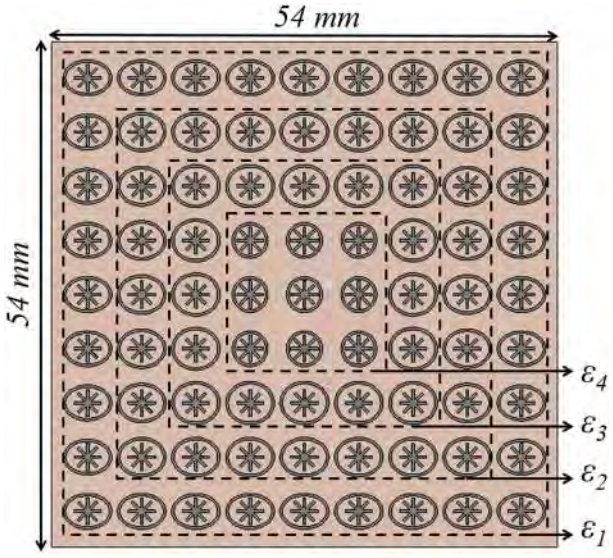


FIGURE 12 Geometry of the proposed two-dimensional metamaterial lens, where $\varepsilon_1 = 0.025$, $\varepsilon_2 = 0.018$, $\varepsilon_3 = 0.012$ and $\varepsilon_4 = 0.007$ at 18 GHz (see Figure 4)

2.4 | Multilayer 2D planar lens design

A 2D planar metasurface lens is similar to what is explained in Section 2.2. The lens is built of five layers of 2D planar lenses. The lens has relative permittivity variation in both azimuth and elevation planes and hence affecting the beam equally at both ends. Such lenses can be created in different architectures such as a concentric rectangular, hexagonal, or circular arrangement. Here, the focus has been on a concentric rectangular structure, where the lens has the highest permittivity in the centre, and it keeps on decreasing as we move outside. This creates an equivalent virtual full convex (or a hemispherical) shape, allowing the beam to be focused in both the azimuth and elevation planes. The proposed architecture for the 2D MTM lens is shown in Figure 12. The lens being symmetrical, it is 54 × 54 × 1.6 mm in dimension with highest relative permittivity in the centre, ε_4, and lowest relative permittivity at the outside, ε_1.

The lens has been shown to be tested with the same microstrip patch antenna described in Section 2.2. The measured radiation pattern of the proposed lens with the microstrip patch antenna is shown in Figure 13. The 2D lens shows a gain improvement of 13.2 dBi and an HPBW reduction of 100° to approximately 25° in both azimuth and horizontal planes. The lens further improves the side lobe level (SLL) of the antenna providing a >15 dB SLL.

Further, the beam can be shown to be more directive because of the lens through the electric field distribution in both azimuth and elevation planes. The simulated E-field distribution in the two planes showing a transition from

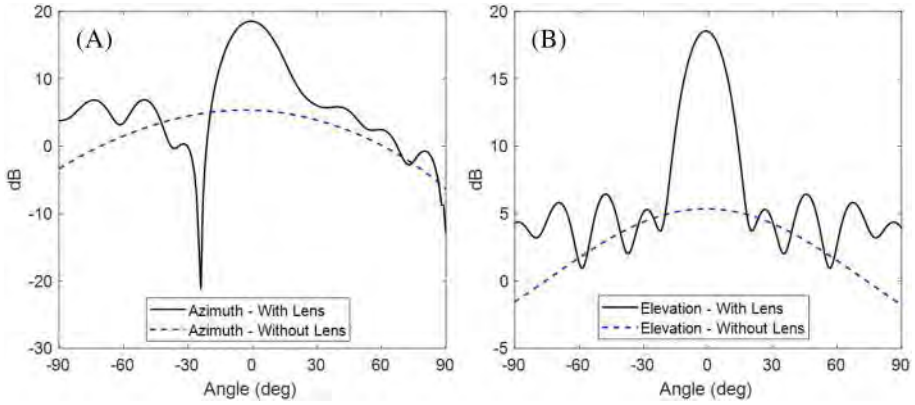


FIGURE 13 Measured radiation pattern for the microstrip patch antenna with the proposed 2D metasurface lens in (A) azimuth and (B) elevation planes

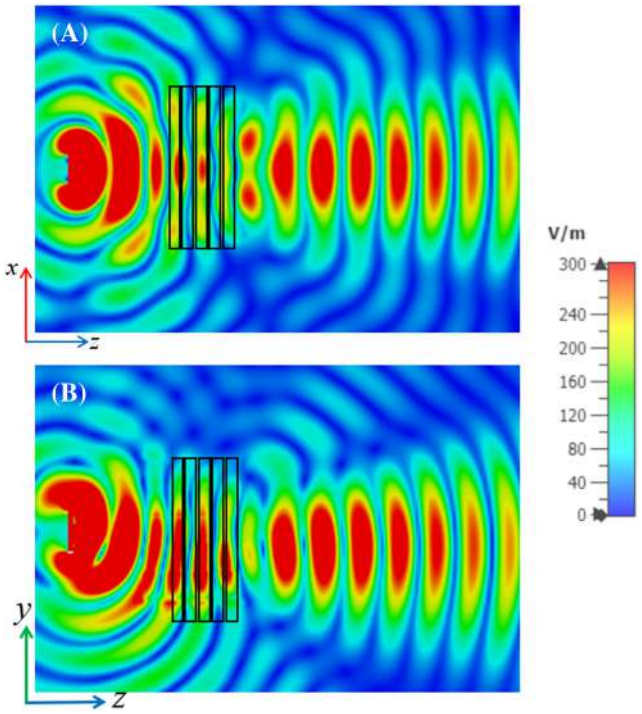


FIGURE 14 Simulated E-field distribution for microstrip patch antenna with a 2D metasurface lens at 15 GHz in (A) azimuth and (B) elevation planes

spherical wavefront to a planar wavefront is demonstrated in Figure 14. This gain improvement is found to be stable than most of the lenses in the literature. The 2D lens again has found to have no effect on the total efficiency of the antenna system. Further, being developed using MTMs, it is important to study their performance with change in frequency band. The same has been discussed in Section 3.

3 | LENS PERFORMANCE WITH HORN ANTENNA

The proposed 2D MTM lens is tested with a wideband horn antenna to study how the change in frequency band affects the performance of the lens. A standard rectangular horn antenna of aperture height, $h_a = 57$ mm, and aperture width, $w_a = 75$ mm, operating at a band of 12–18 GHz has been used to test with the proposed metasurface lens. The S parameters and the gain of the horn antenna are shown in Figure 15.

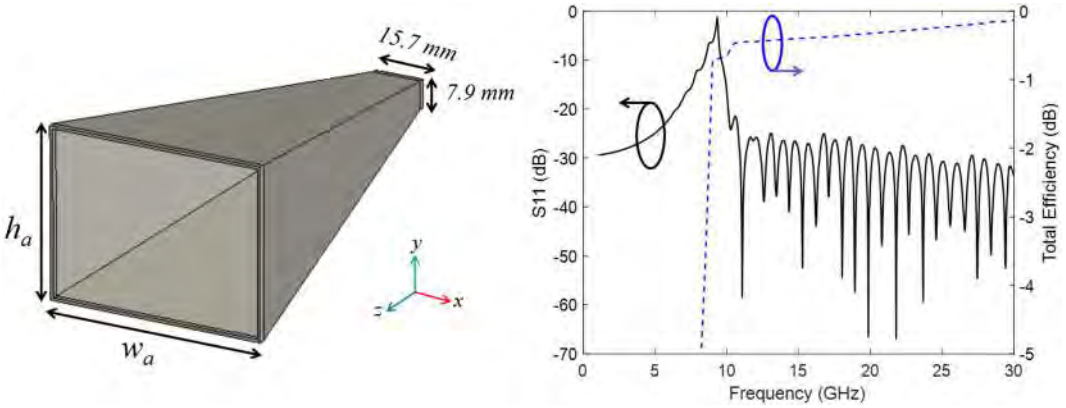


FIGURE 15 Standard horn antenna under use and its measured *S* parameter and total efficiency

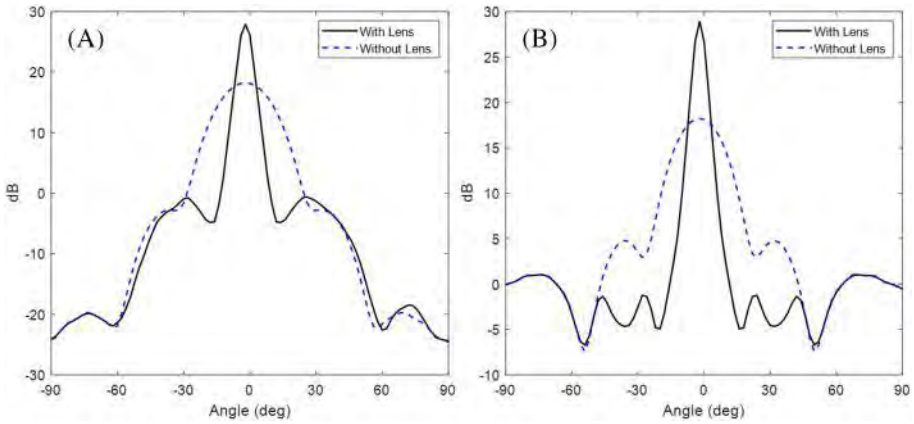


FIGURE 16 Measured radiation pattern for horn antenna with and without the proposed 2D metasurface lens at 15 GHz in (A) azimuth and (B) elevation planes

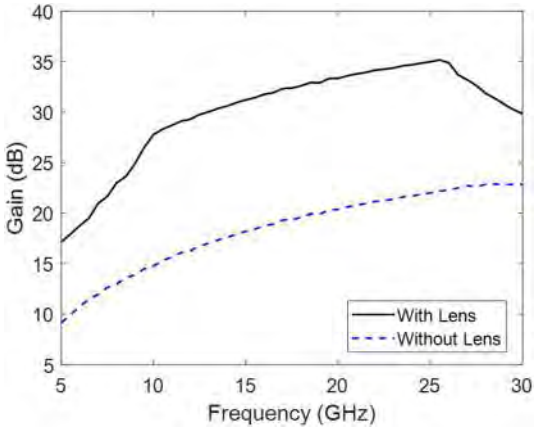


FIGURE 17 Measured gain with and without lens for the horn antenna system

The horn antenna has an average gain of 20 dBi from 12 to 18 GHz. The radiation pattern of the horn antenna used with and without the proposed lens system is shown in Figure 16. The antenna presents a gain improvement of 13.4 dBi and an HPBW reduction of 25° to 7° in both azimuth and elevation planes. The same improvement in the lens is seen consistent throughout the frequency band, demonstrating a gain improvement of 13 ± 0.4 dBi with no significant change in the total efficiency of the antenna system for the bandwidth of 10–25 GHz (see Figure 17). As the MTM is

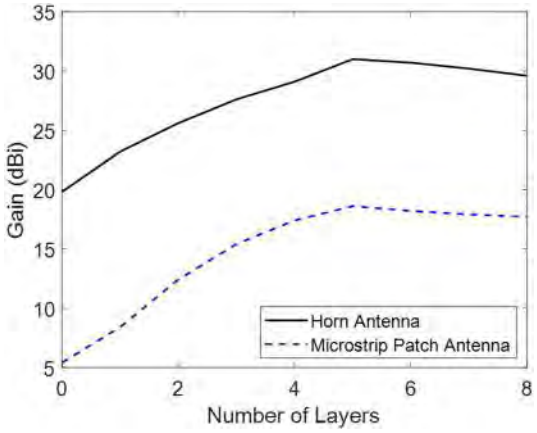


FIGURE 18 Gain enhancement in horn and microstrip patch antenna with the number of layers of the proposed metasurface lens at 15 GHz

TABLE 1 Comparing proposed lens with different lens in the literature

Ref.	Lens description	AUT	Bandwidth	Avg. gain improvement
[3]	Huygens' metasurface lens structure	Horn antenna	23–29 GHz	10.2 dBi
[4]	3D printed dielectric lens with increasing thickness	Waveguide feed	8–12 GHz	9 dBi
[5]	3D printed convex aperture dielectric lens	Horn antenna	25–31 GHz	12 dBi
[13]	Double-ring resonator-based metasurface lens	Vivaldi antenna	5–7 GHz	5.7 dBi
[17]	Maxwell fish-eye lens based on GRIN metamaterial	Microstrip antenna	3–7.5 GHz	11 dBi
This work	ENZ metamaterial-based GRIN Luneberg lens	Microstrip and horn antennas	10–18 GHz	13.4 dBi

designed to follow the gradual increase in permittivity near zero in the selected band. Outside the specified region, the gain improvement starts to decrease significantly and is shown in Figure 17. The antenna system with the lens provides a good SLL of >25 dBi in both azimuth and elevation planes. Further, understanding that the lens is comprised of different layers, it is important to understand the improvement in gain of the antenna with the change in number of layers. The same has been studied and summarised in Section 4.

4 | LENS PARAMETRIC ANALYSIS

The proposed multilayered lens antenna is constructed with five layers of the metasurface with varying relative permittivity. Common sense dictates that the gain can be increased even further, and a much sharper pencil beam can be produced using more layers of the lens. However, this is shown to be false, and the gain tends to be more constant as we increase the layers further. The horn antenna described in Section 3 was tested with the proposed 2D metasurface lens with increasing number of layers. The lens demonstrated an average gain enhancement of 4.7 to 13.4 dBi throughout the frequency band as we go from single-layered lens to five-layered lens and is then found to be reducing with further introduction of layers. This has been presented in Figure 18.

This is believed to be because of stronger reflections from the lens as its thickness increases further, as well as increase in the dielectric losses with introduction of further substrate layers. As seen in Figures 9 and 14, the wavefront is converted from spherical to planar wavefront, making the beam extremely directional. Furthermore, general EM wave theory[29,30] dictates that the change in material introduces a change in the direction of beam and introduces reflections. These reflections start to increase as the lens gets thicker, and hence, as we go above five layers, the gain starts to

decrease slowly. The reduction in gain is low because of very low reflections and losses but will keep on decreasing with further introduction of layers. The same behaviour has been noticed for the rest of frequency band as well. However, it is believed that the number of layers in the lens will vary based on the beam pattern. The wider the beam, the more number of layers will be needed, and the narrow the incident beam, the lesser will be the number of layers being used for the metasurface lens. The performance comparison of the proposed lens with the relevant works is tabulated in Table 1.

5 | CONCLUSION

The paper presented a combination of two commonly used MTMs (ring resonator and microstrip dipole line), combined to create an ENZ nature for a selected frequency band. The proposed ENZ MTM is further demonstrated to have to decrease relative permittivity with change in its dimensions, which is shown to be re-arranged to create an array of permittivity varying metasurface, acting as a virtual EM convex lens. The lens is shown to have an almost constant gain enhancement of 13.4 dBi throughout the operating bandwidth and is shown to be tested with a microstrip patch antenna and a horn antenna. Further, the lens has been studied for its thickness with the introduction of further layers, and a maximum number of layers based on the beam pattern have been defined to achieve optimal performance. The lens is further compared with other works in the literature in Table 1 to further appreciate the novelty of the work. The lens can be used for several applications such as satellite communication, point-to-point communication, vehicular radar and so forth to produce extremely directive beams.

DATA AVAILABILITY STATEMENT

All data generated or analysed during this study are included in this article.

ORCID

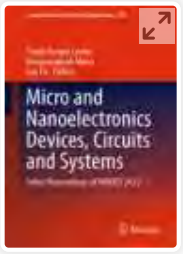
Sri Pravallika Narjala https://orcid.org/0000-0003-2126-9802

REFERENCES

1. Konstantinidis K, Feresidis AP, Constantinou CC, et al. Low-THz dielectric lens antenna with integrated waveguide feed. *IEEE Trans THz Sci Technol*. 2017;7(5):572-581. doi:10.1109/TTHZ.2017.2725487
2. Yurduseven O, Cavallo D, Neto A. Wideband dielectric lens antenna with stable radiation patterns fed by coherent array of connected leaky slots. *IEEE Trans Antennas Propag*. 2014;62(4):1895-1902. doi:10.1109/TAP.2014.2298875
3. Xue C, Lou Q, Chen ZN. Broadband double-layered Huygens' metasurface lens antenna for 5G millimeter-wave systems. *IEEE Trans Antennas Propag*. 2020;68(3):1468-1477. doi:10.1109/TAP.2019.2943440
4. Mahouti P, Belen MA, Güneş F, Yurt R. Design and realization of multilayered cylindrical dielectric lens antenna using 3D printing technology. *Microw Opt Technol Lett*. 2019;61(5):1400-1403. doi:10.1002/mop.31720
5. Ansarudin F, Abd Rahman T, Yamada Y, Rahman NHA, Kamardin K. Multi beam dielectric lens antenna for 5G base station. *Sensors*. 2020;20(20):5849. doi:10.3390/s20205849
6. Gupta R, Bakshi G, Bansal A. Dual-band circularly polarized stacked sapphire and TMM13i rectangular DRA. *Prog Electromagn Res M*. 2020;91:143-153. doi:10.2528/PIERM20012701
7. Wang KX, Wong H. A circularly polarized antenna by using rotated-stair dielectric resonator. *IEEE Antennas Wirel Propag Lett*. 2015;14:787-790. doi:10.1109/LAWP.2014.2385475
8. Dussopt L, Boriskin A, Sauleau R. *Aperture Antennas for Millimeter and Sub-Millimeter Wave Applications*. Springer; 2017.
9. Niroo Jazi M, Chaharmir MR, Shaker J, Sebak AR. Broadband transmitarray antenna design using polarization-insensitive frequency selective surfaces. *IEEE Trans Antennas Propag*. 2016;64(1):99-108. doi:10.1109/TAP.2015.2500230
10. Hsu C-Y, Hwang L-T, Horng T-S, Wang S-M, Chang F-S, Dorny CN. Transmitarray design with enhanced aperture efficiency using small frequency selective surface cells and discrete Jones matrix analysis. *IEEE Trans Antennas Propag*. 2018;66(8):3983-3994. doi:10.1109/TAP.2018.2839755
11. Abdelrahman AH, Elsherbeni AZ, Yang F. High-gain and broadband transmitarray antenna using triple-layer spiral dipole elements. *IEEE Antennas Wireless Propag Lett*. 2014;13:1288-1291. doi:10.1109/LAWP.2014.2334663
12. Reis JR, Copner N, Hammoudeh A, et al. FSS-inspired transmitarray for two-dimensional antenna beamsteering. *IEEE Trans Antennas Propag*. 2016;64(6):2197-2206. doi:10.1109/TAP.2016.2543802
13. Chen Y, Chen L, Yu J-F, Shi X-W. A C-band flat lens antenna with double-ring slot elements. *IEEE Antennas Wireless Propag Lett*. 2013;12:341-344. doi:10.1109/LAWP.2013.2247973
14. Abdelrahman AH, Nayeri P, Elsherbeni AZ, Yang F. Single-feed quad-beam transmitarray antenna design. *IEEE Trans Antennas Propag*. 2016;64(3):953-959. doi:10.1109/TAP.2016.2517660

15. Abdelrahman AH, Nayeri P, Elsherbeni AZ, Yang F. Bandwidth improvement methods of transmitarray antennas. *IEEE Trans Antennas Propag*. 2015;63(7):2946-2954. doi:10.1109/TAP.2015.2423706
16. Yu J, Chen L, Yang J, Shi X. Design of a transmitarray using split diagonal cross elements with limited phase range. *IEEE Antennas Wireless Propag Lett*. 2016;15:1514-1517. doi:10.1109/LAWP.2016.2517019
17. Xu HX, Wang GM, Tao Z, Cai T. An octave-bandwidth half Maxwell fish-eye lens antenna using three-dimensional gradient-index fractal metamaterials. *IEEE Trans Antennas Propag*. 2014;62(9):4823-4828. doi:10.1109/TAP.2014.2330615
18. Ma HF, Cai BG, Zhang TX, Yang Y, Jiang WX, Cui TJ. Three-dimensional gradient-index materials and their applications in microwave lens antennas. *IEEE Trans Antennas Propag*. 2013;61(5):2561-2569. doi:10.1109/TAP.2012.2237534
19. Dhouibi A, Burokur SN, Lustrac A, Priou A. Compact metamaterial-based substrate-integrated Luneburg lens antenna. *IEEE Antennas Wireless Propag Lett*. 2012;11:1504-1507. doi:10.1109/LAWP.2012.2233191
20. Bosiljevac M, Casaletti M, Caminita F, Sipus Z, Maci S. Non-uniform metasurface Luneburg lens antenna design. *IEEE Trans Antenna Propag*. 2012;60(9):4065-4073. doi:10.1109/TAP.2012.2207047
21. Singh AK, Abegaonkar MP, Koul SK. Compact near zero index metasurface lens with high aperture efficiency for antenna radiation characteristic enhancement. *IET Microw Antennas Propag*. 2019;13(8):1248-1254.
22. Li H-P, Wang G-M, Cai T, Liang J-G, Gao X-J. Phase- and amplitude-control metasurfaces for antenna main-lobe and sidelobe manipulations. *IEEE Trans Antennas Propag*. 2018;66(10):5121-5129. doi:10.1109/TAP.2018.2858181
23. Li M, Li QL, Wang B, Zhou CF, Cheung SW. A low-profile dual-polarized dipole antenna using wideband AMC reflector. *IEEE Trans Antennas Propag*. 2018;66(5):2610-2615. doi:10.1109/TAP.2018.2806424
24. Li M, Li Q, Wang B, Zhou C, Cheung S. A miniaturized dual-band base station array antenna using band notch dipole antenna elements and AMC reflectors. *IEEE Trans Antennas Propag*. 2018;66(6):3189-3194. doi:10.1109/TAP.2018.2819890
25. Measurement of dielectric material properties. Rohde & Schwarz. May 2012. https://www.rohde-schwarz.com/file/RAC-0607-0019_1_5E.pdf
26. Giddens H, Hao Y. Multibeam graded dielectric lens antenna from multimaterial 3-D printing. *IEEE Trans Antennas Propagat*. 2020;68(9):6832-6837. doi:10.1109/TAP.2020.2978949
27. Isakov D, Stevens CJ, Castles F, Grant PS. 3D-printed high dielectric contrast gradient index flat lens for a directive antenna with reduced dimensions. *Adv Mater Technol*. 2016;1(6):1600072. doi:10.1002/admt.201600072
28. Bansal A, Gupta R. A review on microstrip patch antenna and feeding techniques. *Int J Inf Tecnol*. 2020;12(1):149-154. doi:10.1007/s41870-018-0121-4
29. Balanis CA. *Antenna Theory: Analysis and Design*. John Wiley & Sons; 2015.
30. Pozar DM. *Microwave Engineering*. John Wiley & Sons; 2011.

How to cite this article: Narjala SP, V R A, K R. High-gain broadband ENZ metasurface-based RF lens with gradient refractive index for microwave and millimetre-wave applications. *Int J Commun Syst*. 2022;35(17):e5325. doi:10.1002/dac.5325



Micro and Nanoelectronics Devices, Circuits and Systems pp 359–366

Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection

M. Lakshmi Prasanna & V. R. Anitha

Conference paper | First Online: 12 September 2022

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 904)

Abstract

Many studies have shown that most MEMS devices use light beams whose deflection response is the main focus before choosing it for specific applications. This paper deals with the simulation of the MEMS-based cantilever beam for low-pressure detection. COMSOL 5.4 multiphysics FEM model is used to study the behavior of beam and compared

with the conventional cantilever beam. Both are analyzed with two different materials SiO_2 and SU-8. The change in the deflection has been analyzed with various pressure values applied on the cantilever beam. Through simulation and analysis results, it is observed that the deflection of the proposed beam is good at lower pressure values.

Keywords

Microcantilever **Deflection**

Rectangular beam **FEM**

References

1. Guo K, Jiang B, Liu B, Xingeng Li, Wu Y, Tian S, Gao Z, Zong L, Yao S, Zhao M, Mi C, Zhu G (2021) Study on the progress of piezoelectric microcantilever beam micromass sensor. IOP Conf Ser Earth Environ Sci 651:022091. https://doi.org/10.1088/1755-1315/651/2/022091

2. Ali S, Bhuvaneswari H, Kumar B (2020) Design and modeling of a large deflection microcantilever using rectangular SCR. AIP Conf Proc 2281:020037. https://doi.org/10.1063/5.0027942

3. Rotake D, Darji A (2020) Design and reliability testing of microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.13140/RG.2.2.25052.92803

4. Ashok A, Nighot R, Sahu N, Pal P, Pandey A (2019) Design and analysis of microcantilever beams based on arrow shape. Microsyst Technol 25:4379–4390. https://doi.org/10.1007/s00542-019-04555-4

5. Passian A, Thundat T, Thanihaichelvan M (2017) Microcantilever sensors. https://doi.org/10.1016/B978-0-12-803581-8.10525-9

6. Rotake D, Darji A, Singh J (2020) Ultrasensitive multi-arm-microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.1109/VDAT50263.2020.9190249

7. Gharge B, Upadhye V, Bodas D (2015) Design and simulation of microcantilevers for detection of pathogens. https://doi.org/10.1109/ISPTS.2015.7220122

8. Parsediya D (2016) Deflection and stresses of effective micro-cantilever beam designs under low mass loading:111–114. https://doi.org/10.1109/ICEPES.2016.7915915

9. Mathew R, Sankar R (2018) A review on surface stress-based miniaturized piezoresistive SU-8 polymeric cantilever sensors. Nano-Micro Lett 10:1–41. https://doi.org/10.1007/s40820-018-0189-1

10. Gopinath PG (2014) Design and characterization of various shapes of microcantilever for human immunodeficiency virus detection. Int J Res Eng Technol 03:24–29. https://doi.org/10.15623/ijret.2014.0310004

11. Schmid S (2009) Electrostatically actuated all-polymer microbeam resonators—characterization and application

12. Lamba M, Singh K, Chaudhary H (2020) Design analysis of polysilicon piezoresistors PDMS (polydimethylsiloxane) microcantilever based MEMS force sensor. Int J Mod Phys B 34:2050072.https://doi.org/10.1142/S0217979220500721

Acknowledgements

A Special thanks to the National MEMS Design Centre, Sree Vidyanikethan Engineering College, A. Rangampet, Tirupati, for providing the research facility.

Author information

Authors and Affiliations

Department of Electronics and Communication Engineering, JNTUA, Ananthapuramu, Andhra Pradesh, India

M. Lakshmi Prasanna

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Yelahanka, Bangalore, Karnataka, 560064, India

V. R. Anitha

Corresponding author

Correspondence to M. Lakshmi Prasanna .

Editor information

Editors and Affiliations

Electronics and Communication Engineering, National Institute Of Technology Silchar, Silchar, India

Trupti Ranjan Lenka

Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA

Durgamadhab Misra

Electronic Materials Engineering, Australian National University, Canberra, ACT, Australia

Lan Fu

Rights and permissions

Copyright information

About this paper

Cite this paper

Lakshmi Prasanna, M., Anitha, V.R. (2023). Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection. In: Lenka, T.R., Misra, D., Fu, L. (eds) Micro and Nanoelectronics Devices, Circuits and Systems.

Lecture Notes in Electrical Engineering, vol 904. Springer, Singapore. https://doi.org/10.1007/978-981-19-2308-1_37

.RIS .ENW .BIB

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-19-2308-1_37	12 September 2022	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-2307-4	978-981-19-2308-1	Engineering Engineering (R0)

DESIGN AND DEVELOPMENT OF LIGHT WEIGHT FRACTAL TREE ANTENNA STRUCTURE

V.R. Anitha[1,*], Shilpa Hiremath[2], G. Shyni[3], J.S. Binoj[4], Shukur Abu Hassan[5], R.A. Ilyas[6]

[1,2]*Department of Electronics and Communication Engineering, BMS Institute of Technology & Management, Yelahanka-560064, Bengaluru, India.*
[3]*Department of Electronics and Communication Engineering, University College of Engineering Nagercoil, Konam - 629004, Tamil Nadu, India.*
[4]*Institute of Mechanical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai - 602105, Tamilnadu, India.*
[5]*Centre for Advanced Composite Materials, Office of Deputy Vice-Chancellor (Research & Innovation), UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*
[5]*School of Mechanical Engineering, Faculty of Engineering, UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*
[6]*School of Chemical and Energy Engineering, Faculty of Engineering, UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*

ABSTRACT

The advancement in wireless communication system and increasing demand to make small size and multiband antenna applications has a great demand in both commercial as well military applications. To fulfill this demand, the reconfigurable fractal antenna has been introduced. The utilization of fractal geometries has a great impact in many areas of science and engineering. The development of an antenna by utilizing a fractal idea will reduce an antenna size without reducing its performance.This paper presents the new shape of fractal tree antenna for multiband applications by reconfigure the shape of fractal tree antenna using PIN diode switches. The proposed antenna covers the service bands such as WLAN (4.9-5.8) GHz and Wi-MAX (3.2- 3.8) GHz during its operation and also operates some other frequencies between 3.12 GHz to 5.4 GHz. The designed antenna is reconfigured by using PIN diodes to connect consecutive branches. The proper switching positions of the PIN diode alter the performance of the antenna at different resonant frequencies. Simulated and measured results show the characteristics of the designed antenna.

Keywords: Fractaltree, Micro strip patch antenna, PIN diodes, Reconfigurability, WLAN, Wi-MAX

INTRODUCTION

In radar and modern communication systems the demand on multi-functional antennas is increasing. The requirements for these antennas are the abilities to have multi radiation patterns, adapting the operation frequency and polarization, keeping the physical dimensions and positioning unaltered. Reconfigurable antennas with switching capability used as a multiple input multiple output (MIMO) system have been used in recent years to full these requirements. By means of switches with compatible antenna elements the antenna and its feed structure can be physically reconfigured to provide radiation pattern, frequency band and polarization diversity so they have more advantage to compare with conventional antennas [1]. The most prevalent implementation about reconfigurable antenna is related to the operation frequency [2] since it might be the easiest feature to alter. Polarization and pattern reconfigurable antennas are also attractive since they can provide diversity features which leads to an increased signal to noise ratio and therefore a higher quality of service of whole systems [3][4][5]. PIN diodes are generally used more than transistors and switches as switching devices for RF and microwave front-end communication systems since they have several crucial properties such as low insertion loss, goodisolation, low power handling and low cost [6].

Although a reconfigurable antenna can take many shapes we will focus on fractal tree antennas in this work. In terminology, fractal means broken or irregular fragments which were originally entitled by Mandelbrot [7] to describe a family of complex shapes that possess an inherent self-similarity in their geometric structure. As a result of small investigation in the environment a lot of example for fractal shapes can be seen as trees, clouds, galaxies, leaves, and much more. Fractal tree structures can be applied into antenna design to produce multiband characteristics [8-10]. A conformal antenna is an antenna that conforms to something or it conforms to prescribed shape. The shape can be some part of a train, airplane or other vehicle. The purpose in conformal antenna is to make surface matched structure so that it becomes integrated with the structure and does not extra drag. Since they have very low profile and can be applied on exile substrates they can behave\hidden" antennas [11]. The target in this paper is showing the design of a new shape of fractal tree antenna for multiband and multi radiation pattern applications. PIN diodes are used as switches for multi frequency and multi beam reconfiguration. By switching PIN diodes, the resonant frequency and radiation pattern variation simulation by using HFSS is presented. The antenna geometry, Fractal Tree, is a plane fractal constructed from Rectangle and Triangle microstrip patch which is connected by using wire for connectivity as shown in Fig. 1. The Parent branch is one pair of rectangle and triangle patches. The child branches are splited from parent branch by 0.5 scale factor branch angle $\theta = 60°$. A 4-iternation, tree is applied as the radiation part here [12]. Length of main stem L1=22mm, width of the stem W1=10mm,

Article history:
Received: 1 July 2022
Accepted: 31 July 2022
Published: 22 August 2022

E-mail addresses:
anithavr@bmsit.in
*Corresponding Author

ISSN: 978-629-96647-0-3

Substrate height h=3.5mm, permittivity of the substrate is 3 Rogers RT/Duroid 5880 (tm) and resonant frequency 3.8 GHZ length and width of each iteration is given in Table 1.

TABLE 1: Dimensions of patch for different iterations.

Iteration Number	Length in mm	Width in mm
First	L1 = 22	W1 = 10
Second	L2 = 11	W2 = 5
Third	L3 = 5.5	W3 = 2.5
Fourth	L4 = 2.75	W4 = 1.25

The geometry shows multi-band and wide bandwidth behavior compared with the existing square patch of same dimensions.

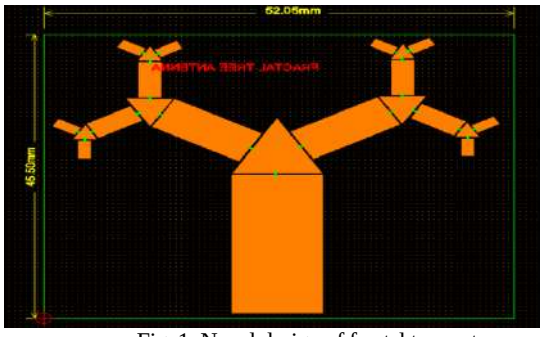
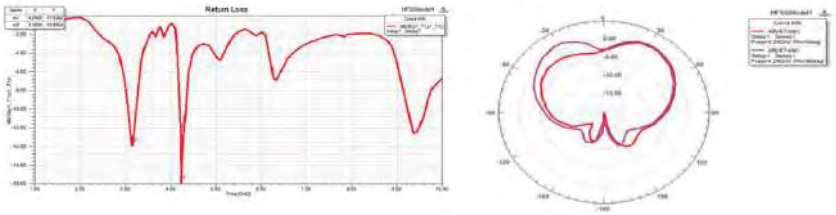


Fig. 1: Novel design of fractal tree antenna

RESULTS AND DISCUSSION

The proposed reconfigurable fractal tree antenna is designed and simulated by using HFSS and tested. By changing the diode conditions, alters the characteristic of antenna and it results shifting in operating frequency and radiation pattern. By different combinations of diode switching positions, many different patterns with different resonance frequencies. When all 21 diodes are in ON state, antenna starts to operate at 2.42 GHz, 2.98 GHz and 3.36 GHz and return loss, radiation pattern and gain are shown. In Fig. 2, all diodes are in OFF state, antenna operates well at 3.68 GHz, 2.44 GHz. The maximum peak gain and bandwidths for different combinations of PIN diodes ON and OFF states are shown in Table 2 and 3.

To validate the result of the automated antenna design, a prototype of the designed antenna has been fabricated and tested. From Fig. 4, the prototype antenna is on RT Duroid substrate, and fed by a 50 Ω coaxial cable. The return loss of the fabricated antenna was measured using Network Analyzer. The measured return loss and VSWR of the designed antenna is shown in Fig. 5. The measured results were met the simulated result.



(a) (b)

Fig. 2: Fractal Tree Antenna with Wire (a) Return Loss (b) Radiation Pattern

Fig. 3: Prototype of Reconfigurable Fractal Tree Antenna

(a) (b)

Fig. 4: Reconfigurable fractal tree antenna (a) Measured Return Loss (b) Measured VSWR

ISSN: 978-629-96647-0-3

TABLE 2: Compilation of Results in ON state

ON State PIN Diodes	Resonant Frequency (GHz)	Max. Peak Gain (dB)	Bandwidth (MHz)
ALL	3.12, 3.8, 6.12	7.428	80
Left side ON	3.2, 3.79	6.4719	250
Right side ON	3.2, 3.8	6.2615	160

TABLE 3: Compilation of results in OFF state

OFF State PIN Diodes	Resonant Frequency (GHz)	Max. Peak Gain (dB)	Bandwidth (MHz)
ALL	3.12, 5.275, 9.6	6.6128	120

CONCLUSIONS

In this paper a multiband frequency reconfiguration is studied by using fractal tree antenna which is controlled by PIN diodes and their equivalent circuit. The effects of the biasing lumped elements on the antenna performance are discussed based on the simulation and measurement results by using HFSS Simulation tool. PIN diodes reconfiguration causes frequency shift in the resonance frequency of the antenna and radiation pattern can be controlled by changing PIN diodes state like ON or OFF states of the respective PIN diodes which is deployed in this antenna.

Different performance parameters of an antenna such as s-parameter or return loss, far-field radiation pattern have been perceived and analyzed by taking into account the simulation and measurement results. The proposed antenna is simulated over the frequency range of 1-10 GHz along with VSWR < 2. The result reveals that the antenna designed resonates at 2.4 GHz, 3.8 GHz, and 5.2 GHz frequency that provides with a minimum return loss of greater than -10 dB which is a good result. A good impedance bandwidth of the designed antenna is achieved for different cases. This shows that the use of reconfigurable fractal geometry shows broadband operations without employing any further modifications for different wireless applications such as Wi-fi, WiMaX, and different satellite mobile communications.

REFERENCES

[1] S. Mili, C. Larbi Aguili, and T. Aguili, “Study of Fractal-Shaped Structures with Pin Diodes using The Multi-Scale Method Combined to The Generalized Equivalent Circuit Modeling” *Progress In Electromagnetics Research B*, Vol. 27, pp. 213-233, 2011, doi: 10.2528/PIERB10110105.

[2] G. H. Huff, J. Feng, S. Zhang, G. Cung, and J. T. Bernhard, “Directional reconfigurable antennas on laptop computers Simulation, measurement and evaluation of candidate integration positions,” *IEEE Trans. Antennas Propagation*, vol. 52, no.12, pp. 3220–3227, Dec. 2004, doi: 10.1109/TAP.2004.836425.

[3] R. Kumar and P. N. Chaubey,”On the Design of Tree-type Ultra WidebandFractal Antenna for DS-CDMA System” Journal of Microwaves, Optoelectronics and Electromagnetic Applications, vol.11, no.1, June 2012, doi: 10.1590/S2179-10742012000100009.

[4] H. Altun, E. Korkmaz, B.Türetken,”Reconfigurable Fractal Tree Antenna for Multiband Applications”, In *2011 XXXth URSI General Assembly and Scientific Symposium*, pp. 1-4. IEEE, 2011, doi: 10.1109/URSIGASS.2011.6050525.

[5] S. V. Khobragade, V. R. Anitha, “Design and Simulation of Fractal Tree Antenna for Wireless Application” *International Journal of Electronics and Communication Engineering & Technology (IJECET)*, Vol. 3, Issue 1, January- June, 2012, doi: 10.1.1.474.9820&rep=rep1&type=pdf.

[6] G. I. Kiani, K. P. Esselle, A. R. Weily, and K. L. Ford, “Active frequency selective surface using PIN diodes," *International Symposium on Antennas and Propagation Society*, pp. 4525-4528, IEEE, 2007, doi: 10.1109/APS.2007.4396549.

[7] D. Langoni, M. H. Weatherspoon, E. Ogunti, and S. Y. Foo, “An overview of recon gurable antennas: Design, simulation, and optimization,” *IEEE 10th Annual Wireless and Microwave Technology Conference 2009*, WAMICON '09, pp. 1-5, 2009.

[8] J. Zhang, A. Wang, and P. Wang, “A survey on reconfigurable antennas,” In *2008 International Conference on Microwave and Millimeter Wave Technology*, Vol. 3, pp. 1156-1159, IEEE 2008, doi: 10.1109/ICMMT.2008.4540632.

[9] C. B. Dietrich, K. Dietze, J. R. Nealy, and W. L. Stutzman, “Spatial, polarization, and pattern diversity for wireless handheld terminals,” *IEEE Trans. Antennas Propag*., Vol. 49, No. 9, pp. 1271-1281, Sept. 2001, doi:10.1109/8.947018.

[10] S. Nikolaou, B. Kim, and P. Vryonides, “Reconfiguring antenna characteristics using PIN diodes,” *3rd European Conference on Antennas and Propagation 2009, EuCAP* 2009, pp. 3748-3752, IEEE, 2009.

[11] J. S. Petko and D. H. Werner, “Miniature reconfigurable threedimensional fractal tree antennas,” *IEEE Transactions on Antennas and Propagation*, Vol 52, No. 8, pp. 1945-1956,2004, doi: 10.1109/TAP.2004.832491.

ISSN: 978-629-96647-0-3

PERFORMANCE ANALYSIS OF 6X6 CNT FIELD-EFFECT TRANSISTOR ARRAY MODEL FOR BIOMEDICAL APPLICATIONS

V.R. Anitha[1,*], Gopinath P.G[2], G. Shyni[3], J.S. Binoj[4], Shukur Abu Hassan[5], R.A. Ilyas[6]

[1]*Department of Electronics and Communication Engineering, BMS Institute of Technology & Management, Yelahanka-560064, Bengaluru, India.*
[2] Department of ECE, Siddharrth Institute of Engineering & Technology, Puttur, Andhra Pradesh, India.
[3]*Department of Electronics and Communication Engineering, University College of Engineering Nagercoil, Konam - 629004, Tamil Nadu, India.*
[4]*Institute of Mechanical Engineering, Saveetha School of Engineering, Saveetha Institute of Medical and Technical Sciences (SIMATS), Chennai - 602105, Tamilnadu, India.*
[5]*Centre for Advanced Composite Materials, Office of Deputy Vice-Chancellor (Research & Innovation), UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*
[5]*School of Mechanical Engineering, Faculty of Engineering, UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*
[6]*School of Chemical and Energy Engineering, Faculty of Engineering, UniversitiTeknologi Malaysia, Johor Bahru, 81310, Johor, Malaysia.*

ABSTRACT

In this paper, different types of the latest Field Effect Transistor (FET) and its electrical characteristics are studied and simulations are performed. Based on the different cases and studies, identified the suitable FET, which is to use in Bio-Medical Applications for sensor design. Nowadays, detection of disease-causing antigen is a big challenge, to overcome the difficulty, FET based biosensor is introduced to improve the sensitivity concerning the disease causing antigen at micro/nano level. Different types of transistor with different layers of material are introduced to identify the suitable transistor. All the simulations are performed in NANOHUB.ORG (Online Tool) and performance is shown as simulation results, graphs and tables.

*Keywords:*Performance Analysis, Field Effect Transistors, Biomedical Applications

INTRODUCTION

The electrical characteristics of Silicon-On-Insulator (SOI) FET-based devices were compared. The ISFET gates and contacts are defined by the wet etching method and passivation and patterned to the gate areas. Gate sensing layers fabricated is10nm Si_3N_4 over the SiO_2. They show the variation of Ids with time for DU145tumor cells [1].The unique properties of single-wall CNTs are due to the confinement of electrons. It has a high value of current densities in the order of 109 A/cm^2. He explained the current relationship of CNTFET to analyze the electrical properties and given output electrical characteristics of 300 nm long CNTFET and transfer characteristics of 800 nm long CNTFET under room temperature. As it has more excellent operating characteristics than silicon devices, it can be used as a biosensor as it is very sensitive to the change in the gate region [2].The nanowire-based field-effect sensor devices to detect the concentration of chemical analyte present in the solution were demonstrated. This sensor device converts chemical action to biological findings through an electronic and digital signal. They pointed out ta hat single virus particle can bind in the Silicon nanowire with the help of antibody coating over it [3]. Silicon Nano Wires and CNTs are very robust materials because of their Young's Modulus and tensile strength values. The advantages and disadvantages of various fabrication, functionalization and assembling procedures of the Nanosensors are reviewed in detail to understand the properties and operations of nano biosensors [4]. An experimental set-up for detecting Salmonella Infantis with a network of CNTFETs functionalized with anti-Salmonella antibodies is explained. Antigen and antibody interaction of these analytes with an SWCNT functionalized with anti-Salmonella antibodies and protected with Tween 20. Anti-Salmonella antibodies were adsorbed on the carbon nanotubes [5]. The detection of biomolecules using CNTFET is explained with the noise suppression methods in CNTFETs by altering the current measurement. CNTFET exhibit high performance as a sensor and in electrical properties and they concluded that CNTFET-based biosensor is promising technology for future biosensors [6]. Aptamer-based immune sensors for the detection of Immunoglobulin E (IgE) using CNTFETs shown an optical and SEM images of arrayed CNTFETs for label-free biosensors. Bio-molecules are detected by the CNT channel conductance modulation in the aptamer-modified CNTFETs [7]. An array of Carbon Nanotube Field Effect Transistors based biosensor to provide nano-biosensing based on cytochromes designed. CNFET biosensor array design proposed based on CNTs, thereby achieving high level of sensitivity due to higher flow of current. The design shows that the improved specificity through multi-gate controller at the nano-scale [8]. The use of graphene films configured as FET for real time molecular sensing. Glucose molecules detected by change in conductance of the graphene transistor as the specific redox enzyme oxidizes them. Glucose in functionalized graphene FET illustrated. Current responses with glucose concentrations had shown that change of current increases with respect to the glucose concentration [9]. Graphene based biosensors exhibited good selectivity and sensitivity towards the detection of small biomolecules. Comparative study of sensing with respect to the efficiency between CNT and graphene-based biosensors are given neatly [10]. CNTFET sensor is used to detect Mycobacterium tuberculosis isolates in the biosensor chip, in which

Article history:
Received: 1 July 2022
Accepted: 31 July 2022
Published: 22 August 2022

E-mail addresses:
anithavr@bmsit.in
*Corresponding Author

ISSN: 978-629-96647-0-3

CNT channels modified with aptamers. This aptamer-based CNTFET device is a promising technology for the development of an integrated, high-specific, multiplexed real-time biosensor [11]. Physics of carbon nanotubes and the calculation of sub-band for zigzag CNT to model small band-gap illustrated and a comparison is made with existing method and proposed the chirality CNT with small band-gap which performs actively in the modeling of nanobiosensor [12].

MATERIALS AND METHODS

Properties of Silicon Nano wire MOSFET and Carbon nanotube MOSFET

As MOS infrastructure continues to scale down deeper into the nanoscale, various non-ideal characteristics, substantially different from MOSFET are aroused. Carbon Nanotube (CNT), the rolled structure of graphene, renders itself to multi-disciplinary applications in nanoelectronic circuits and removes those effects. The field effect transistors consisting of CNT are called CNTFET and recently they have attracted the attention as possible building blocks of future nanoelectronics due to their distinctive structural, electrical and mechanical properties. These provide notable characteristics over MOSFET using the semiconducting properties of CNT like high carrier mobility, large MFP (mean free path), ballistic conduction, suitable contact resistance, fast switching speed and less heat dissipation. Though CNT has high cost and less reliability, it plays a vital role in nanoscale production due to its structural properties. They are three types of terminals. Source, gate and drain. Two types of gates are top gate and bottom gate, top gate act as a sensing device. In CNTFET device antibody coating is applied at the gate region, sample consist of antibody is applied over the gate region in which antibody and antigen reaction takes place resulting in the change of conduction current. This concept is applied to detect in the diseases causing antigen and its concentration.

Carbon nanotubes offer extraordinary characteristics for nano-electronics applications. CNTFETs with semiconducting nanotubes are the most favorable structure with superior characteristics. In order to analyse the performance of the CNTFET in terms of electrical parameters to silicon nanowire FETs is considered and simulated the results in the tool available in nanohub.org. Fig. 1 show the characteristics with respect to the diameter of the conductor with drain current and the model is simulated using online tool nanohub.org. Output drain current is an electrical parameter. It is an important parameter for the device design and modeling which clearly shows that the CNTFET has better value characteristics than the Silicon nanowire FET.

RESULTS AND DISCUSSION

Design of 6x6 Biosensor Array

Carbon nanotubes have many unique properties to obtain unusual high current conduction make them as ansupreme components for the design of future electronics. The design of electronic circuits with CNTs achieves higher performance than the conventional electronic circuits. The CNT is grown through a chemical vapor deposition (CVD) process. Even if nanotubes are precisely positioned, there are different types of nanotubes such as metallic, semiconducting, single-walled, multi-walled cannot be controlled.

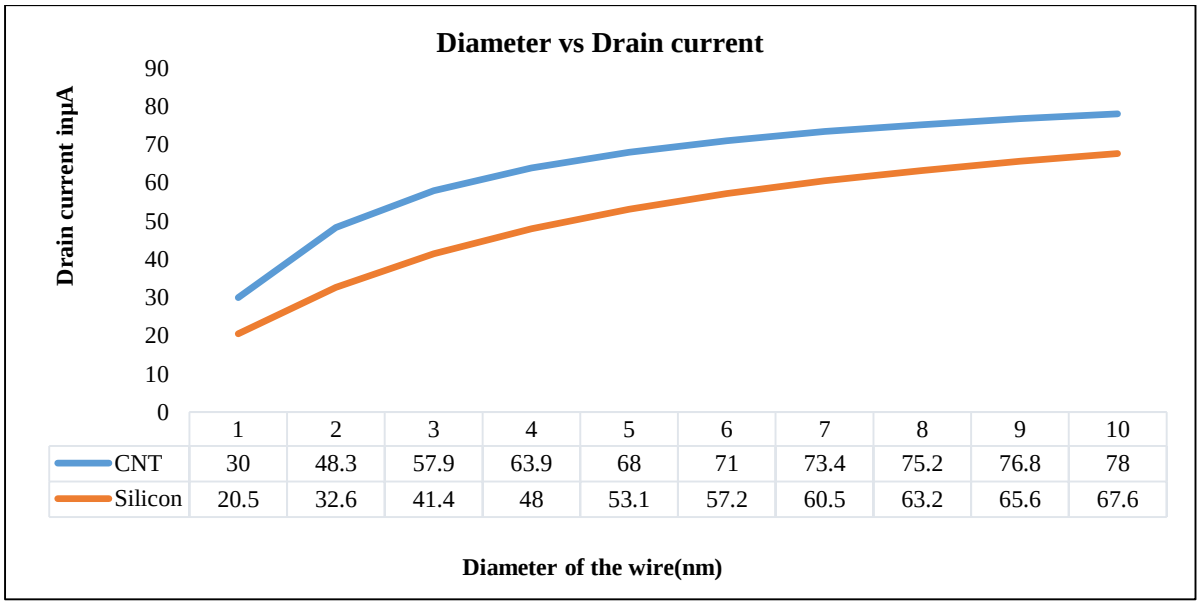


Fig. 1: This is a figure, Drain Current vs Diameter of the Wire

ISSN: 978-629-96647-0-3

Fig 2: Structure of Single Biosensor

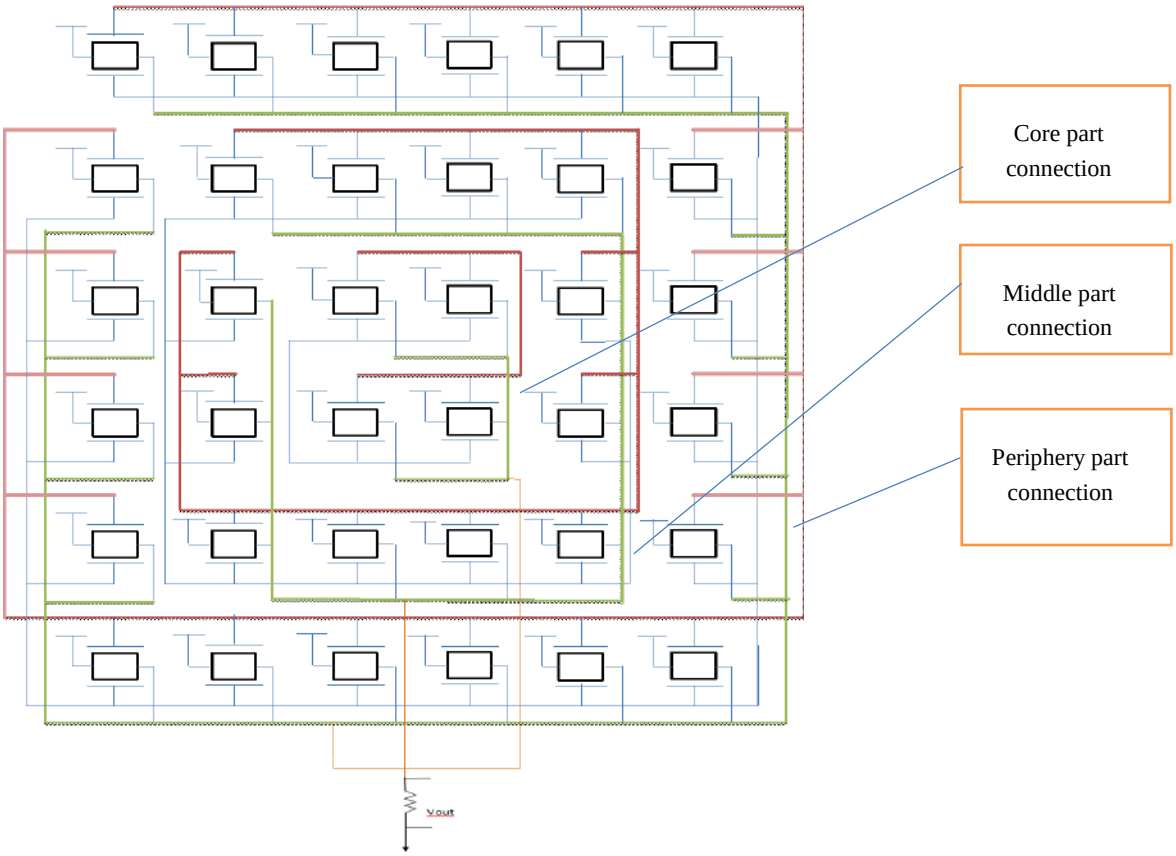


Fig. 3: Design of 6x6 array biosensor

CONCLUSIONS

The biomarker and the signal conditioning circuits are integrated into the sensor array of 6 x 6. The array consists of 36 sensors as discussed in previous section. The designed sensor array is tested for its performance for biomoleucles; as per the design, the sensor array is controlled by control signals to indicate the start of experimentation process. The reference voltages for back gate are applied from the current reference circuits. In the setup, equivalent electro potentials are applied to the top gate of CNTFET. The top gate in CNTFET fixed with the biomarkers that would reach with the antibody present in the analyte solution. However, in this work the equivalent bio-potential that could be generated with the chemical reaction of antigen with the antibody is considered and applied to the sensor array nodes. The sensor array generates three different voltages {VP, VM and VC}, which are compared with predefined voltages {VRP, VRM and VRC}.

REFERENCES

[1] A. Finn, J. Alderman, and J. Schweizer, "Towards an Optimization of FET-Based Bio-Sensors," *European Cells and Materials*, vol. 4, no. Sup 2, pp. 21-23, 2002.

[2] P. Avouris, "Molecular electronics with carbon nanotubes," *Accounts of chemical research*, vol. 35, no. 12, pp. 1026-1034, 2002, doi: 10.1021/ar010152e.

[3] F. Patolsky, G. Zheng, and C. M. Lieber, "Nanowire-based biosensors," pp. 4260-4269, 2006.

[4] A. K. Wanekaya, W. Chen, N. V. Myung, and A. Mulchandani, "Nanowire-based electrochemical biosensors," *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis*, vol. 18, no. 6, pp. 533-550, 2006, doi: 10.1002/elan.200503449.

[5] R. A. Villamizar, A. Maroto, F. X. Rius, I. Inza, and M. J. Figueras, "Fast detection of Salmonella Infantis with carbon nanotube field effect transistors," *Biosensors and Bioelectronics*, vol. 24, no. 2, 279-283, 2008, doi: 10.1016/j.bios.2008.03.046.

ISSN: 978-629-96647-0-3

[6] Y. Yamamoto, K. Maehashi, Y. Ohno, and K. Matsumoto, "Highly sensitive biosensors based on high-performance carbon nanotube field-effect transistors," *Sensors and Materials*, vol. 21, no. 7, pp. 351-361, 2009.

[7] K. Maehashi, K. Matsumoto, Y. Takamura, and E. Tamiya, "Aptamer-based label-free immunosensors using carbon nanotube field-effect transistors," *Electroanalysis: An International Journal Devoted to Fundamental and Practical Aspects of Electroanalysis*, vol. 21, no. 11, pp. 1285-1290, 2009, doi: 10.1002/elan.200804552.

[8] S. Bobba, S. Carrara, and G. De Micheli, "Design of a CNFET array for sensing and control in P450 based biochips for multiple drug detection," In Proceedings of 2010 IEEE International Symposium on Circuits and Systems, pp. 1711-1714. IEEE, 2010, 10.1109/ISCAS.2010.5537523.

[9] Y. Huang, X. Dong, Y. Shi, C. M. Li, L. J. Li, and P. Chen, "Nanoelectronic biosensors based on CVD grown grapheme," *Nanoscale*, vol. 2, no. 8, pp. 1485-1488, 2010, doi: 10.1039/C0NR00142B.

[10] T. Kuila, S. Bose, P. Khanra, A. K. Mishra, N. H. Kim, and J. H. Lee, "Recent advances in graphene-based biosensors," *Biosensors and bioelectronics*, vol. 26, no. 12, pp. 4637-4648, 2011, doi: 10.1016/j.bios.2011.05.039.

[11] H. W. Wang, J. T. Huang, and C. C. Lin, "Real-Time Detecting Concentration of Mycobacterium Tuberculosis by CNTFET Biosensor," *International Journal of Biomedical and Biological Engineering*, vol. 7, no. 7, pp. 410-412, 2013, doi: 10.5281/zenodo.1087370.

[12] S. Farhana, A. H. M. Z. Alam, and S. Khan, "Small band-gap-based CNT for modeling of nano sensor," *Procedia Computer Science*, vol. 42, pp. 122-129, 2014, doi: 10.1016/j.procs.2014.11.042.

[13] W. Kim, H. C. Choi, M. Shim, Y. Li, D. Wang, and H. Dai, "Synthesis of ultralong and high percentage of semiconducting single-walled carbon nanotubes," *Nano Letters 2*, no. 7, pp. 703-708, 2002, doi: 10.1021/nl025602q.

[14] J, Kong, H.T. Soh, A.M. Cassell, F. Calvin. Quate, and H. Dai, "Synthesis of individual single-walled carbon nanotubes on patterned silicon wafers," *Nature 395*, no. 6705, pp. 878-881, 1998, doi: 10.1038/27632.

[15] R, Martel, T. Schmidt, H. R. Shea, T. Hertel, and Ph Avouris, "Single-and multi-wall carbon nanotube field-effect transistors," *Applied physics letters* 73, no. 17, pp. 2447-2449, 1998, doi: 10.1063/1.122477.

[16] M. Freitag. "Carbon nanotube electronics and devices." In Carbon nanotubes properties and applications, CRC Press, 2006, pp. 83-118.

[17] S. J. Wind, J. Appenzeller, R. Martel, V. Derycke, and P. Avouris, "Vertical scaling of carbon nanotube field-effect transistors using top gate electrodes," *Applied Physics Letters*, vol. 80, no. 20, pp. 3817–3819, 2002, doi: doi.org/10.1063/1.1480877.

ISSN: 978-629-96647-0-3

St. XAVIER'S
CATHOLIC COLLEGE OF ENGINEERING
(Autonomous)
Chunkankadai, Nagercoil - 629003
Kanyakumari District, Tamil Nadu

AICTE - SPICES Sponsored

International Conference on

Energy Efficient Technologies for Sustainability

Certificate of Participation

This is certify that Dr./Mr./Ms./Mrs. V. R. Anitha, Professor / ECE of BMS Institute of Technology and Management has presentad a paper titled Cloud Based Face and speech Recognition for Access control Applications in the International Conference ICEETS '23 Organised by the Department of Electrical and Electronics Engineering and the Department of Mechanical Engineering held at St.Xavier's Catholic College of Engineering on 27th & 28th April 2023.

Organizing Secretary

Convener

Principal

Chapter

SRR-loaded octagonal Sierpinski-based carpet-shaped antenna for multiband application

By K. Yogaprasad (/search?contributorName=K. Yogaprasad&contributorRole=author&redirectFromPDP=true&context=ubx), M. Nanda Kumar (/search?contributorName=M. Nanda Kumar&contributorRole=author&redirectFromPDP=true&context=ubx), V. R. Anitha (/search?contributorName=V. R. Anitha&contributorRole=author&redirectFromPDP=true&context=ubx), Anil Kumar Nayak (/search?contributorName=Anil Kumar Nayak&contributorRole=author&redirectFromPDP=true&context=ubx)

Book Printed Antennas (https://www.taylorfrancis.com/books/mono/10.1201/9781003347057/printed-antennas?refId=1b0bd2b7-143c-4af5-ab97-05c1018e4150&context=ubx)

Edition	1st Edition
First Published	2022
Imprint	CRC Press
Pages	14
eBook ISBN	9781003347057

ABSTRACT

< Previous Chapter (chapters/edit/10.1201/9781003347057-5/microstrip-interconnect-design-modeling-using-reverse-approach-obtain-efficient-wideband-ms-line-rwg-hybrid-transition-varshney-sharma-roshan-kumar?context=ubx)

Next Chapter > (chapters/edit/10.1201/9781003347057-7/antenna-rfid-smart-systems-shailendra-shastri-archana-deshpande?context=ubx)

Connect with us


(https://www.linkedin.com/company/taylor-&-francis-group/)


(https://twitter.com/tandfnewsroom?lang=en)


(https://www.facebook.com/TaylorandFrancisGroup/)


(https://www.youtube.com/user/TaylorandFra

Registered in England & Wales No. 3099067
5 Howick Place | London | SW1P 1WG

Emerging Research in Computing, Information, Communication and Applications pp 263–273

Conference paper

Analysis of Deep Learning Architecture-Based Classifier for the Cervical Cancer Classification

R. Chandraprabha & Seema Singh

Conference paper | First Online: 13 December 2022

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 928)

Abstract

The incidence and mortality rates of cervical cancer in low-income states are more due to the constrained healthcare resources. For quick screening and early diagnosis, a very efficient and accurate intelligent system has to be developed which can diminish the incidence and mortal rate. Various image processing

with machine learning algorithms and deep learning algorithms are applied for the diagnosis of cervical cancer to provide technical support for the healthcare pathologist department. Deep learning architecture achieves the classification task at higher accuracy. This paper brings the analysis of a fine-tuned deep learning architecture applied for the cervical cancer diagnosis. The proposed pre-trained model Alexnet and GoogleNet are fine-tuned with transfer learning for the classification task. The dataset considered is the HErlev Dataset (standard) and real-time clinical data comprising of 917 images and 10 images, respectively. The proposed model is trained, validated, and tested with 06 epochs and 30 epochs for the standard dataset. Accuracy, error rate, precision, and area under the curve are evaluated to analyze the proposed architectures. An accuracy achieved for the proposed architectures with 30 epochs is 93.99 and 89.07%. The proposed architectures trained with 30 epochs are validated on the clinical data. An accuracy of 0.9508 and 0.5 for the 30 epochs is presented by the proposed architecture with clinical dataset.

Keywords

Cervical cancer **Deep learning**

Convolutional neural network **Augmentation**

Clinical data **Transfer learning**

References

1. Jemal A, Center MM, DeSantis C, Ward EM (2010) Global patterns of cancer incidence and mortality rates and trends. Cancer Epidemiol Biomark Prev 19(8):1893–1907

2. Cohen PA, Jhingran A, Oaknin A, Denny L (2019) Cervical cancer. Lancet 393(10167):169–182.

https://doi.org/10.1016/S0140-6736(18)32470-X

3. Canavan TP, Doshi NR (2000) Cervical cancer. Am Fam Physician 61(5):1369–1376

4. Saslow D, Solomon D, Lawson HW, Killackey M, Kulasingam SL, Cain J, Garcia FA, Moriarty AT, Waxman AG, Wilbur DC, Wentzensen N, Downs LS Jr, Spitzer M, Moscicki AB, Franco EL, Stoler MH, Schiffman M, Castle PE, Myers ER, ACS-ASCCP-ASCP Cervical Cancer Guideline Committee, American Cancer Society, American Society for Colposcopy and Cervical Pathology, American Society for Clinical Pathology (2012) Screening guidelines for the prevention and early detection of cervical cancer. CA Cancer J Clin 62(3):147–172. https://doi.org/10.3322/caac.21139. Epub 2012 Mar 14. PMID: 22422631; PMCID: PMC3801360

5. Nayar R, Wilbur DC (2017) The Bethesda system for reporting cervical cytology: a historical perspective. Acta Cytol 61(4–5):359–372. https://doi.org/10.1159/000477556. Epub 2017 Jul 11. PMID: 28693017

6. Reddy S, Allan S, Coghlan S, Cooper P (2020) A governance model for the application of AI in

health care. J Am Med Inform Assoc 27(3):491–497. https://doi.org/10.1093/jamia/ocz192

7. Iwai, Tanaka T (2017) Automatic diagnosis supporting system for cervical cancer using image processing. In: 2017 56th annual conference of the society of instrument and control engineers of Japan (SICE), pp 479–482. https://doi.org/10.23919/SICE.2017.8105610

8. William W, Ware A, Basaza-Ejiri AH, Obungoloch J (2018) A review of image analysis and machine learning techniques for automated cervical cancer screening from pap-smear images. Comput Methods Programs Biomed 164:15–22. https://doi.org/10.1016/j.cmpb.2018.05.034

9. Hussain E, Mahanta LB, Das CR, Talukdar RK (2020) A comprehensive study on the multi-class cervical cancer diagnostic prediction on pap smear images using a fusion-based decision from ensemble deep convolutional neural network. Tissue Cell 65:101347

10. Supriyanto E, Pista NA, Ismail L, Rosidi B, Mengko T (2011) Automatic detection system of cervical cancer cells using color intensity classification

11. Sokouti B, Haghipour S, Tabrizi AD (2012) A pilot study on image analysis techniques for extracting early uterine cervix cancer cell features. J Med Syst 36(3):1901–1907. https://doi.org/10.1007/s10916-010-9649-y

12. Ashok B, Aruna D (2016) Comparison of feature selection methods for diagnosis of cervical cancer using SVM classifier

13. Su J, Xu X, He Y, Song J (2016) Automatic detection of cervical cancer cells by a two-level cascade classification system. Anal Cell Pathol 2016:1–11. Article ID 9535027. https://doi.org/10.1155/2016/9535027

14. Sharma M, Singh S, Agrawal P, Madaan V (2016) Classification of clinical dataset of cervical cancer using KNN. Indian J Sci Technol

15. Kumar R, Srivastava R, Srivastava S (2015) Detection and classification of cancer from

microscopic biopsy images using clinically significant and biologically interpretable features. J Med Eng 2015:457906. https://doi.org/10.1155/2015/457906

16. Singh S, Tejaswini V, Murthy RP, Mutgi A (2015) Neural network based automated system for diagnosis of cervical cancer. Int J Biomed Clin Eng

17. Mustafa N, Mat Isa NA, Mashor MY, Othman NH (2007) New features of cervical cells for cervical cancer diagnostic system using neural network. In: International symposium on advanced technology

18. Chen YF, Huang PC, Lin KC, Lin HH, Wang LE, Cheng CC, Chen TP, Chan YK, Chiang JY (2014) Semi-automatic segmentation and classification of pap smear cells. IEEE J Biomed Health Inform 18(1):94–108. https://doi.org/10.1109/JBHI.2013.2250984

19. Shin HC, Roth HR, Gao M, Lu L, Xu Z, Nogues I, Yao J, Mollura D, Summers RM (2016) Deep convolutional neural networks for computer-aided detection: CNN architectures, dataset characteristics and transfer learning. IEEE Trans

Med Imaging 35(5):1285–1298. https://doi.org/10.1109/TMI.2016.2528162

20. Park YR, Kim YJ, Ju W et al (2021) Comparison of machine and deep learning for the classification of cervical cancer based on cervicography images. Sci Rep 11:16143. https://doi.org/10.1038/s41598-021-95748-3

21. Tan X, Li K, Zhang J et al (2021) Automatic model for cervical cancer screening based on convolutional neural network: a retrospective, multicohort, multicenter study. Cancer Cell Int 21:35. https://doi.org/10.1186/s12935-020-01742-6

22. Wu M, Yan C, Liu H, Liu Q, Yin Y (2018) Automatic classification of cervical cancer from cytological images by using convolutional neural network. Biosci Rep 38(6):BSR20181769. https://doi.org/10.1042/BSR20181769

23. Krizhevsky A, Sutskever I, Hinton G (2012) ImageNet classification with deep convolutional neural networks. Adv Neural Inf Process Syst. https://doi.org/10.1145/3065386

24. Han J, Kamber M, Pei J (2012) Classification: basic concepts, chap 8. In: Han J, Kamber M, Pei J (eds) Data management systems, data mining, 3rd edn. The Morgan Kaufmann series. Morgan Kaufmann, pp 327–391. ISBN 9780123814791. https://doi.org/10.1016/B978-0-12-381479-1.00008-3. https://www.sciencedirect.com/science/article/pii/B9780123814791000083

25. Szegedy C, Liu W, Jia YQ, Sermanet P, Reed S, Anguelov D (2015) Going deeper with convolutions. In: Proceedings of the IEEE conference on computer vision and pattern recognition (CVPR), Boston, MA, pp 1–9

26. Han J, Kamber M, Pei J (2012) Classification: advanced methods, chap 9. In: Han J, Kamber M, Pei J (eds) Data management systems, data mining, 3rd edn. The Morgan Kaufmann series. Morgan Kaufmann, pp 393–442. ISBN 9780123814791. https://doi.org/10.1016/B978-0-12-381479-1.00009-5. https://www.sciencedirect.com/science/article/pii/B9780123814791000095

27. Chandraprabha R, Hiremath S (2021) Computer processing of an image: an introduction. In:

Handbook of research on deep learning-based image analysis under constrained and unconstrained environments. IGI Global, pp 1–22

28. Mikołajczyk, Grochowski M (2018) Data augmentation for improving deep learning in image classification problem. In: 2018 international interdisciplinary PhD workshop (IIPhDW), pp 117–122. https://doi.org/10.1109/IIPHDW.2018.8388338

29. Ying X (2019) An overview of over fitting and its solutions. J Phys Conf Ser 1168:022022

30. Jantzen J, Dounias G (2006) Analysis of pap-smear image data

31. Chandraprabha R, Singh, S (2016) Artificial intelligent system for diagnosis of cervical cancer: a brief review and future outline. J Latest Res Eng Technol

32. Jones OT, Calanzani N, Saji S, Duffy SW, Emery J, Hamilton W, Singh H, de Wit NJ, Walter FM (2021) Artificial intelligence techniques that may be applied to primary care data to facilitate earlier diagnosis of cancer: systematic review. J

Med Internet Res 23(3):e23483. https://doi.org/10.2196/23483. PMID: 33656443; PMCID: PMC7970165

Author information

Authors and Affiliations

Electronics and Comunication Engineering, BMS Institute of Technology and Management, Bangalore, India

R. Chandraprabha

Electronics and Telecomunication Engineering, BMS Institute of Technology and Management, Bangalore, India

Seema Singh

Corresponding author

Correspondence to R. Chandraprabha .

Editor information

Editors and Affiliations

Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

N. R. Shetty

National Institute of Advanced Studies (NIAS), Bengaluru, Karnataka, India

L. M. Patnaik

Nitte Meenakshi Institute of Technology, Bengaluru, Karnataka, India

N. H. Prasad

Rights and permissions

Reprints and Permissions

Copyright information

About this paper

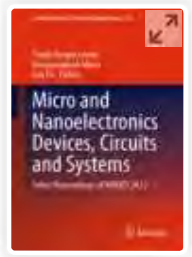
Cite this paper

Chandraprabha, R., Singh, S. (2023). Analysis of Deep Learning Architecture-Based Classifier for the Cervical Cancer Classification. In: Shetty, N.R., Patnaik, L.M., Prasad, N.H. (eds) Emerging Research in Computing, Information, Communication and Applications. Lecture Notes in Electrical Engineering, vol 928. Springer, Singapore. https://doi.org/10.1007/978-981-19-5482-5_23

.RIS .ENW .BIB

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-19-5482-5_23	13 December 2022	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-5481-8	978-981-19-5482-5	Computer Science Computer Science (R0)



Micro and Nanoelectronics Devices, Circuits and Systems pp 359–366

Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection

M. Lakshmi Prasanna & V. R. Anitha

Conference paper | First Online: 12 September 2022

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 904)

Abstract

Many studies have shown that most MEMS devices use light beams whose deflection response is the main focus before choosing it for specific applications. This paper deals with the simulation of the MEMS-based cantilever beam for low-pressure detection. COMSOL 5.4 multiphysics FEM model is used to study the behavior of beam and compared

with the conventional cantilever beam. Both are analyzed with two different materials SiO_2 and SU-8. The change in the deflection has been analyzed with various pressure values applied on the cantilever beam. Through simulation and analysis results, it is observed that the deflection of the proposed beam is good at lower pressure values.

Keywords

Microcantilever **Deflection**

Rectangular beam **FEM**

References

1. Guo K, Jiang B, Liu B, Xingeng Li, Wu Y, Tian S, Gao Z, Zong L, Yao S, Zhao M, Mi C, Zhu G (2021) Study on the progress of piezoelectric microcantilever beam micromass sensor. IOP Conf Ser Earth Environ Sci 651:022091. https://doi.org/10.1088/1755-1315/651/2/022091

2. Ali S, Bhuvaneswari H, Kumar B (2020) Design and modeling of a large deflection microcantilever using rectangular SCR. AIP Conf Proc 2281:020037. https://doi.org/10.1063/5.0027942

3. Rotake D, Darji A (2020) Design and reliability testing of microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.13140/RG.2.2.25052.92803

4. Ashok A, Nighot R, Sahu N, Pal P, Pandey A (2019) Design and analysis of microcantilever beams based on arrow shape. Microsyst Technol 25:4379–4390. https://doi.org/10.1007/s00542-019-04555-4

5. Passian A, Thundat T, Thanihaichelvan M (2017) Microcantilever sensors. https://doi.org/10.1016/B978-0-12-803581-8.10525-9

6. Rotake D, Darji A, Singh J (2020) Ultrasensitive multi-arm-microcantilever-based piezoresistive sensor for BioMEMS application. https://doi.org/10.1109/VDAT50263.2020.9190249

7. Gharge B, Upadhye V, Bodas D (2015) Design and simulation of microcantilevers for detection of pathogens. https://doi.org/10.1109/ISPTS.2015.7220122

8. Parsediya D (2016) Deflection and stresses of effective micro-cantilever beam designs under low mass loading:111–114. https://doi.org/10.1109/ICEPES.2016.7915915

9. Mathew R, Sankar R (2018) A review on surface stress-based miniaturized piezoresistive SU-8 polymeric cantilever sensors. Nano-Micro Lett 10:1–41. https://doi.org/10.1007/s40820-018-0189-1

10. Gopinath PG (2014) Design and characterization of various shapes of microcantilever for human immunodeficiency virus detection. Int J Res Eng Technol 03:24–29. https://doi.org/10.15623/ijret.2014.0310004

11. Schmid S (2009) Electrostatically actuated all-polymer microbeam resonators—characterization and application

12. Lamba M, Singh K, Chaudhary H (2020) Design analysis of polysilicon piezoresistors PDMS (polydimethylsiloxane) microcantilever based MEMS force sensor. Int J Mod Phys B 34:2050072.https://doi.org/10.1142/S0217979220500721

Acknowledgements

A Special thanks to the National MEMS Design Centre, Sree Vidyanikethan Engineering College, A. Rangampet, Tirupati, for providing the research facility.

Author information

Authors and Affiliations

Department of Electronics and Communication Engineering, JNTUA, Ananthapuramu, Andhra Pradesh, India

M. Lakshmi Prasanna

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Yelahanka, Bangalore, Karnataka, 560064, India

V. R. Anitha

Corresponding author

Correspondence to M. Lakshmi Prasanna .

Editor information

Editors and Affiliations

Electronics and Communication Engineering, National Institute Of Technology Silchar, Silchar, India

Trupti Ranjan Lenka

Electrical and Computer Engineering, New Jersey Institute of Technology, Newark, NJ, USA

Durgamadhab Misra

Electronic Materials Engineering, Australian National University, Canberra, ACT, Australia

Lan Fu

Rights and permissions

Copyright information

About this paper

Cite this paper

Lakshmi Prasanna, M., Anitha, V.R. (2023). Design and Simulation Analysis of a Piezoresistive Cantilever Beam for Low-Pressure Detection. In: Lenka, T.R., Misra, D., Fu, L. (eds) Micro and Nanoelectronics Devices, Circuits and Systems. Lecture Notes in Electrical Engineering, vol 904. Springer, Singapore. https://doi.org/10.1007/978-981-19-2308-1_37

.RIS .ENW .BIB

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-19-2308-1_37	12 September 2022	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-19-2307-4	978-981-19-2308-1	Engineering Engineering (R0)

Proceedings of the 7th International Conference on Trends in Electronics and Informatics (ICOEI 2023)
IEEE Xplore Part Number: CFP23J32-ART; ISBN: 979-8-3503-9728-4

A Novel IoT Framework and Device Architecture for Efficient Smart city Implementation

Sheetalrani R Kawale[1]
Assistant Professor, Department of Computer Science,
Karnataka State Akkamahadevi Women's University,
Vijayapura, Karnataka, India.
Email: sheetalrkawale@gmail.com

KDV Prasad[2]
Assistant Professor (Research)
Symbiosis Institute of Business Management (SIBM)
Symbiosis International (Deemed University) (SIU)
Off Bangalore Highway, Kothur Mandal
Village, Mahabubnagar, Hyderabad, Telangana, India
Email: kdv.prasad@sibmhyd.edu.in

D Palanikkumar[3]
Professor, Department of CSE, Dr NGP Institute of
Technology, Coimbatore, Tamilnadu, India.
Email: palanikkumard@gmail.com

P. Arockia Mary[4]
Professor, Department of Information Technology, V.S.B.
Engineering College, Karur, Tamilnadu, India.
Email: mary.vsbec@gmail.com

A.Yasmine begum[5]
Associate Professor, Department of Electronics and
Instrumentation Engneering, Mohan Babu University (Erst
while Sree Vidyanikethan Engineering College)
A.rangampet, Andhra Pradesh, India.
Email: a.yasminebegum@vidyanikethan.edu

Dankan Gowda V[6]
Department of Electronics and Communication Engineering,
BMS Institute of Technology and Management,
Bangalore, Karnataka, India.
Email: dankan.v@bmsit.in

Abstract: **Internet of Things (IoT) apps will provide fast and effortless management of smart cities. This networked service management is dependent on the Internet Service Provider (ISP). This framework reduces physical human interaction in banking, railway counters, water and power counters, and so on. The IoT based applications makes the smart cities more sustainable.**
Recently, the Indian government is more focused and supportive towards building a reliable IoT based framework for developing innovative products, and businesses that rely mostly on cyberspace. This study discusses about the advantages, disadvantages and future possibilities of smart cities. The importance of the Internet of Things (IoT) in the development of India's smart cities, as well as how widespread adoption may swiftly disrupt people's everyday routines are also discussed.

Keywords: Internet of Things (IoT), Smart city, cloud, Sensor, Development, and Data Collection.

I. INTRODUCTION

The Internet of Things (IoT) concept is gaining significant research attention as it is applied in a wide range of industries, including the transformation of wealthy metropolitan communities, critical infrastructure management, transportation, communication, and so on. As the utility grows, more and more structured data may be explored and the data can be secured and shared under various conditions [1].

In recent years, there has been a major shift in the most effective methods of securing and retrieving data both personally and publicly. With the advanced research, IoT applications are implemented on a large scale [2]. The IoT users are connected in a distributed manner whereas these connections may also introduce network and data complexity challenges [3]. The web related inquiries needs to explored thoroughly explored. Some analysis has also been proposed to perform content appropriation. Recently, in the IoT infrastructure, everything between objects is black boxed [4]. Therefore, most clients are not aware of its intrinsic features [5]. IoT and internet based gadgets are also used in healthcare applications in different forms [6]. The energy of the IoT will be completely acknowledged to enhance the human wellbeing, condition and profitability. The cloud based devices are used to monitor the air quality or it may be provided in the form of wearable devices to monitor patients in healthcare domain [7]. In our daily lives, we face a number of challenges while travelling, for example, movement, air contamination, time delay, and so on, and we also face a number of issues in toll accumulation, for example, long waiting times to pay the sum. With IoT technologies, the number of vehicles passing through a tollbooth can be efficiently monitored and organized on the national highway in a smart and productive way [8]. The primary goal of electronic toll collection is to eliminate delay on toll roads by collecting charges electronically with no client [9]. They have enabled the things to not only perceive data but also interact with the physical world. With rapid population growth, there is always a demand on transportation offices. This challenges can be eliminated by using a

2023 7th International Conference on Trends in Electronics and Informatics (ICOEI) | 979-8-3503-9728-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICOEI56765.2023.10125677

smart transportation architecture [10]. Singapore's MRT planning structure is particularly successful in terms of group management and punctuality. However, this study has considered the transport framework in Singapore because it isn't particularly advanced in terms of innovation, and there is a requirement for travellers to analyse various transportation alternatives effectively [11]. There ought to be some sort of invention that can guide the passenger in this situation.

II. LITERATURE SURVEY

The topic of urban manageability has gone to the front of research in the previous decade. It is coming about because of the way that the extent of urban populace consistently increment, as per the estimations 70% of the populace will be urban subject by 2050. Then again urban areas assume a huge part in manageability, atmosphere insurance and the lessening of unsafe discharges [12]. A few creative activities were performed on the region of the foundation of carbon-unbiased urban areas of zero vitality utilization and maintainable districts. The motivation behind Smart City ideas is to boost urban life quality and to expand vitality effectiveness, to diminish unsafe outflows and to improve the nature of urban administrations using the potential outcomes gave by the accessible ICT-gadgets (Information and Communication Technology). These days the state of the urban condition, the urban issues can be always estimated and observed by methods for instruments and with the assistance of information investigation and process figures we can present to-date answers for them. Insightful gadgets have gone to the concentration while examining city associations [13]. Thus the significance of examinations inquiring about issue discernment does not just intend to demonstrate the urban legislative issues, how the neighborhood populace thinks about some urban inquiries, however political responses can be additionally advanced and general society engagement can be likewise developed identified with an issue separately [14].

Transport data framework of Singapore transport generally have data related to transport numbers and transport stops. Be that as it may, hunting down applicable data is very agonizing in this labyrinth of information. Additionally, a traveler is uninformed of how much time a transport would take in the present movement conditions to reach his beginning stop [17]. Web of things can help in having a collaboration between the travelers what's more, the transport through the methods for a cell phone application. This will be followed by the transports running along that specific course. They will detect the movement on that course to give continuous appraisals [18]. This same innovation empowers the client to know when the transport would land at his stop of starting point. One of the significant wellsprings of income for some, nations is transportation. Human transportation, co-ordinations, private activity shapes some portion of transportation. The creating and created nations are endeavoring to enhance the standard and proficiency of transportation framework. In light of the congested driving conditions, and absence of appropriate activity administration framework, time and cash of the open is being squandered. Products transportation, hardware and human transportation are the key components which impact the advancement of enterprises [19]. The improvement of activity observing and controlling framework is an essential prerequisite in every one of the nations. The blue print of Intelligent Traffic System. As there is a fast development in the rush hour gridlock nowadays, the specialists need to discover diverse approaches to maintain a strategic distance from these issues . Measures, for example, growing new streets, over extensions, under scaffolds, fly-overs, burrow streets and so forth are a few choices [20]. Aside from this, utilization of trains, for example, Light Rapid Transit [LRT] trains is helpful.

A remote sensor based framework which will be arranged in the city streets and read the activity information and send it to the showcases or street signs which are advanced driven sheets giving data about all information [21]. To stay away from this water flood sensor will read the water information of the stream scaffold and send to control station which will send this to dynamic driven street signs where that will be shown [22]. The fourth arrangement is for the avalanches occurring in the uneven regions which are the reason for roads turned parking lots and substantial misfortune. The territories where avalanche happens are situated in remote parts so not very many specialized gadgets are accessible there. In this utilizing ultrasonic sensor based avalanche locator sensor will be put at such places which will recognize avalanche and send data to the calamity administration framework utilizing GSM or XBee.

III. ELEMENTS OF IOT

Examine how the IoT is evolving and how this provides a general understanding of its importance and convenience. In the following sections, we go over six crucial components that are predicted to transmit the efficiency of the IoT as shown in Figure 1.

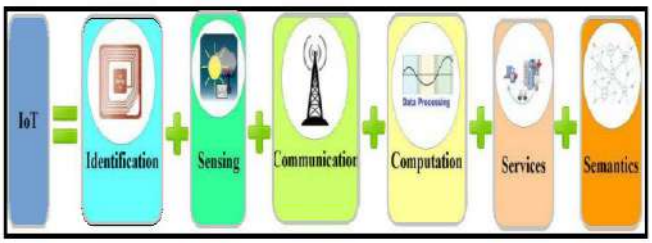


Figure.1. IoT Elements

Identification

For the Internet of Things to successfully term and challenge established businesses, unambiguous proof is of the utmost importance. The identification and placement of things in the Internet of Things must be monitored so that problems may be identified and resolved. If you have a query about a specific heat sensor, for instance, its ID will suggest its name, and its convey will provide its

location inside an exchanges orchestrate. Moreover, IoT device monitoring infrastructures should include support for both IPv6 and IPv4. Internet Protocol version 6.0 (IPv6) monitoring is made reliable for low-control distant systems by the 6LoWPAN's use of an additional weight instrument on top of IPv6 headers.

Sensing

Companies like Wemo, revolv, and Smart Objects provide excellent hubs and flexible software that allow users to control and monitor a wide variety of smart devices and mechanical systems in their homes from their smartphones.

Communication

The IoT communication drives several object queries in unison to bypass otherwise exceptional service providers. The IoT's nerve centres need to use low-manipulate, internally-sighted, lossy, and noisy communication joints on a regular basis. Semiactive and dynamic markings sometimes make use of board control. As a subculture, NFC operates in the high repetition band at 13.56 MHz and can provide data rates of up to 424 kbps. Correspondence between specific per-customer details and inactive markings or active per-customer details is optimal at a range of up to 10 cm.

Computation

Another fundamental IoT technique is outlined by cloud platforms. All of these steps make it possible for businesses to outsource their data storage and management to the cloud, allowing for more efficient management of large files and, ultimately, better insights for end users. A variety of public and private cloud deployment models and settings are available for use by IoT businesses.

Services

As a rule, there are many categories under which IoT services may be requested. These categories include services linked to individuality, data aggregation administrations, collaboratively-minded products and services, and always-present administrations. Of the many types of enterprises, the most fundamental and important are those with which a person is linked. Every programme that has to send verified information onto the Internet must be able to identify the data as genuine.

Semantics

Semantic in the IoT implies the capability to evacuate adapting cleverly by altered technologies to give the required organizations. Data withdrawal fuses finding besides, utilizing assets and showing data. Furthermore, it fuses seeing and exploring records to grasp the correct decision to give the right organization. Subsequently, semantic addresses the intelligence of the IoT through directing solicitations to the other side supply. Also, it diminishes transmission limit requirements without upsetting related assets, for instance, battery life, code measure, imperativeness ate up for taking care of, and memory assess. EXI changes over XML messages to parallel to decrease the required information transmission and point of confinement the required accumulating size.

IV. SENSOR TECHNOLOGY

The sensors detect various changes, analyze them, store the data, and then show them all together for analysis and management. Sensors play a major role in designing and implementing a smart home.

A PIC controller, light sensors, and a PIR sensor were all purchased locally to create this smart system, which gives us command over our dwelling's electrical appliances. The mix of the interconnected and canny intra vehicle correspondence frameworks and the vehicle to foundation into the general IoT benefit stages will offer the likelihood to grow new applications and Society of Automotive Engineers,services it is normal that 80% of vehicles in Europe will be two-path associated by 2018.

The developing Smart Grid is anticipated to use a different concept of transmission arrange that may successfully route the energy that is introduced from both concentrated and dispersed plants to the ultimate customer with high security and nature of deliver models. Therefore, it is expected that the smart Grid will involve the utilisation of a type of "net" in which energy packages are managed similarly to data packages across switches and portals that can independently choose the best route for the parcel to achieve its goal with the highest possible degree of trustworthiness. here the "net of power" idea is defined as a device foundation due to common knowledge.

Figure 2 demonstrates the Internet of Things (IoT) based sensor technology. A variety of software-based game plans, most of which are not compatible with one another, are now on display, illustrating the market for prosperity checking gadgets. Although it is reasonable to assume some degree of variability in unit costs, reaching the overall agreement goal of reduced development costs across existing and new divisions without a more robust strategy would be very difficult. A patient in a hospital whose physiological status calls for careful attention may have such close attention maintained with the help of the loT by means of continuous, constant, non-invasive monitoring.

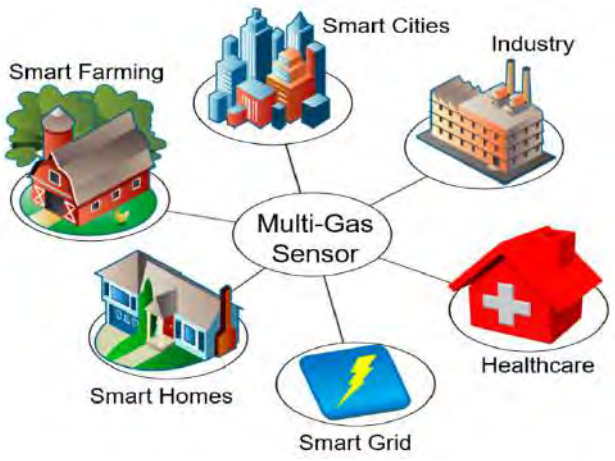


Figure.2. IoT Sensor Technology

Furthermore, the technology is increasingly being employed for remote monitoring by employing low-cost, remote setups connected through the loT. The Ambient Assisted Living (AAL) frameworks are incorporated with a view to provide administrations in day-to-day sports, wellbeing and movement tracking, wellness and security by having access to medicinal and disaster frameworks, and encouraging to provide health support.

V. IoT FRAMEWORK

The typical engineering of IoT setups is often far more complex than the design of most enterprise frameworks. The fact that the foundation of most frameworks—the server farm is only a small part of multifaceted IoT frameworks. For IoT setups to function, the field-operated devices must be coordinated. While the concept behind these devices is completely distinct from that of online, desktop, or even mobile clients, we need a bridging component between the field devices and the server. Figure.3 shows the IoT Framework.

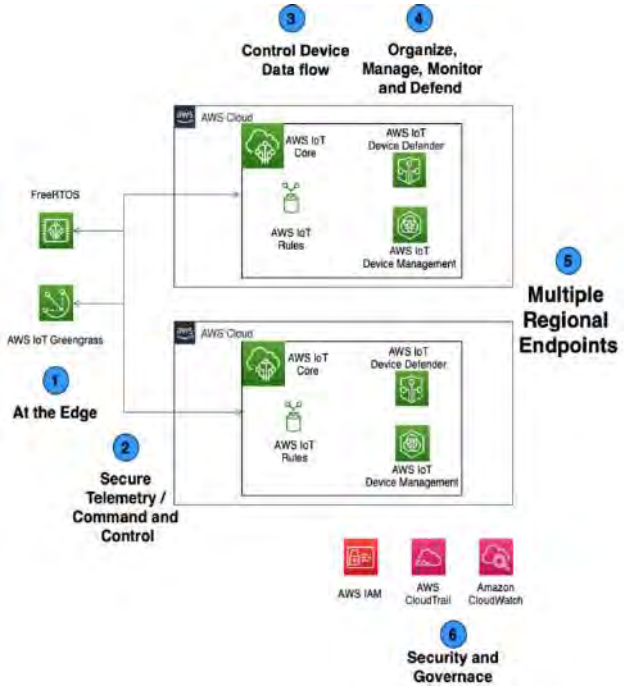


Figure.3. IoT Framework

Both sensors and actuators are used to control and modify the selected environment. Just unlike the problems of data collection and management, communication may provide its own set of challenges. In any case, the issue with this approach is its unbending nature, which means rigidity when it comes to business forms utilizing SOA. On the other hand, semantic interoperability can be accomplished by giving suitable semantic arbiters (interpreters) at every user end. Internet of Things is essential to recall that numerous things, which will not have changeless system network yet depend on supporting knowledge in their neighbourhood condition or in remote data frameworks. The IoT Device Architecture is presented in Figure.4. Things will require the capacity to impart their area, state and necessities to data frameworks that have more changeless or on the other hand more solid system network. Through data frameworks, an advanced partner of the thing can be checked or indeed, even showed in a virtual portrayal, to such an extent that remote approved elements could inquiry or refresh the condition of a person thing or impact its fate. Handheld gadgets may likewise be utilized by lodge group amid air ship turnaround tasks, to quickly watch that all required security gear (life coats, breathing devices, fire quenchers) are available furthermore, remedy and not lost.

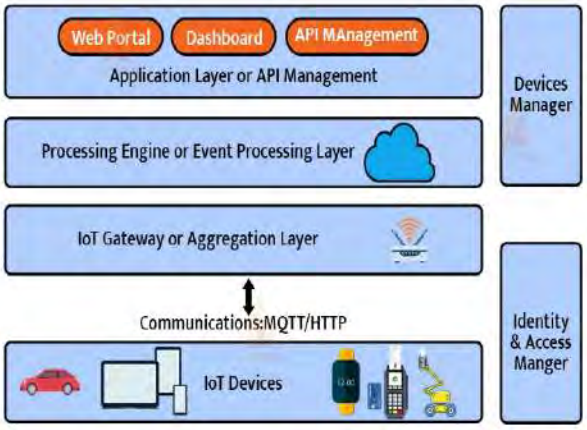


Figure.4. IoT Device Architecture

In such situations, one can conceive pre-situating onto the handheld gadget the data about the normal show of wellbeing gear or insights about the total upkeep history of each part known to be introduced on that particular airplane, to such an extent that the pre-stored data is instantly accessible at the time and place of communication with the protest. The memory of the handheld gadget can likewise be utilized to briefly record any updates, for example, alterations to the parts, side effects watched, missing wellbeing gear, and so on., so that those updates can be synchronized to the organize when the hand-held gadget returns inside scope of system network, for example, Wi-Fi.

VI. PROCESS FOR IOT

Since smart devices receive a massive volume of sensor data, registration and storage assets are required to evaluate, load, and process this information. Cloud-based registration and storage assets are most popular due to the cloud's massive data management, flexibility, and adaptability. Yet, due to the reasons described below, this will not be sufficient to fulfil the criteria of many IoT requests. The first characteristic is adaptability; many cutting-edge technologies may be employed in a variety of ways. Due to the dynamic system circumstances across various locations, communication with the cloud server becomes difficult. Continuous and trustworthy activation should be established through cloud communication.

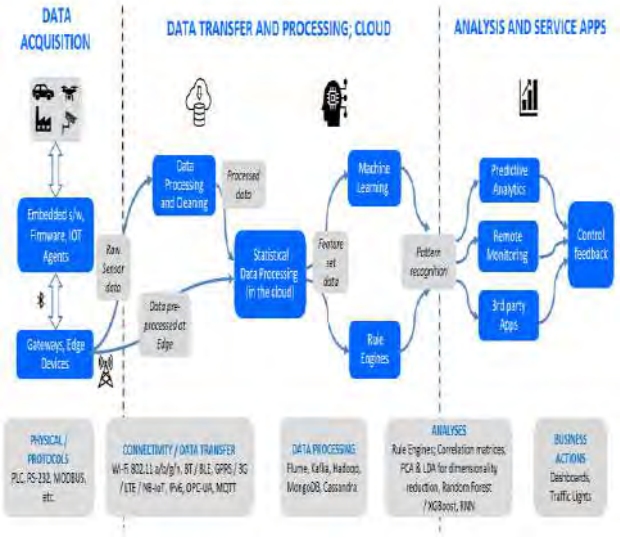


Figure.5. Process for IoT

Investing a lot of money to generate responses is a waste of time and money. Applications that are very sensitive to latency and need constant feedback may be out of reach with this framework. As an additional complication, remote connections may cause the communication to be lossy leading to inaccurate data.

VII. RESULTS AND DISCUSSION

Raspberry Pi is a new advanced framework that will capture and collect various environmental aspects as said above with the help of sensors attached; Memory chip on Pi keeps the accumulated data as identical memory chip. Next, on the yield side, a liquid crystal display is linked for displaying the outcome except ON OFF exchanges for server access.

Figure.6. Sensor Data on IoT Service

To monitor a particular place remotely, the client can sign in via internet by authenticating name and secret key provided by the customer for a specific server. After verifying the secret key, the yield graphical depiction will be shared with the user. Raspberry pi will be continuously store data on a cloud server, and customers will become better acquainted with the saved data on hourly and regular basis.

The Raspberry Pi board may function as a web server and data encryption system inside the framework. It takes readings from a variety of sensors, including those for temperature, humidity, pressure, altitude, brightness, and rainfall. This data is subsequently sent to the client end via the HTTP standard. Thingspeak.com gives customers access to real-time data from any location in the world. Connecting to the board's network is done through LAN via its Ethernet connection or a USB dongle via its USB port. Thingspeak.com displays its findings in a visually appealing layout. In this location, a single channel is constructed, and all six fields are placed within it. Pressure, temperature, and humidity are shown in fields 2 and 4, whereas barometric pressure is shown in field 1. The picture depicts real-time data from each of the four domains. I'd like to refer you to Figure 6.

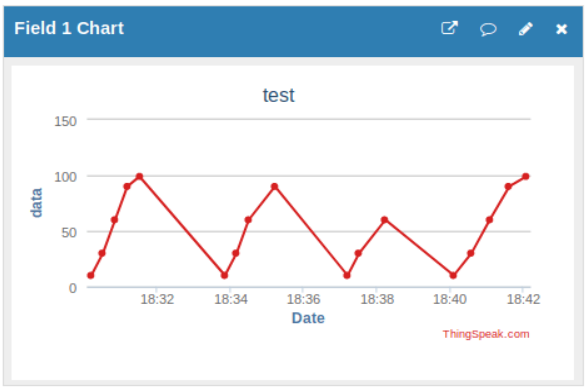


Figure.7. Logging Data

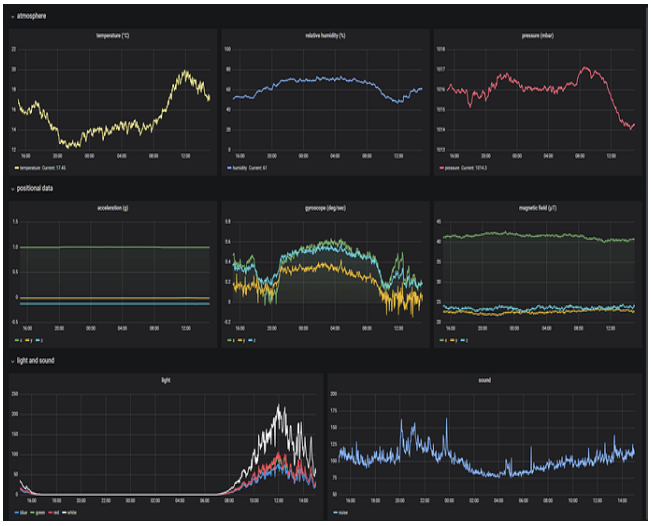


Figure.8. Sensor Data from Mirco weather station.

As a result of being a smaller yet effective device, we can conclude that the Raspberry Pi productively record every important aspect required for a climate monitoring station. RPI GPIO plugins are necessary to provide access to raspberry PI's GPIO's. RPI GPIO collection is also required for developing a conclusive Application and the HTTPLIB to establish raspberry web association and concluded with a web information transfer to the server.

VIII. CONCLUSION

IoT frameworks provide various opportunities and benefits to people and society. These frameworks must be safe, powerful, and useful in order to be successful. Although progress has been made in improving the security of existing frameworks, IoT frameworks need considerably better quality by introducing additional functionalities. Current strategies for developing high-trustworthy frameworks, such as air ship control programming, etc. Several IoT frameworks will have equal credibility requirements, but will have far less time-consuming and costly development forms. As previously said, significant advancements in practices, devices, and

improvement forms are essential to achieve the greatest benefits from IoT. More automated analysis of multi-component frameworks under complex and dynamic situations should be developed. IoT frameworks will have dynamic registration and function under various situations that may contain inadequate components.

REFRENCES

1. A. Zanella, N. Bui, A. Castellani, L. Vangelista and M. Zorzi, "Internet of Things for Smart Cities," in *IEEE Internet of Things Journal*, vol. 1, no. 1, pp. 22-32, Feb. 2014, doi: 10.1109/JIOT.2014.2306328.
2. H. Rajab and T. Cinkelr, "IoT based Smart Cities," *2018 International Symposium on Networks, Computers and Communications (ISNCC)*, Rome, Italy, 2018, pp. 1-4, doi: 10.1109/ISNCC.2018.8530997.
3. K. Prasad, A. Y. Begum, R. C. Tanguturi and B. Ahmed, "Implementation of an Early Warning System using Internet of Things and Rainfall Threshold," *2022 6th International Conference on Electronics, Communication and Aerospace Technology*, Coimbatore, India, 2022, pp. 491-496, doi: 10.1109/ICECA55336.2022.10009211.
4. Priyadarshani Shivkumar Mali, Hemant. A. Tirmare, Varsha Amol Suryawanshi and Abhay Chaturvedi (2022), Novel Predictive Control and Monitoring System based on IoT for Evaluating Industrial Safety Measures. IJEER 10(4), 1050-1057. DOI: 10.37391/IJEER.100448.
5. S. Reddy P. P. S. Patwal, "Data Analytics and Cloud-Based Platform for Internet of Things Applications in Smart Cities," *2022 International Conference on Industry 4.0 Technology (I4Tech)*, 2022, pp. 1-6, doi: 10.1109/I4Tech55392.2022.9952780.
6. S. B. M, P. Pavankumar, N. K. Darwante, "Performance Monitoring and Dynamic Scaling Algorithm for Queue Based Internet of Things," *2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES)*, 2022, pp. 1-7, doi: 10.1109/ICSES55317.2022.9914108.
7. Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (eds) Recent Advances in Metrology . Lecture Notes in Electrical Engineering, vol 906. Springer, Singapore. https://doi.org/10.1007/978-981-19-2468-2_3.
8. B. Kameswara Rao, Ravi Shankar, Parismita Sarma, Abhay Chaturvedi, Naziya Hussain, Industrial quality healthcare services using Internet of Things and fog computing approach, Measurement: Sensors, Volume 24, 2022, 100517, ISSN 2665-9174, https://doi.org/10.1016/j.measen.2022.100517 .
9. A. Singla, N. Sharma, "IoT Group Key Management using Incremental Gaussian Mixture Model," *2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC)*, 2022, pp. 469-474, doi: 10.1109/ICESC54411.2022.9885644.
10. A. Sharma, K. S and M. R. Arun, "Priority Queueing Model-Based IoT Middleware for Load Balancing," *2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS)*, 2022, pp. 425-430, doi: 10.1109/ICICCS53718.2022.9788218.
11. Namitha A R, Manu Y M, Rashmi G R and Veera Sivakumar Chinamuttevi (2022), IOT Based Smart Health Care System to Monitor Covid-19 Patients. IJEER, 10(1), 36-40. DOI: 10.37391/IJEER.100105.
12. M Sandeep Prabhu, M Ramesha, Jayashree M Kudari and Ansuman Samal (2021), Smart Agriculture and Smart Farming using IoT Technology, Journal of Physics: Conference Series, 2089(1), pp. 012038. https://doi.org/10.1088/1742-6596/2089/1/012038.
13. Arudra Annepu, M Ramesha,K Prashantha Kumar and Pallavi Singh (2021), IoT Enabled Smart Lighting System for Smart Cities, Journal of Physics: Conference Series, 2089 (1), pp. 012037. https://doi.org/10.1088/1742-6596/2089/1/012037.
14. Pankaj Mudholkar, Megha Mudholkar, B S Puneeth Kumar and S. Srinivasulu Raju (2021), Smart Villages: IoT Technology Based Transformation, Journal of Physics: Conference Series, 2070(1), pp. 012128. https://doi.org/10.1088/1742-6596/2070/1/012128.
15. Ramesh Naidu and N. Guruprasad (2021), A High-Availability and Integrity Layer for Cloud Storage, Cloud Computing Security: From Single to Multi-Clouds, Journal of Physics: Conference Series, 1921 (1), pp. 012072. https://doi.org/10.1088/1742-6596/1921/1/012072.
16. K. Jeevan and B. M. Sathisha, "Implementation of IoT Based Wireless Electronic Stethoscope," *2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)*, 2020, pp. 103-106, doi: 10.1109/MPCIT51588.2020.9350476.
17. G. Naveena Pai, M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.
18. S. B. Sridhara, K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," *Adv. Math. Sci. J.*, vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.
19. N. Guruprasad, "Design and implementation of cryptcloud system for securing files in cloud," *Adv. Math. Sci. J.*, vol. 9, no. 7, pp. 4485–4493, 2020, doi: 10.37418/amsj.9.7.17.
20. Shivashankar, and S. Mehta, "MANET topology for disaster management using wireless sensor network," in *International Conference on Communication and Signal Processing, ICCSP 2016*, 2016, pp. 0736–0740, doi: 10.1109/ICCSP.2016.7754242.
21. M. Talebkhah, A. Sali, M. Marjani, M. Gordan, S. J. Hashim and F. Z. Rokhani, "IoT and Big Data Applications in Smart Cities: Recent Advances, Challenges, and Critical Issues," in *IEEE Access*, vol. 9, pp. 55465-55484, 2021, doi: 10.1109/ACCESS.2021.3070905.
22. R. R. Harmon, E. G. Castro-Leon and S. Bhide, "Smart cities and the Internet of Things," *2015 Portland International Conference on Management of Engineering and Technology (PICMET)*, Portland, OR, USA, 2015, pp. 485-494, doi: 10.1109/PICMET.2015.7273174.

Proceedings of the 7th International Conference on Computing Methodologies and Communication (ICCMC-2023)
IEEE Xplore Part Number: CFP23K25-ART; ISBN: 978-1-6654-6408-6

IoT Wearable Breast Temperature Assessment System

B. Ashreetha[1]
Assistant Professor, Department of Electronics & Communication Engineering College, Sree Vidyanikethan Engineering College, Tirupati,Andhra Pradesh.
Email:ashreetha.b@vidyanikethan.edu

Dankan Gowda V[2]
Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, Karnataka.
Email: dankan.v@bmsit.in

Harishchander Anandaram[3]
Assistant Professor, Centre for Excellence in Computational Engineering and Networking, Amrita Vishwa Vidyapeetham, Coimbatore, Tamilnadu.
Email:a_harishchander@cb.amrita.edu

Nithya B A[4]
Assistant Professor, Department of Computer Science, School of Applied Sciences, REVA UNIVERSITY, Bengaluru, Karnataka. Email: nithya.ba@reva.edu.in

Neeraj Gupta[5]
Professor, Department of Computer Science and Engineering, SGT University, Gurugram, Haryana.
Email:neerajgupta3729@gmail.com

Basant Kumar Verma[6]
Professor, Department of Computer Science and Engineering, SGT University, Gurugram, Haryana.
Email: bkverma.3474@gmail.com

2023 7th International Conference on Computing Methodologies and Communication (ICCMC) | 978-1-6654-6408-6/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICCMC56507.2023.10083511

Abstract: **Cancer is an undesirable cell with odd characteristic varies from normal cell of the breast tissue. This will develop swiftly and infiltrate surrounding tissue and forms as tumor which happens in both men and women. After lung cancer, breast cancer has become the largest cause of malignancies in women which raise the mortality rate. In this article, Sensor enabled wearable device framework to detect breast temperature abnormalities is suggested. Early work on wearable sensors supported to monitor breast temperature is presented. The concept of a wearable garment with sensor-enabled patches is also considered. Detailed architecture of the proposed framework is explained, and implementation specifics are examined. Later the statistical highlights help to process the sensor-generated breast temperature data are explained. The performance of the proposed framework is examined by taking abnormal and normal patient temperature dataset and the findings are described using attractive figure representations.**

Keywords: Abnormalities, Cancer, Detection, Internet of Things, Temperature, Sensor.

I. INTRODUCTION

Internet of Things (IoT) offers an environment that links different physical items like sensors and actuators to the Internet and enables these things to collect and send the data via a network. The IoT may be well-defined as an universe of connected objects that are suited for sensing, actuating, and communicating among themselves and with the environment. IoT also delivers the potential to exchange information and autonomously respond to the physical world occurrences by activating processes and creating services with or without human participation. A smart object is a physical item in which a processor, data loading system, sensor logic, and communication mechanism are implanted [1]. The smart objects may be anyplace nowadays because to its low-cost and compact size, its applications are not limited to consumer medical devices, electrical gadgets, cameras, household appliances, and all types of data-generating devices. Healthcare, agriculture, transit, industrial control, water management, environment safety, security, and item tracking are important application areas of IoT.

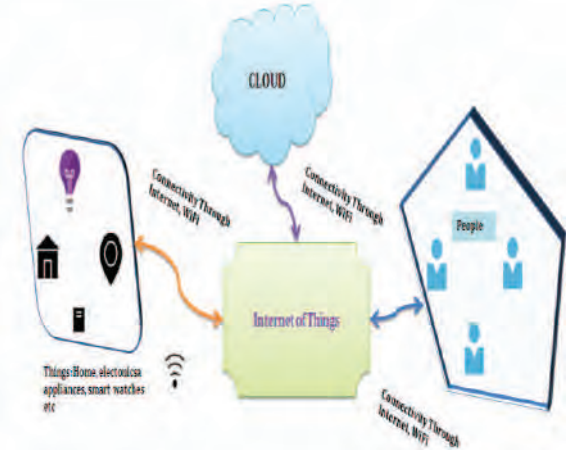


Fig.1. General Architecture of IoT System

People, devices, and cloud services are all connected through the Internet in Fig.1 to create new use cases and business models. Sensors, computing devices, and other mechanical or digital items having unique identifiers (UIDs) are examples of "things," which are defined as devices that can detect, calculate, send, and receive data across a system without requiring human interaction. When a system generates a large amount of data, the cloud offers a service to store and manage it[2]. The Internet may be used to interface with devices and get data from the cloud. As a result of this system, several acts may be performed in a closed loop around the planet, which is detailed in detail. In addition, smart systems are another

name for IoT buildings. Sensors and actuators make up the IoT system. Built-in sensors and actuators in IoT devices allow it to assess and take appropriate action at the appropriate moment. So, the Internet of Things (IoT) is the network of interconnected devices that connects billions of devices worldwide [3].

A well-defined IoT system is a self-organizing worldwide system structure built on normal and interoperable communication procedures, where physical and simulated things have identities, physical qualities, and virtual characters and use smart analysis, and are seamlessly integrated into the information system [4]. The Internet of Things (IoT) is a relatively new concept that can be described as a network of physical objects that are linked together through the Internet and enabling users to access the Internet from anywhere and at any time. Sensors, actuators, software, etc. are integrated into the physical object. The IoT's structure is different from that of a traditional processing approach. It considers hardware and software to be part of a single system and offers a path for new tools to be added. In the section-II describes the literature survey, section III presents various wearables in health monitoring, IoT based health care system is described in section IV. Architecture of Proposed Model is presented in section V. Results and Discussion and followed by conclusion is presented in section VI. and VII. Respectively.

II. LITERATURE SURVEY

Due to IoT rising in the medical area the patient health care becomes easier and smart. It provides a more personalised kind of treatment where users are self-handling and self-examining their health and clinicians may enhance the experience of care. Smart health care systems enable remote connections among patients and clinicians to deliver important services like health monitor, health conditions monitoring and recording of critical medical information [7]. Medical IoT plays a major role in the health monitoring of chronic illness patients and senior care. IoT is utilized in monitoring the patient health responses during surgery and how the patient's health reacts to treatment. If emergency increases automated signals provided to medical specialists or corresponding persons. IoT delivers telemedicine services to patients in distant places with less medical facilities and gives first-aid assistance during emergency scenarios [6]. Smart health care is the key use of IoT discussed here. According to World Health Organization's (WHO) inquiry reports on Cancer, every year forty thousand plus women perished due of breast cancer. As per international study, 14 percent breast cancer mortality rate occurred out of 23 percent (more than 1.5 million patients) in 2010. In 2013, 39,620 breast cancer patients are perished out of 232,340 cases in US. In 2014, out of 232670 new incidences roughly 62570 instances are in situ breast tumor as per the U.S. cancer society. In 2018, as per the USA cancer data 252,710 females impacted owing to breast cancer out of them 40,610 fatalities occurred. In 2019, as per WHO Figs roughly 62,7,000 women fatalities happed owing to breast cancer, this sum is almost 15 percent of total cancer deaths among woman. Current incidences of breast cancer are greater in developed nations as comparison to developing countries and the incidence rate is larger in developing countries as compared to developed countries. Approximately 1,55,000 new breast cancer cases are recorded every year in India and this Fig projected to be double by 2030. Premature discovery of abnormal breast tumor might notify the patients to attend better therapy and recovery. Breast cancer statistics in India is presented in fig.2.

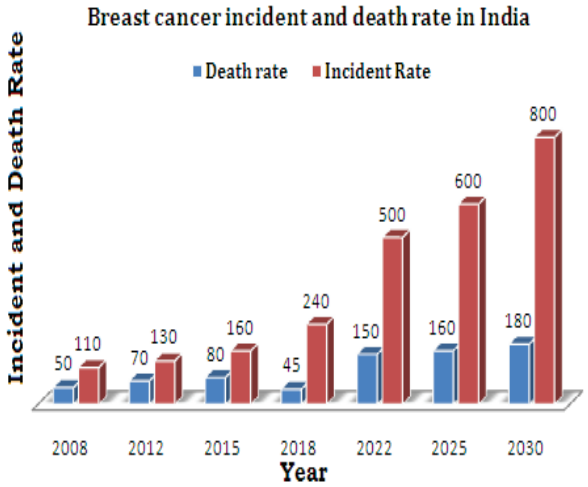


Fig.2. Breast cancer incident and death rate in India

Breast cancer screening aids to discover tumor in breast[8]. The screening methods available to identify breast cancer include breast self-examination, clinical breast examination, mammography, thermography, MRI, and ultrasound. Especially, for the woman who is having family history frequent breast health monitoring is necessary to find the breast abnormalities at early stage. As per medical statistics, the breast cancer family history lady has to attend periodic screening from the age of 20 years to check breast health relatively.

BSE refers a physical examination of own breast tissue by sight and feel on breast, that implies person alone may examine both breasts by following the BSE user guide. Another alternative screening method is thermography, which is useful for detecting abnormalities in the breasts by the use of a thermal picture (thermogram). Temperature irregularities may be detected using a FLIR E30 Infrared Camera thermogram. Doctors say the malignancy will not strike both breasts simultaneously. For a typical patient, the left and right breast thermograms are identical. Breast abnormalities may be diagnosed if the temperature patterns of the left and right breasts are somewhat different. With the guidance of a doctor, the patient must expose her breasts in front of a camera in order to take this test[9]. As a result, the patient is left feeling uneasy. It is a less intrusive, less costly, and less radioactive alternative to mammography. If the tumour is deeper in the breast, it does not improve outcomes. To obtain images of the breast, Breast Magnetic Resonance Imaging (MRI) employs magnets and radio waves. Screening for breast cancer in high-risk women includes both mammography and magnetic resonance imaging (MRI). One drawback of

this test is that it is inconvenient and costly for women who are at average risk. Using high-frequency sound waves, ultrasound delivers pictures of the breast to a computer monitor, where they may be examined. With the help of different screening methods, the breast tumour may be detected and removed [10].

It is common practise to employ mammography and thermography as complimentary breast cancer screening procedures, which implies that neither can be utilised alone. Tumor screening techniques such as MRI and biopsies are quite beneficial in the latter stages of the disease. The vast majority of these tests make the user uncomfortable. For various examinations, patient periodic enduring is not an option because of its ionic nature and expensive expenses. However, according to medical records, patients who have a family history of breast abnormalities should have their breasts checked on a regular basis after the age of 20 years. As a result, there is a clinical need for patients to be able to monitor and forecast breast abnormalities at an early stage using technology.

An image segmentation approach developed by the author in [11] increases the accuracy of thermography test results by segmenting pictures. Images of thermograms are segmented for further processing. Edge detection and hough transform algorithms may be used to automatically and precisely segment the thermographic breast image for asymmetry analysis of thermal patterns in the opposite breasts.

Computer-assisted breast cancer detection using cytological pictures from tiny needle biopsies was described in [12]. The biopsies are then classified as either benign or malignant based on this method. Circular high transform and support vector machine algorithms are utilised to identify and categorise observed circles in this method.

Machine learning was suggested by the authors in [13] as a way to identify breast cancer early on. Genetic mutations, chronic discomfort, size, form, and look of the breast portion are all indicators of breast abnormalities. In order to handle the Wisconsin data set, the authors suggested a supervised classification method decision tree and compared the accuracy results with other current machine learning techniques[14].

It has been shown in different research studies that the currently accessible screening modalities for breast cancer such as mammography (including thermography and ultrasound) have a limited degree accuracy. Several researchers described an advanced automated technique for detecting breast cancer at an early stage utilising current screening procedures[15]. In today's world, sensors are critical for keeping tabs on people's health. Sensors provide continuous monitoring of health indicators, making it feasible to discover health abnormalities early on and see how therapy affects the patient's health.

III. WEARABLES IN HEALTH MONITORING

Many E-health demands promise to predict and cure major illnesses by receiving and managing functional information through wearable medical sensors with wireless capabilities.

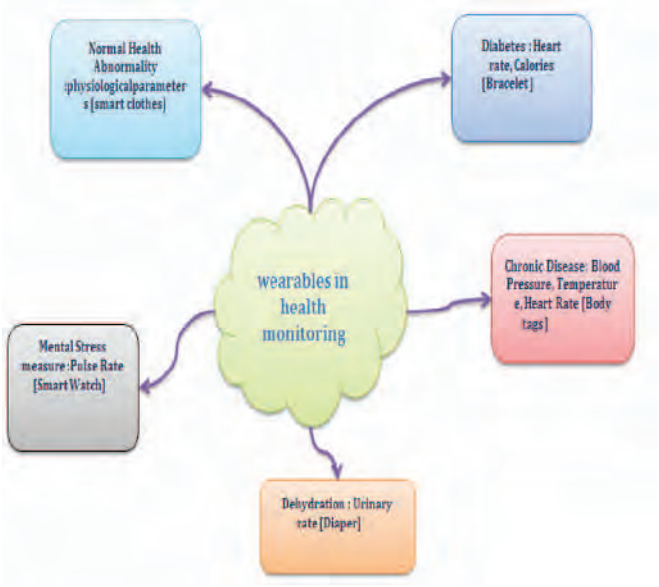


Fig.3. wearable gadgets using in health monitoring.

Sensors integrated in fabrics, decorations, etc. are used to create wearable devices for user comfort and assistance for remote health monitoring. For example, the wearable devices include everything from earphones to bracelets, which can track a user's daily activities, such as steps, calories consumed, and even how well they sleep. It is possible to connect these devices to smartphones [16].

Wearable technology is poised to play a significant role in health care as the Internet of Things develops. Sensor technology has made great strides in the field of human health monitoring, as seen in this section[17]. For many life-threatening disorders, the Internet of Things (IoT) is an emerging technology. Because of their capacity to detect, analyse, and correlate, wearable IoT is presented. Individuals will be able to monitor their own health in unprecedented detail thanks to this new technology, which will revolutionise the healthcare sector. Health care has entered a new phase of the Internet of Things (IoT) because to recent technical advancements in wireless sensors and communication methods. Healthcare IoT's primary goal is to enable individuals to live a healthy and productive life by wearing connected devices. IoT sensors often have limited processing and storage capabilities[18]. Data collected by sensors in IoT devices may be stored and processed on the cloud, and specialised programmes can be used to do so. Using a web page application, medical practitioners or other participants may have access to cloud-based health data. Remote healthcare systems might be utilised to monitor the health of patients in their own homes[19]. It improves access to healthcare in remote locations or allows elderly people to stay in their homes longer, allowing them to maintain their independence. In other words, it allows everyone to have more control over their health. Integration of IoT devices with current

medical technology and patients' environments has provided new opportunities for data collecting and enhanced decision-making in the healthcare industry[20]. Wearable IoT devices assist physicians in constantly monitoring patients' illness symptoms and help patients take the proper measures toward therapy, which helps indirectly to raise survival rates even in death case conditions. There are a few wearable devices that may be used for health monitoring as shown in Fig 3.

IV. IoT BASED HEALTH CARE SYSTEM

The Internet of Things (IoT) plays an important role in patient healthcare, including remote monitoring of patients' health, old people's pharmaceutical assistance, and early illness detection based on the data of gathered symptoms. A healthcare system based on the Internet of Things (IoT) has a number of criteria.

To create an IoT device for patient health monitoring, sensors are the most important part of the process. Wirelessly-enabled small sensors are now on the market. The first step in designing an IoT device for patient health monitoring is to choose the right sensors. Bluetooth and ZigBee are employed to transport sensor data to the coordinator system, whereas long-range communications such as LTE are used to transfer data from the coordinator system to the server in an IoT healthcare system. Hospital-based or cloud-based servers may be used for the server. It is possible to store and handle vast amounts of data with the cloud service. Algorithms for machine learning assist in the processing of data in the cloud and the generation of useful output for end users. This system has a wide range of users. Authorized professionals may easily access patient data saved in cloud storage from any location and at any time. Doctors, nurses, caretakers, governmental officials, researchers, etc. are the primary consumers of the IoT system. Fig.4 shows a conceptual structure for an IoT-based healthcare system based on an existing survey.

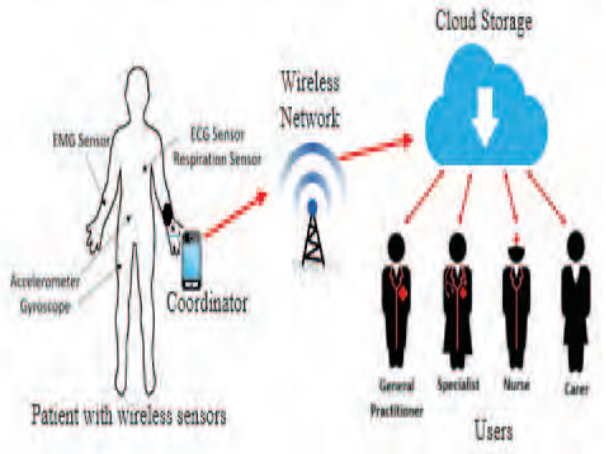


Fig.4. IoT based Healthcare System general framework

V. ARCHITECTURE OF PROPOSED SYSTEM

Recording the test results is important for subsequent illness investigation. As a consequence, cloud or hospital-based servers are ideal for storing the test results. The open-source cloud saves hospitals money by eliminating the need to purchase and install expensive server equipment. It also delivers built-in services to hospital patients.

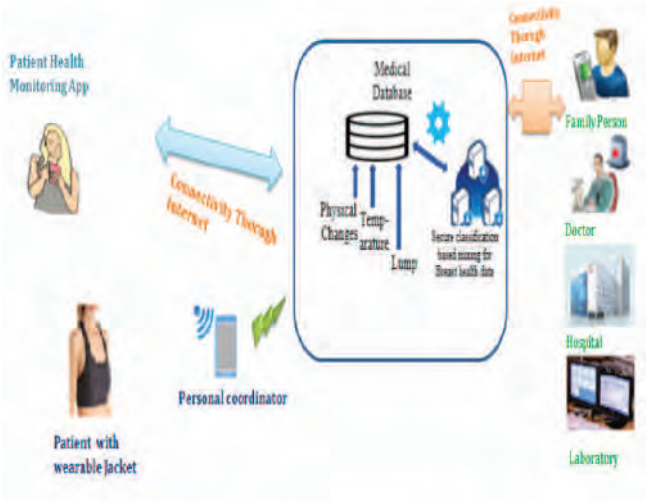


Fig.5. Architecture of proposed System

Patients' medical conditions and treatment regimens may be studied further using the data stored at this location. While SaaS provides product applications as a service to examine patient information in a distributed environment, PaAS provides the client with the ability to transmit and store the results of necessary applications in the cloud for analysis, IaAS provides virtualized data centres to include storage as an administration, all of these services were discussed in detail in [8]. These approaches allow clients to transmit and execute programmes. Fig.5 depicts the proposed framework's full design. Blue tooth or Zigbee are preferred for short-range communication between sensors in the wearable network and the coordinator. When transferring processed data from the coordinator to the server for long-term storage, it is best to use long-range connections like LTE or Wi-Fi. Patients, their loved ones, physicians, nurses, clinical labs, hospitals, and government agencies all have a stake in this system. The coordinator receives spreadsheets including the left and right breast temperature data. The dataset is initially preprocessed to remove any irregularities. In certain cases, the sensors may detect data that is irrelevant, such as excessive temperature readings or null entries, due to the effects of the environment in the test area Breasts with a positive test result are considered abnormal, whereas those with a negative test result are considered normal. This data is transferred to the server and stored there indefinitely for future use. The dataset's mean is calculated by multiplying the total number of data points by the dataset's average temperature. The term "mode" refers to a value that occurs the most times in the data collection. The median is a number that divides a collection of data into two equal portions. Difference in temperature between largest and smallest values is known as range. The sum of squared deviations between the mean and its data items divided by the total number of data items is referred to as the variance

of the dataset. The square root of variance is used to calculate the standard deviation.

VI. RESULTS AND DISUSSIONS

Using a breast temperature data set, the suggested framework is tested. In order to streamline the processing procedure, we randomly selected 150 observations from the whole dataset. The sample mean is used to replace any data with a null or irrelevant value. The acceptable left and right breast samples are subjected to the statistical measures of mean, mode, median, range, variance, and standard deviation. Data from left and right breast samples were also compared to see whether their respective maximum and minimum values differed. To look for an anomaly, all the measurements are compared between the right and left breasts. The proposed approach is validated using temperature datasets from healthy and ill patients. To make sense of the data, statistical measurements are used to the left and right breast temperatures. The statistical measurements for the normal patient are almost same for the left and right breasts. There is a significant difference in breast symmetry between the normal and atypical patients. This asymmetrical finding will aid in the early detection of breast abnormalities. Fig.6 depicts the Temperature Data Processing Algorithm and Fig.7 depicts the statistical analysis of abnormal patient breast temperature data.

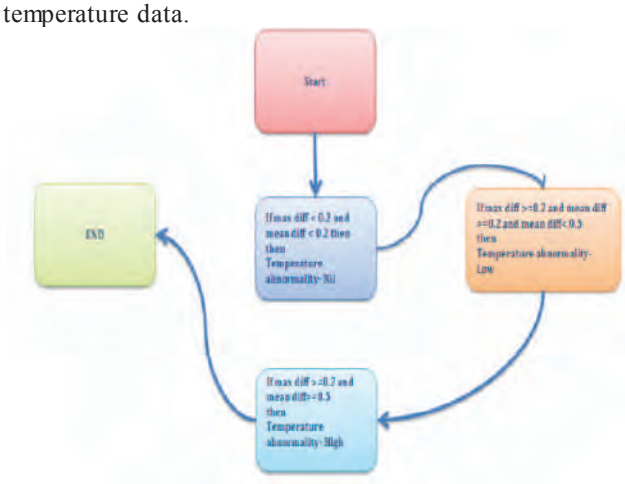


Fig.6. Algorithm of Temperature Data Processing

Table.I. Normal patient breast temperature data statistical analysis

Parameter	Variance	Standard Deviation	Range
Left Breast	<0.1	<0.2	<1.2
Right Breast	<0.1	<0.22	<1.25

Table. I and Table. II exhibit a graphical depiction of the statistical assessment of normal and abnormal patient temperature data. In the coordinator site, an analytic engine will analyse all these statistical estimations of the left vs right breast and remedy the anomaly. Final findings are sent to a server for long-term storage by the coordinator. The outcome may be seen by the system's authorised users.

Table.II. Abnormal patient breast temperature data statistical analysis

Parameter	Variance	Standard Deviation	Range
Left Breast	>0.1	>0.2	>1.5
Right Breast	>0.125	>0.18	>1.2

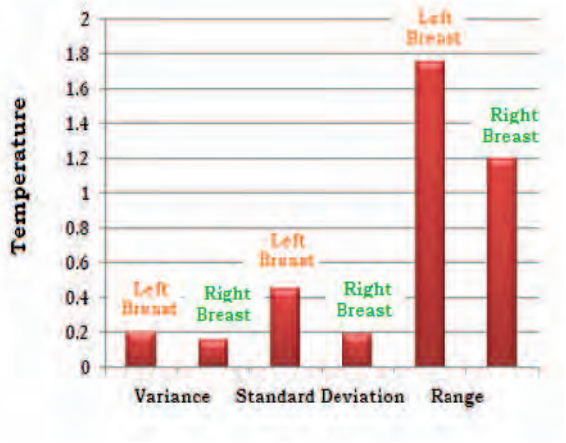


Fig.7. Abnormal patient breast temperature data statistical analysis

The proposed framework is verified by considering normal patient and abnormal patient temperature dataset. Table.III. shows the statistical elements results for accepted left and right breast data of a normal and abnormal patients.

Table.III. Normal and Abnormal breast heath patient temperature data processing using statistical highlights [21]

Statistical Feature	Normal Patient		Abnormal Patent	
	Left	Right	Left	Right
Mean	36.86	36.82	37.62	36.86
Mode	36.87	36.82	37.75	36.87
Median	36.87	36.82	37.75	36.87
Range	1.2	1.24	1.77	1.2
Variance	0.037	0.045	0.178	0.036
Standard Deviation	0.193	0.211	0.422	0.189
Maximum Difference	0.05		0.87	
Mean Difference	0.042		0.76	

VII. CONCLUSION

Designing a wearable device framework that can detect abnormal breast temperatures was explored in this study. Sensor-enabled patches on a wearable jacket are also being considered. The proposed framework's design and implementation details are laid out in depth. With the use of an aberrant and normal patient temperature dataset, the proposed framework's performance is evaluated by using graphic representations.

REFERENCES

[1]. A. Elmoufidi, "Deep Multiple Instance Learning for Automatic Breast Cancer Assessment Using Digital Mammography," in *IEEE Transactions on Instrumentation and Measurement*, vol. 71, pp. 1-13, 2022, Art no. 4503813, doi: 10.1109/TIM.2022.3177141.

[2]. A. Elmoufidi, K. El Fahssi, S. Jai-Andaloussi, A. Sekkaki, Q. Gwenole and M. Lamard, "Anomaly classification in digital mammography based on multiple-instance learning", *IET Image Process.*, vol. 12, no. 3, pp. 320-328, Mar. 2018.

[3]. S. R, S. D, U. K N and R. V M, "Arduino based COVID-19 Suspect Detection Device," 2022 6th International Conference on Electronics, Communication and Aerospace Technology, Coimbatore, India, 2022, pp. 158-163, doi: 10.1109/ICECA55336.2022.10009600.

[4]. M. R. G. and H. Anandaram, "Extraction of Fetal ECG Using ANFIS and the Undecimated-Wavelet Transform," 2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT), 2022, pp. 1-5, doi: 10.1109/GCAT55367.2022.9972078.

[5]. S. R. Kawale, S. P. Diwan and D. G. V, "Intelligent Breast Abnormality Framework for Detection and Evaluation of Breast Abnormal Parameters," 2022 International Conference on Edge Computing and Applications (ICECAA), 2022, pp. 1503-1508, doi: 10.1109/ICECAA55415.2022.9936206.

[6]. K. R. Swetha, Namitha A R, Manu Y M, Rashmi G R and Veera Sivakumar Chinamuttevi (2022), IOT Based Smart Health Care System to Monitor Covid-19 Patients. IJEER, 10(1), 36-40. DOI: 10.37391/IJEER.100105.

[7]. F. Bakshi, G. A, A. B. Naik and N. HG, "Covid-19 Prevention Kit Based on an Infrared Touchless Thermometer and Distance Detector," *2021 5th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2021, pp. 358-362, doi: 10.1109/ICECA52323.2021.9676014.

[8]. M. Ramesha, K. Jeevan and B. M. Sathisha, "Implementation of IoT Based Wireless Electronic Stethoscope," *2020 Third International Conference on Multimedia Processing, Communication & Information Technology (MPCIT)*, 2020, pp. 103-106, doi: 10.1109/MPCIT51588.2020.9350476.

[9]. J. C. Nunes, Y. Bouaoune, E. Delechelle, O. Niang and P. Bunel, "Image analysis by bidimensional empirical mode decomposition", *Image Vis. Comput.*, vol. 21, no. 12, pp. 1019-1026, 2003.

[10]. G. Naveena Pai, M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," *2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA)*, 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.

[11]. S. B. Sridhara, K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," Adv. Math. Sci. J., vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.

[12]. M. Penna, J. J. Jijesh, and Shivashankar, "Design and implementation of automatic medicine dispensing machine," in RTEICT 2017 - 2nd IEEE International Conference on Recent Trends in Electronics, Information and Communication Technology, Proceedings, 2017, vol. 2018-Janua, pp. 1962–1966, doi: 10.1109/RTEICT.2017.8256941.

[13]. S. B. Kulkarni, Y. Singh and A. A. Junnarkar, "Design of Physician Recommendation System for Patients," *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2022, pp. 1-8, doi: 10.1109/CONIT55038.2022.9847943.

[14]. S. Ciatto et al., "Second reading of screening mammograms increases cancer detection and recall rates. Results in the Florence screening programme", *J. Med. Screening*, vol. 12, no. 2, pp. 103-106, Jun. 2005.

[15]. G. Ouellec, M. Lamard, M. Cozic, G. Coatrieux and G. Cazuguel, "Multiple-instance learning for anomaly detection in digital mammography", *IEEE Trans. Med. Imag.*, vol. 35, no. 7, pp. 1604-1614, Jul. 2016.

[16]. A. Elmoufidi, K. El Fahssi, S. Jai-Andaloussi, N. Madrane and A. Sekkaki, "Detection of regions of interest's in mammograms by using local binary pattern dynamic k-means algorithm and gray level co-occurrence matrix", *Proc. Int. Conf. Next Gener. Netw. Services (NGNS)*, pp. 118-123, May 2014.

[17]. S. B. Kulkarni, Y. Singh and A. A. Junnarkar, "Automatic Physician Recommendation Framework using Robust Data Mining Techniques," *2022 2nd International Conference on Intelligent Technologies (CONIT)*, Hubli, India, 2022, pp. 1-9, doi: 10.1109/CONIT55038.2022.9847849.

[18]. A. S. Araujo, T. A. E. da Silva, M. B. H. Moran and A. Conci, "Using Series of Infrared Data and SVM for Breast Normality Evaluation," *2019 IEEE/ACS 16th International Conference on Computer Systems and Applications (AICCSA)*, 2019, pp. 1-8, doi: 10.1109/AICCSA47632.2019.9035222.

[19]. W. R. Rashmi, R. M. A. Zayed, A. R. Riya and A. I. Nahid, "Effects of Tumor Radius, Metabolic Heat Rate and Heat Transfer Coefficient on the Temperature Distribution of Tumor Affected Breast," *2021 3rd International Conference on Sustainable Technologies for Industry 4.0 (STI)*, 2021, pp. 1-4, doi: 10.1109/STI53101.2021.9732616.

[20]. M. Frize, C. Herry and N. Scales, "Processing thermal images to detect breast cancer and assess pain," *4th International IEEE EMBS Special Topic Conference on Information Technology Applications in Biomedicine, 2003.*, 2003, pp. 234-237, doi: 10.1109/ITAB.2003.1222520.

[21]. M. Gautherie, A. Kotewicz and P. Gueblez, "Accurate and objective Evaluation of Breast Thermograms: Principles and New Advances with Special Reference to an Improved Computer-Assisted Scoring System", *Proc. Int. Conf. on Thermal Assessment of Breast Health*, pp. 74-97, 1983.

[22]. A. Ghanmi and I.A. Abbas, "An analytical study on the fractional transient heating within the skin tissue during the thermal therapy", *J. Therm. BioL.*, vol. 82, no. December 2018, pp. 229-233, 2019.

2023 International Conference for Advancement in Technology (ICONAT)
Goa, India. Jan 24-26, 2023

Implementation of a Machine Learning-based Model for Cardiovascular Disease Post Exposure prophylaxis

Dankan Gowda V[1]
Department of Electronics and Communication Engineering, BMS Institute of Technology and Management,
Bangalore, Karnataka, India.
Email: dankan.v@bmsit.in

KDV Prasad[2]
Assistant Professor (Research), Symbiosis Institute of Business Management, Hyderabad, Symbiosis International (Deemed University),
Pune, India
Email: Kdv.prasad@sibmhyd.edu.in

Anil Kumar N[3]
Assistant Professor
Department of Electronics & Instrumentation Engineering
Sree Vidyanikethan Engineering College, Mohan Babu University,
Tirupati, Andhra Pradesh,India.
Email: anilkumar.n@vidyanikethan.edu

S Venkatakiran[4]
Professor, Department of Electronics & Communication Engineering
Sri Venkatesa Perumal College of Engineering and Technology.
Puttur. Andhra Pradesh,India.
Email: venkatkiran95@gmail.com

B. Ashreetha[5]
Assistant Professor
Department of Electronics & Communication Engineering
Sree Vidyanikethan Engineering College, Mohan Babu University,
Tirupati, Andhra Pradesh,India.
Email:
ashreetha.b@vidyanikethan.edu

N Sudhakar Reddy[6]
Professor, Department of Electronics & Communication Engineering, MJR College of Engineering and Technology,
Pileru, Andhra Pradesh,India.
Email: sudhakar1784@gmail.com

2023 International Conference for Advancement in Technology (ICONAT) | 978-1-6654-7517-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICONAT57137.2023.10080833

Abstract: **According to research, characteristics taken from ultrasound imaging may help with atherosclerosis diagnosis and decision-making. Atherosclerosis is estimated by parameters like elasticity, stiffness, lumen diameter, distension and IMT which can be used as an indicator of cardiovascular disease. Experienced radiologists are required to measure these parameters from the ultrasound images for the correct diagnosis. If a system could be automated to measure these parameters and thereby diagnose the CVDs, no doubt that this would be a milestone in the efforts taken to prevent cardiovascular disease. With the use of machine learning techniques, the classification of CCA anomalies from longitudinal ultrasonography B-mode images is progressed in this research. Acquisition of CCA images is done by ultrasound machine. For the segmentation of the layers of CCA, thresholding and other edge detection methods are used in spatial domain. Statistical, size and shape measurements are done from normal and abnormal images and features are identified.**

Keywords: Machine Learning, Cardiovascular, Prevention, ultrasound images and Control.

I. INTRODUCTION

The top causes of mortality and disability worldwide, according to the World Health Organization's (WHO) Global Atlas on Cardiovascular Disease Prevention and Control, are cardiovascular diseases (CVDs). Nearly 23.6 million individuals will pass away from CVDs by 2030, mostly from stroke and cardiovascular disease.

According to projections, CVD will continue to be the largest cause of mortality. In low- and medium-income nations, CVDs account for more than 80% of all fatalities worldwide. People in low- and moderate-income nations are more likely to be exposed to the risk factors for CVDs and other non-communicable illnesses. Compared to residents of high-income nations, they are less exposed to preventative initiatives. Less efficient and equitable health-care services that meet their requirements are available to them. As a consequence, many individuals in low- and moderate-income nations pass away often during their prime years of productivity[1]. Heart disease is facilitated by changes in lifestyle brought on by urbanisation, industrialisation, and globalisation. These risk factors consist of poor eating, inactivity, and cigarette use. Due to the dramatic increase in life expectancy in emerging nations, people are exposed to these risk variables for longer periods of time. The cost and duration of clinical CVD care are high. People who have CVD in their prime middle age disturb the destiny of the families. By denying key human resources at their peak years of productivity, it also jeopardises the growth of whole countries [2]. The macroeconomic impact of CVDs on low- and middle-

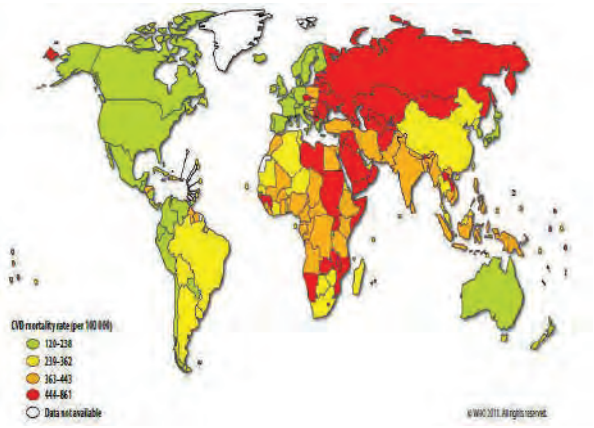


Fig. 1. Male CVD death rates throughout the world

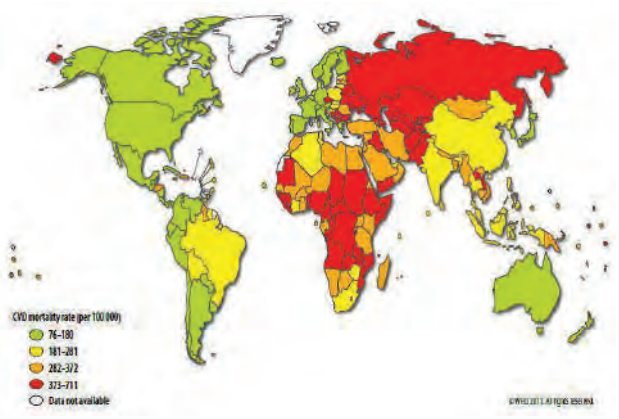


Fig. 2. Female CVD death rates throughout the world

income nations' economy is significant. According to estimates, diabetes, heart disease, and stroke would each cause low- and middle-income nations' GDP to drop anywhere from 1% to 5%. Figures 1 and Figure.2 are world maps provided by the WHO that depict the distribution of CVD death rates among men and women throughout the world.

Globally, the WHO initiative on cardiovascular problems focuses on cardiovascular disease prevention, care, and monitoring. It seeks to provide cost-effective and egalitarian health care technologies for the treatment of cardiovascular illnesses in order to build worldwide strategies to lower the incidence, morbidity, and death of cardiovascular diseases [3]. Even if several techniques are being researched for the study of CVD, an integrated system that enables automated diagnosis must yet be developed. Therefore, an attempt is made in this study to create a system that can diagnose CVDs at a reasonable cost.

II. LITERATURE SURVEY

A sample of reviews from the earlier literature is presented related to this work and discussed to highlight the importance of the proposed work. The focus is to list out the state of art research work using machine learning techniques in the field of computer aided diagnosis of cardiovascular diseases. The authors of [4] have suggested a machine learning-based segmentation approach of the carotid far wall in order to accurately and automatically quantify the Intima-Media thickness. They classified the pixels using Radial Basic Function Networks in order to identify the IMT contours. The ideally pruned-Extreme learning machine was used to solve the network learning process and choose the ideal number of radial units. A transfer learning technique to image segmentation was proposed in [5] by the authors, allowing for the supervised segmentation of pictures obtained using various imaging procedures. They have shown four transfer classifiers that, in addition to a bigger number of additional training data with somewhat different properties, can train a classification framework using only a modest amount of relevant training data.

The untapped potential of kinetic characteristics to help in the detection of carotid atherosclerosis within the context of a computer-aided diagnostics tool has been examined by the authors in [6]. Multiple cross-validation approaches were used to completely assess the motion-based CAD performance, which was likewise shown to be superior than texture-based characterisation of carotid amyloid plaques. The ability to simulate blood circulation in realistic anatomy models created from medical imaging images has advanced significantly in recent years. To comprehend how the circulatory system reacts to biomechanical stresses, it is crucial to quantify blood circulation and pressure fields using mathematical models. To forecast the possibility of plaque development and the probability of plaque breakup, knowledge of the biomechanics of blood flow is necessary [7]. A large range of potential telematics application for cardiovascular or respiratory disease may be found thanks to the development of CAD. An ultrasound image-based computational model of a carotid artery common with a plaque was created by the authors in [8]. In a case with an early atheromatous creating a minor asymmetrical stenosis, they utilised ultrasound imaging from a CCA. From a series of 2D cross-sectional photographs, they were able to recreate the 3D vasculature geometry of the sick artery segment and create computational models for the flowing and wall domains[9]. Results from this work have shed significant light on the flow and wall behaviour in slightly stenosed arteries during physiological settings, despite the fact that their model makes a number of assumptions. It is evident from the thorough literature review that diabetes, hypertensive, and atherosclerosis all influence the shape and functional activity of CCA. Particularly the IMT is suggested as a potent and standalone risk factor for atorvastatin and therefore cardiovascular disease [10]. The IMT is assessed by the separation of intima and mainstream press layers of CCA. The study of vascular disease and stroke will benefit from information of the circulation flow and pressure in healthy and sick arteries. It has been shown that the research of cardiovascular disease will benefit by understanding the elastic and mechanical characteristics of arteries. Most segmentation algorithms attempt to merely extract the image's borders and outlines. Only a few articles have successfully extracted boundaries perfectly. Multi wavelets are found to have a variety of beneficial characteristics that provide significant promise for their usage in image processing applications. Most of the research listed above is done in the spatial domain, and very little is done in the transform domain employing multiple wavelets [11]. Machine learning is applied in the proposed study using both spatial and transform domain information. According to research, characteristics taken from ultrasound imaging may help with atherosclerosis diagnosis and decision-making. Indicators of cardiovascular disease such as elasticity, stiffness, lumen diameter, distension, and IMT are used to assess atherosclerosis. For the right diagnosis, experienced radiologists must measure these factors from the ultrasound pictures [12]. Undoubtedly, it would be a significant step forward in the fight against cardiovascular disease if a system could be developed that could automatically monitor these indicators and, in turn, detect CVDs. With the use of machine learning techniques, the classification of CCA anomalies from longitudinal ultrasonography B-mode pictures is progressed in this study. The CCA pictures are analysed in this suggested study in both the spatial and transform domains. The near-

far wall, lumen, and intima-media areas of the CCA are segmented in the spatial domain. Lumen Diameter and IMT are assessed for both normal and pathological pictures after segmentation [13]. The categorization of normal and diseased arteries is then done using these characteristics to create classifiers. Multiwavelets are used for CCA image analysis in the transform domain. The energy of each band is calculated. Following that, classifiers are created using these sub band energies to diagnose faulty arteries.

III. ACQUISITION OF IMAGES AND DATABASE

In this investigation, 250 normal and abnormal patients were utilised. The ultrasound device records the artery motions. The values of the ultrasonic scanner settings are a significant problem that affects the accuracy of recording picture sequences. They may have an impact on how anatomical structures look, which would then have an impact on how the photos should be interpreted.

The term "spatial domain" refers to the actual picture plane, and image processing techniques in this area are based on direct pixel manipulation. The ability to quickly see spatial processes is their greatest benefit. The most significant obstacle in computer assisted clinical applications is carotid artery wall segmentation [14]. The adventitia layer shows brilliant in B-Mode longitudinal pictures, whereas the layer mainly is dark grey. Typically, it is impossible to tell the intima layer apart from the media layer. Using the Aphelion imaging software package, the following processes are used to segment and identify the boundaries of near-far wall, lumen, and intima-media areas. The colour picture is first transformed into a grey image. It is customary to employ a smoothing filter to reduce excessive picture noise. A Gaussian filter is used here. In two dimensions, Gaussian filters are logically symmetric. The filter does the same amount of smoothing in every direction [15]. A Gaussian filter won't prejudice future edge detection in any specific way since it is logically symmetric. The size of the neighbourhood, which is defined by the number of pixels, is calculated by this filter as a weighted average of the pixel values. The border of the lumen and other layers of the CCA are segmented using the thresholding approach. The simplest and most used segmentation technique is global thresholding [16]. On the basis of the threshold value, it is the process of extracting the interest-bearing areas. This straightforward approach performs well when there is a discernible trough between the histogram modes corresponding to objects and background. This can be seen in Figure 3. The region in the profile that contains the boundary's pixel values is darkened.

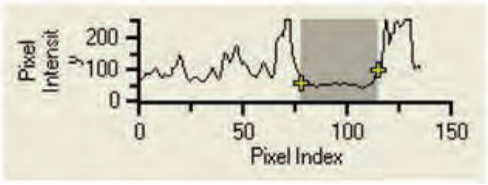


Fig. 3. Intensity profile

Given that the lumen region is dark, principal component analysis is used to find dark objects with pixel values between 0 and 70. The thresholding range is retained at 0 to 140 to segment the area between near and distant walls, and it is 70 to 140 to segment the intima-media region. The segmented and filtered picture is shown in Figure 4.

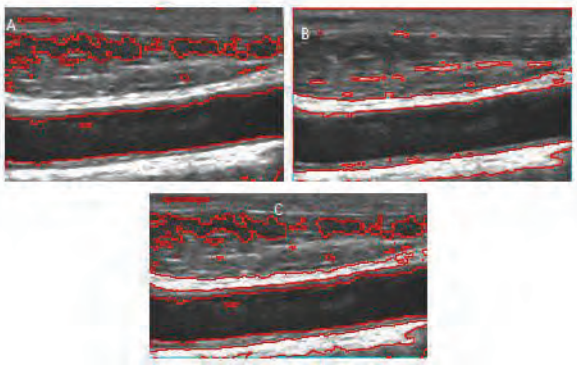


Fig. 4. Filtered and segmented image (A) Lumen (B) NW-FW (C) IMT

IV. COMPARISON OF EDGE BASED SEGMENTATION

Boundaries of intima, media and adventitia layers are well separated using region based segmentation technique, thresholding. Some edge based algorithms like Roberts, Prewitts, Sobel and Canny edge detection techniques are also applied to segment the layers of CCA using Aphelion imaging software suite to select the best [17]. Sobel and Prewitt are edge detection algorithms which help in emphasizing the edge of an image.

Both operators calculate the gradient approximations of the image intensity function and are discrete differential operators. The purpose of the operators is to do a reasonably simple and low-cost horizontal and vertical convolution of the picture with a tiny, separable, integer-valued filter [18]. Prewitt is easier to implement than Sobel, although it is more noise-sensitive. Prewitt is more sensitive to vertical and horizontal edges whereas Lbp operator is more responsive to diagonal edges.

Canny operator utilizes non-maxima suppression and hysteresis thresholding in order to detect weak and stronger edges and produce a 1 pixel thick edge. The gradient size of Roberts operator represents the strength of the edge. Roberts operator gives the information of the edges that are vertical to the direction of the gradient. The result of applying these edge detection operators is shown in the Figure.5.

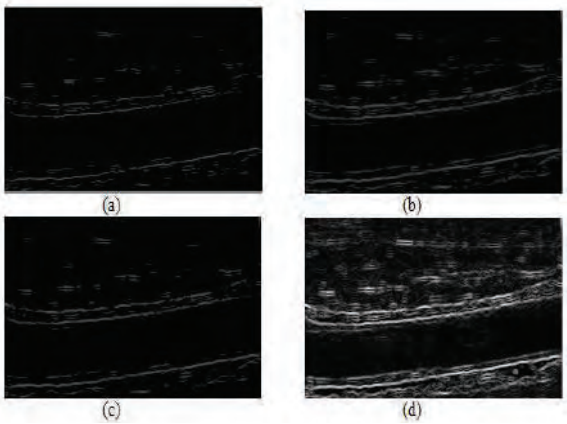


Fig. 5. Segmentation using (a) Roberts (b) Prewitts (c) Sobel (d) Canny edge detection techniques

Segmentation results obtained through edge detection techniques also prove the same. Thus, it can be said that thresholding, when used with the AphelionTM imaging software package, provides the greatest performance for the delineation of lumen and intestinal mucosa areas from ultrasound B-mode pictures of CCA.

V. MATHEMATICAL MODELLING

Researchers have been able to study, create, and solve extremely accurate mathematical models that are capable of providing insight into the physical phenomena occurring in the human cardiovascular system thanks to recent advancements in computational modelling and mathematical simulation techniques, as well as the growing compute capabilities offered by personal computers.

Vascular diseases are the main factors of mortality, and atherosclerosis is the most common vascular disease. A study of vascular disease and, in particular, of atherosclerosis impact on haemodynamics may be based on mathematical models and numerical simulation.

Insight into healthy and unhealthy blood artery conditions is provided by computer modelling of blood flow in the pulmonary circulation, which can have uses in the design of medical devices and surgical planning.

The variations in the equivalent diameter shown in Figure.6. are with respect to frame number. This can also be represented in terms of cardiac cycles. If a polynomial is derived from the known values of attributes, it is possible to estimate the missing as well as future data. Hence, a mathematical model is developed to derive a polynomial for equivalent diameter of CCA as a function of cardiac cycle time. Though the frames are taken only for particular time intervals, equivalent diameter at any time can be computed using the derived polynomial.

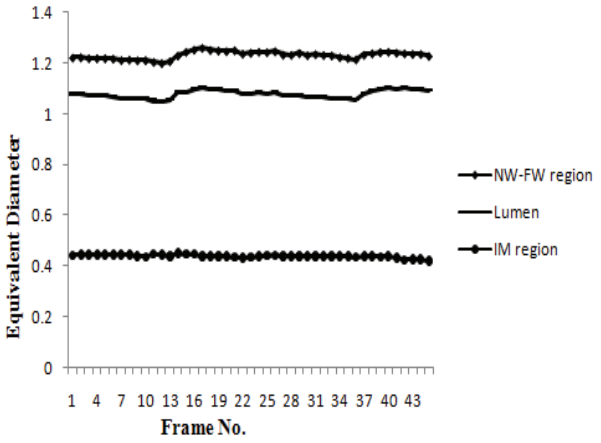


Fig. 6. Variation in equivalent diameter

Interpolation is the technique of calculating function intermediate values from a collection of function values. The process of extrapolation involves locating values outside of the interval. But generally speaking, both techniques utilise the term interpolation. There are many ways to locate these interpolating polynomials. In this work, three interpolated techniques—forward Newton's interpolation formula, Newton's split difference formula, and Lagrange's interpolation formula are employed to generate polynomials. If y = f(x) denotes a function which takes the values y0, y1,….yn corresponding to the values x0, x1,....xn respectively of x.

VI. RESULTS AND DISCUSSION

Results of the various methods applied in the spatial domain for the analysis of CCA images and performance of the classifiers using spatial domain features are discussed. On 250 typical image frames, the suggested thresholding segmentation technique is used. The boundaries of near-far wall, lumen and intimamedia regions are detected. In general thresholding needs manual intervention. But it works well in AphelionTM imaging software suite. 60 frames are enough to analyse the vascular characteristics of CCA in two cardiac cycles since the average cardiac cycle lasts 0.8 seconds and may vary from 0.75 to 1 second. Due to its superior visibility in B-mode pictures, measuring of IMT is often restricted to the far-end wall. IMT is thus segmented only in the far-end wall. Segmented and detected boundaries of near-far wall, lumen and intima-media regions for a sample of 3 frames of normal CCA are shown in the Figure.7. Since intima –media region does not expand or contract during the cardiac cycle, numerical attributes measured for the lumen region alone are plotted and analyzed for the analysis of elasticity.

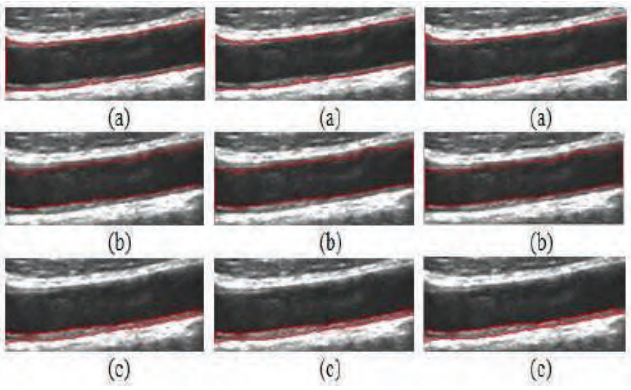


Fig. 7. Segmented (a) NW-FW (b) lumen (c) IMT regions for a sample of 3 frames of normal CCA

Analysis of elasticity of CCA is done only to validate and check whether the measured attributes follow the behaviour of CCA so that they can be used as features for the classifier. The numerical characteristics derived from the identified lumen border and other places precisely agree with the function of CCA during a cardiac cycle, as shown by measurements and attribute graphs. Hence the measured values of Lumen Diameter (LD) and IMT using the proposed segmentation algorithm are used as the features to characterize the CCA and to classify normal and abnormal CCA. The primary sign of atherosclerosis is thickening of the carotid intima layer. When the thickening is limited to a small arterial segment and lowers the lumen diameter, it is referred to as plaque. Since measuring IMT allows for the detection and evaluation of plaque, the suggested segmentation technique is used to separate the intima-media area from anomalous CCA pictures. For a sample of aberrant CCA pictures, segmented images and recovered intima-media areas with plaque are shown in Figure 8.

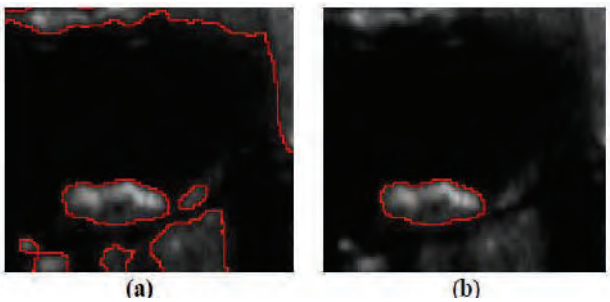


Fig. 8. Sample 1 of four (a) Segmented abnormal CCA images, (b) Extracted intima-media regions with plaque

For spatial domain features, the lowest accuracy is given by RBFN. But for transform domain features, MLP gives the lowest accuracy when compared to other classifiers. But, MLP gives 100 % sensitivity. Error rate is very high for MLP, when compared with other classifiers. SVM gives 100% specificity and precision. RBFN offers 100% sensitivity. LVQ also gives 100% specificity and precision. The value of Euclidean distance for all the classifiers is very close to the perfect classifier. Thus the proposed classifiers provide better classification of normal and abnormal CCA images using the diagonal energy coefficients obtained from the GHM and HM multiwavelet decomposition as the transform domain features.

VII. CONCLUSION

The border of the near-far wall, lumen, and intima-media area are segmented, and a technique is presented to estimate the numerical properties from transverse B-mode ultrasound images. By using the thresholding approach, the borders of the lumen, intima-media, and near-far wall areas are recognised and segmented. Measured values of IMT are used to identify the progression of atherosclerosis since IMT increases due to the presence of plaque. Validation of the measured attributes is done by the analysis of elasticity. Proposed segmentation algorithm precisely extracts the plaque area when it is applied on abnormal CCA images. The measured values of Lumen Diameter and IMT using the proposed segmentation algorithm are used as the spatial domain features to characterize the CCA and to classify normal and abnormal CCA. All the proposed classifiers give more than 90% accuracy for both spatial domain features and transform domain features. Error rate of all the classifiers is very low. Measured value of Euclidean distance for all the classifiers is very close to the perfect classifier.

REFRENCES

[1] Baldewsing, RA, Danilouchkine, MG, Mastik, F, Schaar, JA, Serruys, PW & van der Steen, AF 2018, An inverse method for imaging the local elasticity of atherosclerotic coronary plaques‘, IEEE Transactions on Information Technology in Biomedicine, vol. 12, no. 3, pp. 277-289.

[2] Destrempes, F, Meunier, J, Giroux, MF, Soulez, G & Cloutier, G 2019, Segmentation in ultrasonic B-mode images of healthy carotid arteries using mixtures of Nakagami distributions and stochastic optimization‘, IEEE Transactions on Medical Imaging, vol. 28, no. 2, pp. 215-229.

[3] Gastounioti, A, Makrodimitris, S, Golemati, S, Kadoglou, N, Liapis, C, & Nikita, K 2015, A novel computerized tool to stratify risk in carotid atherosclerosis using kinematic features of the arterial wall‘,IEEE Journal of Biomedical and Health Informatics, vol. 19, no. 3, pp. 1137-1145.

[4] Hasegawa, H & Kanai, H 2016, Modification of the phased-tracking method for reduction of artifacts in estimated artery wall deformation‘, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 53, no. 11, pp. 2050-2064.

[5] Sharma, R, M. R. G. and H. Anandaram, "Extraction of Fetal ECG Using ANFIS and the Undecimated-Wavelet Transform," *2022 IEEE 3rd Global Conference for Advancement in Technology (GCAT)*, 2022, pp. 1-5.

[6] Hernandez-Sabate, A, Gil, D, Fernandez-Nofrerias, E, Radeva, P & Marti, E 2009, Approaching artery rigid dynamics in IVUS‘, IEEE Transactions on Medical Imaging, vol. 28, no. 11, pp. 1670-1680.

[7] Hossain Md Murad, AlMuhanna Khalid, Zhao Limin, LalBrajesh, K & Sikdar Siddhartha 2015, Semiautomatic segmentation of atherosclerotic carotid artery wall volume using 3D ultrasound imaging‘, vol. 42, no. 4, pp. 2029–2043.

[8] Ranjeet Suryawanshi, Revanna C R, B. Kameswara Rao and Parismita Sarma (2022), Enhanced Diagnostic Methods for Identifying Anomalies in Imaging of Skin Lesions. IJEER 10(4), 1077-1085. DOI: 10.37391/IJEER.100452.

[9] Jayanthi, KB & Banu, RSD 2019, Neural network based classifier for cardio vascular diseases based on vascular aging‘. Proceedings of the IEEE Annual International Conference on Engineering in Medicine and Biology Society, pp. 1109-1112.

[10] A. P, A. Sharma, S. R. Kawale, S. P. Diwan and D. G. V, "Intelligent Breast Abnormality Framework for Detection and Evaluation of Breast Abnormal Parameters," *2022 International Conference on Edge Computing and Applications (ICECAA)*, 2022, pp. 1503-1508.

[11] Larsson, M, Kremer, F, Claus, P, Kuznetsova, T, Brodin, LA & D'hooge, J 2020, Ultrasound-based radial and longitudinal strain estimation of the carotid artery: a feasibility study‘, IEEE transactions on ultrasonics, ferroelectrics, and frequency control, vol. 58, no. 10, pp. 2244-2251.

[12] Avinash Sharma, Rajesh L, Mirzanur Rahman, Ghazaala Yasmin, Parismita Sarma, A. Azhagu Jaisudhan Pazhani, A novel method of data compression using ROI for biomedical 2D images, Measurement: Sensors, Volume 24, 2022, 100439, ISSN 2665-9174.

[13] Ley, O & Deshpande, CV 2009, Comparison of two mathematical models for the study of vascular reactivity‘, Journal of Computers in biology and medicine, vol. 39, no. 7, pp. 579-589.

[14] Loizou, CP, Pattichis, CS, Nicolaides, AN, & Pantziaris, M 2019, Manual and automated media and intima thickness measurements of the common carotid artery‘, IEEE Transactions on Ultrasonics, Ferroelectrics and Frequency Control, vol. 56, no. 5, pp. 983-994.

[15] Avinash Sharma, M. Nagabushanam, H. G. Govardhana Reddy & K. Raghavendra (2022) Vector space modelling-based intelligent binary image encryption for secure communication, Journal of Discrete Mathematical Sciences and Cryptography, 25:4, 1157-1171.

[16] Molinari, F, Zeng, G & Suri, JS 2018, A state of the art review on intimamedia thickness (IMT) measurement and wall segmentation techniques for carotid ultrasound‘, Journal of Computer methods and programs in biomedicine, vol. 100, no. 3, pp. 201-221.

[17] Pai, G.N., Sridhara, S.B., Shashidhara, K.S., Gangadhara, “Signal Analysis and Filtering using one Dimensional Hilbert Transform,” Journal of Physics: Conference Series 1706(1),2020, https://doi.org/10.1088/1742-6596/1706/1/012107.

[18] Opbroek, AV, Ikram, MA, Meike W Vernooij & Bruijne, MD 2015, Transfer Learning Improves Supervised Image Segmentation Across Imaging Protocols,‘ IEEE Transactions on Medical Imaging, vol. 34, no. 5, pp. 1018-1030.

2023 International Conference for Advancement in Technology (ICONAT) | 978-1-6654-7517-4/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICONAT57137.2023.10080332

Technologies for Comprehensive Information Security in the IoT

Dankan Gowda V[1]
Department of Electronics and Communication Engineering, BMS Institute of Technology and Management,
Bangalore, Karnataka, India.
Email: dankan.v@bmsit.in

Sheetalrani R Kawale[2]
Assistant Professor, Department of Computer Science, Karnataka State Akkamahadevi Women's University,
Vijayapura, Karnataka, India.
Email: sheetalrkawale@gmail.com

KDV Prasad[3]
Assistant Professor (Research), Symbiosis Institute of Business Management, Hyderabad, Symbiosis International (Deemed University),
Pune, India
Email: Kdv.prasad@sibmhyd.edu.in

Anil Kumar N[4]
Assistant Professor
Department of Electronics & Instrumentation Engineering
Sree Vidyanikethan Engineering College, Mohan Babu University,
Tirupati, Andhra Pradesh,India.
Email: anilkumar.n@vidyanikethan.edu

N Sudhakar Reddy[5]
Professor, Department of Electronics & Communication Engineering, MJR College of Engineering and Technology,
Pileru, Andhra Pradesh,India.
Email: sudhakar1784@gmail.com

B. Ashreetha[6]
Assistant Professor
Department of Electronics & Communication Engineering
Sree Vidyanikethan Engineering College, Mohan Babu University,
Tirupati, Andhra Pradesh,India.
Email: ashreetha.b@vidyanikethan.edu

Abstract: **Due to the difficulty and significance of the Internet of Things (IoT), the highest standards of data protection are essential to the successful growth of any smart city or technical endeavour that relies on the IoT. This paper focuses on the efficient security management techniques for IoT that can be applied in smart city, smart homes, industry and many applications. Initially, an introduction to IoT and Blockchain as security methods are discussed, as well as the sensors and wireless sensor networks associated with IoT are also introduced. A comprehensive survey on IoT and its security issues are presented. IoT concerns about privacy and security are a necessary procedure as well as a difficult challenge for researchers. The blockchain technology enhanced and motivated the recent security parameters and it has been validating various technical sectors since its inception.**

Keywords: IoT, Security, Information, Smart city, Privacy issues and Block chain.

I. INTRODUCTION

The IoT is a network of everyday physical objects, household appliances, automobiles, gadgets, and many embedded electronics items that enable data transmission among these things in order to allow for a high level of integration between the physical world and digital system [1]. As a result, the economy gains, efficiency increases, and the need for human labour drops. As shown in Figure.1, the Internet of Things encompasses an expanded network that extends beyond conventional computing devices to include a wide variety of physical gadgets and ordinary things that are not typically online [2]. Devices that are part of the Internet of Things (IoT) are equipped with technology that allows them to be remotely supervised and controlled and to exchange data with one another through the Internet [3]. IoT represents a fastest flourishing technology paradigm containing plenty of devices that produce, process and communicate large amount of data. Therefore, several types of intrusions and threats can target IoT devices in the network [4]. Many lightweight IoT devices have low energy capacity and it is consumed by its applications processing and core functionalities, ultimately, the limited power becomes the challenging issue for security and privacy [5]. With reference to sensing, processing and communication, the traditional security methods and protocols are in applicable to low powered IoT devices and networks. Due to the distributed characteristics, scalability and single point failure of the IoT network most centralized based security frameworks are not appropriate for IoT. Therefore, distributed, scalable and lightweight security solutions are required in IoT network [6]. To overcome these shortcomings, blockchain based distributed record keeping technology is required. Unpredictable behaviour can be observed in the network traffic due to intrusions and attacks, so security is the primary step for any network deployment [7].The blockchain technology attracted many researchers, government, financial, industrial, real estate, healthcare sectors across the world. Without any central authentication, the high level of security can be provided using blockchain technology as a decentralized manner [8]. Currently, there exists many blockchain and applications based on IoT, such as agriculture, logistics, smart homes, shared homes, smart cities, Internet of smart trains and energy sectors as presented in Figure.2. In the blockchain based security network, the distributed ledger blocks concepts are used to exchange data among users with worrying about to trust each other [9].

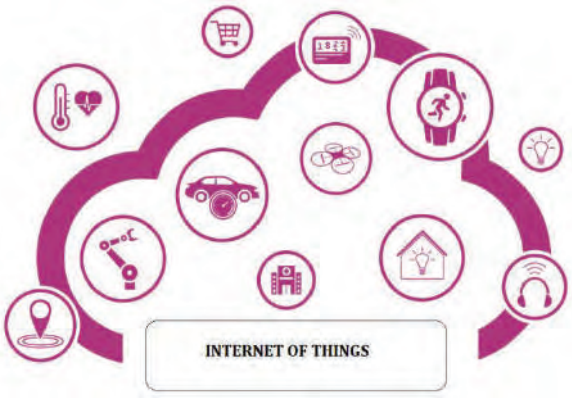


Fig. 1. Internet of Things

Without any third party trusted entity, the cooperating transactions among the users can be very fast. To maintain all secure, authenticate and trusted interactions in the network, cryptographic smart-contracts is used by the blockchain technology. The blockchain has become futuristic sustainable infrastructure for many modern social innovations in the IoT world. The collaborative co-operativity of Blockchain with Internet of Things presents an promising conceptualization which can communicate various challenging issues in Figure.2. Integrated Blockchain and IoT based Applications the research of IoT. Particularly, the IoT-blockchain integrated technology can raise various advanced technological prospects [10].The operating methods of Internet of Things can be modified by blockchain as a distributed ledger technique, it provides new concepts of networking and stable framework for social initiation and creativeness.

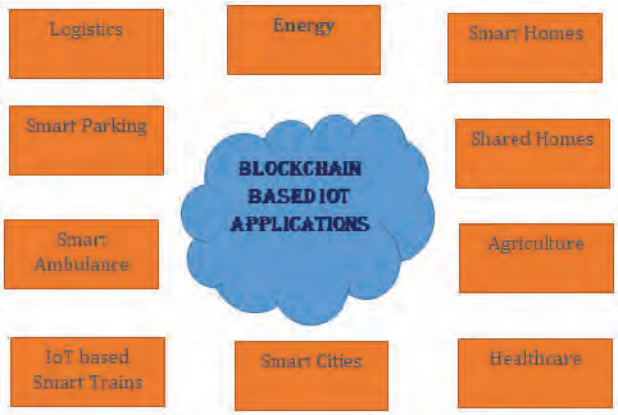


Fig. 2. Integrated Blockchain and IoT based Applications

The device spoofing and deception can be avoided by IoT-blockchain integration that will enable the secure mesh networkfor reliable communication and interaction among active devices in IoT network [11].Without requiring any third party certification authority, gateway or central broker, each legitimate registered IoT device under blockchain can easily authenticate and identify each other in the network. Furthermore, without any external agentand management, the IoT network can be expanded to support plenty of new devices.

II. LITERATURE SURVEY

In recent years, researchers have been addressing the difficulties of integrating blockchain with IoT network. To authorize new blocks, In [12] author's proposed a lightweight technique for secure devices using attest or observer from blockchain based IoT network. Without complex calculations, the IoT devices can validate and process using Bloom filters. Their experimental simulation shows that the blockchain security can be achieved on lowpower devices.

In [13] proposed infrastructure to generate a local peer-to-peer network for allowing Blockchain ledger to expand across all nodes within a network. The experimental outcomes acquired from implementation and testing represents remarkable improvement in TPS and ledger degree is brought out. For broad level of IoT business transactions, this framework assign better scalability and addresses the issues of memory demand for blocks storage. SMER framework consist of resource exchange layer exchange rules layer and financial infrastructure layer. The resource exchange layer provides the management of device module and resource module. The exchange rules layer manages smart contract module and contract execution module. Finally, the financial infrastructure layer provides encrypted coin module and device account management. It also manages resource exchange issues and responsible for producing and handling accounts for each IoT device. This layer is also responsible for account security, transfer and transaction records.

In this part, you'll learn about cutting-edge techniques for sniffing out security gaps and blocking attacks in IoT networks [14].In order to create a blockchain-based security model, they examined a wide variety of existing countermeasures and weighed their relative merits and shortcomings. Techniques such as data-control revision, protocol revision at the network layer, and machine-learning application are used to counteract these threats. They also presented a blockchain-based safe architecture to avoid privacy and security problems in the Healthcare IoT. (H-IoT) [15]. Their technique was put to the test in the Internet of Things-Fog network test bed. The testing model located inside the fog node network may discover infiltration and attack more quickly than under cloud-located conditions. This semi-supervised security technique may also be used to manage the challenge of categorising large datasets. In [16], the authors provide a formulation that helps with DDoS traffic identification and countermeasures. They looked at four distinct classifier types using a predetermined dataset and achieved accuracy. To protect the victim's website hosted on Apache, they built a framework in Linux on a Raspberry Pi and examined the data packets using Wire shark. With the goal of detecting intrusions in IoT networks, the authors of [17] centred their design efforts on creating a lightweight IDS (IDS). Down this case, they zeroed in on a typical kind of DDoS assault. The first is the data characteristics attained for signal categorization, and the second is the classifier based on ML, both of which are at the heart of this intrusion detection system (IDS). Within this framework, the node's packet-receiving speed is an important metric to monitor.

IoT device recognition, network traffic classification, and intrusion detection are all covered in [18], where the authors propose a framework. They presented AI in this context to overcome the security and complexity of QoS management systems. In order to evaluate different machine learning strategies, this framework runs a series of experiments using data acquired from the IoT network, including both normal and attack traffic.

The conventional decentralized techniques face the issue of network congestion and single-point failure. To overcome these shortcomings, Stochastic based blockchain method is proposed for verifying data integrity in IoT network [19].The proposed method minimizes the quantity of cooperating IoT devices and trust onedge devices to produce the block, which reduces the computational and transmission time. The simulation outcomes present an enhanced success rate against large-area IoT networks with a low amount of cooperating devices.RSA based

cryptography is applied in communication links between gateways and clients of the mesh network with existing malicious devices[20].The existing data routing method works for static mesh devices and monitors transmission links, which can create a deficient effect on the performance of the network and increases the chance of packet drop ratio. Their proposed Infrastructure constitutes a mesh network of mobile clients to perform better network-exposure in data transmission lines considering factors for packet drop reduction and a high ratio of data delivery[21]. This model defeats the communication cost using the flooding of distance vector routing over periodic time intervals by mobile mesh clients. Their simulation outcomes present high data reliability as well as a low computational operating expense in different topological networks.

III. IOT SECURITY ISSUES AND POTENTIAL SOLUTIONS

Defining the IoT's underlying architecture is a contentious topic, with competing definitions coming from writers using a wide variety of terms. At now, the IoT is supported by three main levels: the perception layer, the network layer, and the application server (framework might be inserted between the final two layers). These partitions provide for modularity and facilitate the efficient growth of systems [22]. Unfortunately, this opens the door for bad actors—in this case, external attackers acting inside the framework of the stated threat model—to take advantage of flaws in the design of each layer of the Internet of Things. Since the many levels of the IoT are built on top of diverse technologies, each has its own set of strengths and drawbacks that are function- and use-specific. Exploitation of these flaws has led to the compromise of millions of IoT devices, providing hackers with the tools they needed to carry out one of the most widespread outages in the history of the Internet. The inherent basic concept of energy efficiency and the limited computational power available on portable systems are in some way antagonism to the existing cryptographic algorithms principles, creating a more difficult environment for the Internet of Things and its community despite the concerns of security researchers.

IV. PROPOSED SYSTEM

Due to the various drawbacks of centralised security approaches, this section presents a blockchain-based decentralised security system for the IoT-based network. High integrity, no participation of third party protection, secure peer-to-peer identification and connections, and more are just a few of the ways in which blockchain technology outperforms more traditional defensive systems.

Figure 3 depicts the proposed blockchain-based Internet of Things network for secure communication and protection against hacking and intrusion. Every Internet of Things gadget has its own block in the network, complete with information, hash, and date. Each device's block data is shown as a blockchain in Figure.4.

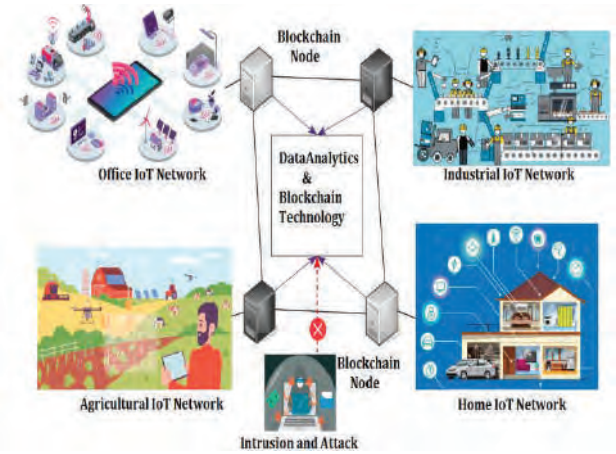


Fig. 3. IoT Blockchain Proposed Network

Fig. 4. Each device's information as a blockchain

The unauthorized activities can modify or delete the original data stored on IoT, therefore, for transaction record, blockchain based identity credential scheme for each device in IoT must be registered during device manufacturing and deployment for use. By using smart contract, accessing protocols for each IoT device should be assigned and enforced .Consequently, unauthorized performance and activities on stored data also be removed in the IoT network. When blockchain technology assures data integrity and authenticity, also prohibit unauthorized access, then third party centralized authority like cloud service provider is not required for stored data safety at different network locations in IoT. There are several benefits of integration and adoption of IoT with blockchain as presented in Figure.5. The benefits includes publicity, decentralization, security, resiliency, cost saving, speed, anonymity among users and immutability. Message transmissions among IoT devices are considered as financial transactions of Bitcoin network. Message transmissions among IoT devices are authorized by smart contracts, that provides the agreement between two entities. The basic potentiality of blockchain is completely decentralized, secure ledger and database of all activities and transactions within a network is maintained.

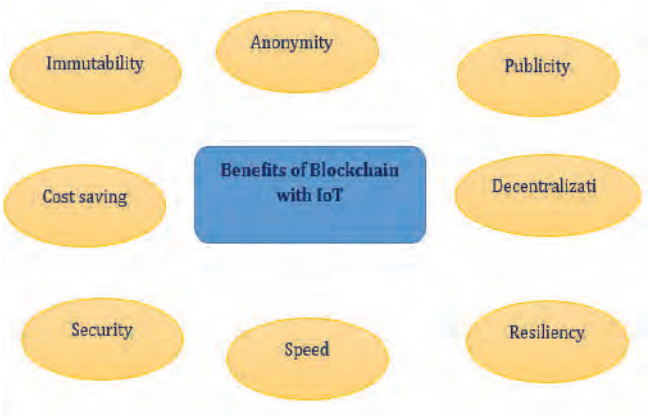


Fig. 5. Benefits of IoT-Blockchain Integration

V. RESULTS AND DISCUSSION

Integrity was verified with the use of blockchain technology and the widely-used message digest 5 and secure hash algorithm 1. Integrity validation simulation with MD-5, blockchain and SHA-1, based on degree centrality is shown in Figure.6. If you look at the graph, you'll see that blockchain (the blue line) is far closer to the baseline than MD5 or SHA-1, indicating that it has a very high level of integrity. Figure.7 shows another another model of integrity validation based on MD5, blockchain's and SHA-1. When comparison to MD5 and SHA-1, blockchain also displays good integrity since it is closer to the baseline. Data such as network flows, sensor readings, log entries, and transaction histories are created by IoT devices and may be verified using cryptographic hash methods. The findings indicate that the performance offered by blockchain integrity validation is much higher than that offered by MD5 and SHA-1. The immediacy with which every change or update to the ledger is mirrored throughout the network is made possible by the decentralised and distributed structure of the blockchain.after a network device has been compromised, the blockchain may be used to validate and update the compromised data, restoring the device's original state. Cryptographic transaction ledgers like blockchain were originally developed to ensure the safety of Internet of Things connections.

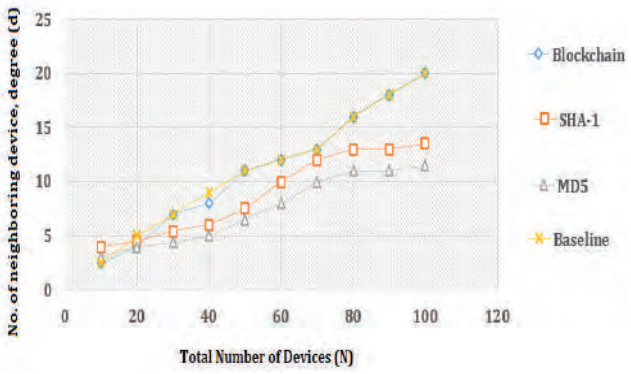


Fig. 6. Integrity Validation-1.

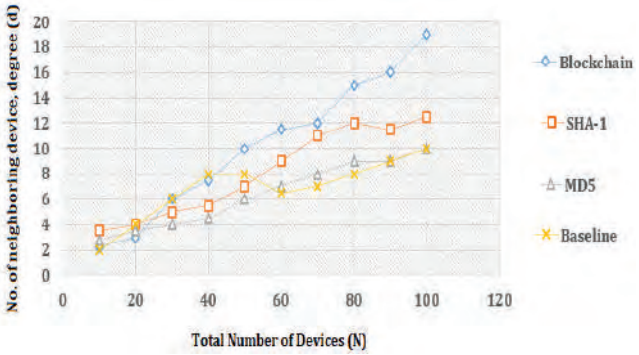


Fig. 7. Integrity Validation-2

VI. CONCLUSION

The model presents, modern blockchain technique surpassthe traditional security system in efficient and secure IoT deployment at various conditions. Finally, to validate the signification of blockchain in IoT network, the avalanche effect is calculated and compared with Triple-DES, AES and Blowfish cryptographic algorithms. The result presents significant output changes in hash for blockchain IoT integrated model as compared to other cryptographic algorithms. Blockchain and the IoT are two rapidly developing and expanding technologies. Blockchain technology provides a robust decentralised security solution that can be linked into IoT-based networks to protect them from a wide range of assaults, threats, and intrusions. This paper explores effective security management strategies for the IoT, with applications ranging from smart cities and homes to factories and offices. The IoT and Blockchain are presented, together with the accompanying sensors and wireless sensor networks, and addressed in terms of their security. An in-depth analysis of IoT security concerns is provided.

REFRENCES

[1] R. Mahmoud, T. Yousuf, F. Aloul and I. Zualkernan, "Internet of things (IoT) security: Current status, challenges and prospective measures," 2015 10th International Conference for Internet Technology and Secured Transactions (ICITST), 2015, pp. 336-341, doi: 10.1109/ICITST.2015.7412116.

[2] M. Jun, "Blockchain government - a next form of infrastructure for the twenty-first century,"Journal of Open Innovation: Technology, Market, and Complexity, vol. 4, no. 1, p. 7, Feb2018.

[3] A. Sharma, Y.-H. Hung, P. Agarwal, and M. Kalra, "A study on blockchain applications inhealthcare," in Frontier Computing, J. C. Hung, N. Y. Yen, and L. Hui, Eds. Singapore:Springer Singapore, 2019, pp. 623–628.

[4] A. Dorri, S. S. Kanhere, and R. Jurdak, "Towards an optimized blockchain for IoT," in 2017IEEE/ACM Second International Conference on Internet-of-Things Design and Implementation(IoTDI), April 2017, pp. 173–178.

[5] X. Duan and X. Wang, "Authentication handover and privacy protection in 5G hetnets using software-defined networking", Communications Magazine, vol. 53, pp. 28-35, 2015.

[6] K. Haseeb, I. Ud Din, A. Almogren, N. Islam, and A. Altameem, "RTS: A robust and trustedscheme for IoT-based mobile wireless mesh networks," IEEE Access, vol. 8, pp. 68 379–68 390,2020.

[7] Priyadarshani Shivkumar Mali, Hemant. A. Tirmare, Varsha Amol Suryawanshi and Abhay Chaturvedi (2022), Novel Predictive Control and Monitoring System based on IoT for Evaluating Industrial Safety Measures. IJEER 10(4), 1050-1057. DOI: 10.37391/IJEER.100448.

[8] A. P, A. Sharma, S. Reddy P, P. S. Patwal, "Data Analytics and Cloud-Based Platform for Internet of Things Applications in Smart Cities," 2022 International Conference on Industry 4.0 Technology (I4Tech), 2022, pp. 1-6, doi: 10.1109/I4Tech55392.2022.9952780.

[9] A. P, A. Sharma, S. B. M, P. Pavankumar, N. K. Darwante, "Performance Monitoring and Dynamic Scaling Algorithm for Queue Based Internet of Things," 2022 International Conference on Innovative Computing, Intelligent Communication and Smart Electrical Systems (ICSES), 2022, pp. 1-7.

[10] Puneeth Kumar, B.S., Ramesh Naidu, P., Sridhara, S.B. (2023). Internet of Things and Cognitive Radio Networks: Applications, Challenges and Future. In: Yadav, S., Chaudhary, K., Gahlot, A., Arya, Y., Dahiya, A., Garg, N. (eds) Recent Advances in Metrology . Lecture Notes in Electrical Engineering, vol 906. Springer.

[11] H. F. Atlam, A. Alenezi, M. O. Alassafi, and G. B. Wills, "Blockchain with internet of things: benefits, challenges, and future directions," International Journal of Intelligent Systems and Applications, vol. 10, no. 6, pp. 40–48, June 2018.

[12] J. Huang, C. Xing, S. Y. Shin, F. Hou, and C. Hsu, "Optimizing M2M communications and quality of services in the IoT for sustainable smart cities," IEEE Transactions on Sustainable Computing, vol. 3, no. 1, pp. 4–15, Jan 2018.

[13] Avinash Sharma, B. Kameswara Rao, Ravi Shankar, Parismita Sarma, Abhay Chaturvedi, Naziya Hussain, Industrial quality healthcare services using Internet of Things and fog computing approach, Measurement: Sensors, Volume 24, 2022, 100517

[14] P. K. Verma, R. Verma, A. Prakash, A. Agrawal, K. Naik, R. Tripathi, M. Alsabaan, T. Khalifa, T. Abdelkader, and A. Abogharaf, "Machine-to-Machine (M2M) communications: A survey," Journal of Network and Computer Applications, vol. 66, pp. 83–105.

[15] P, A. Sharma, A. Singla, N. Sharma, "IoT Group Key Management using Incremental Gaussian Mixture Model," 2022 3rd International Conference on Electronics and Sustainable Communication Systems (ICESC), 2022, pp. 469-474.

[16] H. Ghayvat, J. Liu, S. C. Mukhopadhyay, and X. Gui, "Wellness sensor networks: A proposal and implementation for smart home for assisted living," IEEE Sensors Journal, vol. 15, no. 12, pp. 7341–7348, Dec 2015.

[17] A. P, A. Sharma, A. Sharma, K. S and M. R. Arun, "Priority Queueing Model-Based IoT Middleware for Load Balancing," 2022 6th International Conference on Intelligent Computing and Control Systems (ICICCS), 2022, pp. 425-430.

[18] M Sandeep Prabhu, M Ramesha, Jayashree M Kudari and Ansuman Samal (2021), Smart Agriculture and Smart Farming using IoT Technology, Journal of Physics: Conference Series, 2089(1), pp. 012038. https://doi.org/10.1088/1742-6596/2089/1/012038.

[19] Arudra Annepu, M Ramesha,K Prashantha Kumar and Pallavi Singh (2021), IoT Enabled Smart Lighting System for Smart Cities, Journal of Physics: Conference Series, 2089 (1), pp. 012037. https://doi.org/10.1088/1742-6596/2089/1/012037.

[20] Pankaj Mudholkar, Megha Mudholkar, B S Puneeth Kumar and S. Srinivasulu Raju (2021), Smart Villages: IoT Technology Based Transformation, Journal of Physics: Conference Series, 2070(1), pp.012128. https://doi.org/10.1088/1742-6596/2070/1/012128.

[21] G. Naveena Pai, M. Swathi Pai, M. Shruthi and B. Naveen K, "Internet of Things: A Survey on Devices, Ecosystem, Components and Communication Protocols," 2020 4th International Conference on Electronics, Communication and Aerospace Technology (ICECA), 2020, pp. 611-616, doi: 10.1109/ICECA49313.2020.9297458.

[22] S. B. Sridhara, K. B. Naveen, M. Ramesha, and G. N. Pai, "Internet of things: Internet revolution, impact, technology road map and features," Adv. Math. Sci. J., vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.

Proceedings of the International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT 2023)
IEEE Xplore Part Number: CFP23CV1-ART; ISBN: 978-1-6654-7451-1

2023 International Conference on Intelligent Data Communication Technologies and Internet of Things (IDCIoT) | 978-1-6654-7451-1/23/$31.00 ©2023 IEEE | DOI: 10.1109/IDCIoT56793.2023.10053473

Customer Churn Analysis in Financial Domain using Deep Intelligence Network

Dr.Sandeep Kumar Hegde
Assistant Professor
Department of Computer Science and Engineering
NMAM Institute of Technology, Nitte (Deemed to be University)
Karkala Taluk, Udupi District, India
sandeep.hegdey@gmail.com

Rajalaxmi Hegde
Assistant Professor
Department of Computer Science and Engineering
NMAM Institute of Technology, Nitte (Deemed to be University)
Karkala Taluk, Udupi District, India
laxmi123.prabhu@gmail.com
(Corresponding Author)

Sudhansu Sekhar Nanda
Associate Professor
Kirloskar Institute of Management
Karnataka, India.
nandasudhansusekhar.87@gmail.com

Gururaj Phatak
Assistant Professor
Kirloskar Institute of Management
Harihar, India.
gururajphatak4u@gmail.com

Pushpa Hongal
Assistant Professor
Kousali Institute of Management Studies
Karnatak University
Dharwad, India
pushpaphd2022@gmail.com

Dankan Gowda V
Assistant Professor
Department of Electronics and Communication Engineering,
BMS Institute of Technology and Management
Bangalore, Karnataka, India.
Email: dankan.v@bmsit.in

***Abstract*—Customer churn analysis is regarded as a crucial indicator that determines the revenues and profitability of the organisation in the modern day due to the advancement of technology and business models. Regardless of the company's size including startups, it is relevant to all business sectors, mainly because 65% of the company's revenue comes from its current clients. The aim of the analysis of customer churn is to identify the clients most likely to leave the existing organisation. In order to accurately predict customer churn in the banking industry, this research aims to develop an optimized deep neural network model. Five predictive analytical models from various categories of machine learning algorithms are also compared to the suggested model in the comparison study such as Gaussian Naive Bayes algorithms, decision trees, logistic regression and Artificial neural networks. The outcome demonstrates that, when compared to all other predictive models, the proposed deep intelligence network model exceeds them all in terms of accuracy.**

Keywords—Predictive Model, Deep Neural Network, Customer Churn, Banking Sector

I. INTRODUCTION

Retaining profitable customers is thought to be one of the most important factors for the success of business in this cutthroat commercial era caused by the rapid advancement of technology and globalisation. Any industry that depends on customer involvement, including telecom, advertising, finance, and healthcare, can use the customer churn analysis. One of the important facts is that keeping existing customers costs 15 times less than getting new ones. All company sectors are consequently focusing on client churn issues in order to increase their effectiveness and profitability [1]. Customer churn analysis is regarded as a significant research area in data science. Given that each specific banking business has a number of clients, the volume of the customer descriptions data may be bigger if the banking sector is taken into account. These data can be utilised to forecast client churn by identifying hidden patterns and understanding the connections between the data, which will increase the banking industry's productivity. To extract a meaningful pattern from produced data that is still in its raw form, efficient machine learning algorithms are needed. In the banking industry, the phrase "churning" or "churn" refers to consumers who quit conducting business with the bank and/or shut all of their accounts. There has already been research on predicting client turnover or churn for a number of businesses, including insurance, retail, and the telecommunications sector. However, there haven't been many research done on the banking industry. Due to rising globalisation and competition in the financial markets, banking sectors must embrace customer-oriented practises in order to prosper in the commercial era of banking. The more money and profits a bank may make, the longer it keeps a valuable customer. In order to keep its valuable clients and remain competitive on the market, the banking sector must identify client turnover rates [2]. Clients routinely change banks in order to purchase the same products from several banks, thus the bank must predict which clients are most likely to migrate. Deep neural network-based methods are utilized in the proposed study to build the prediction models. Deep neural networks are a promising technique for predictive analytics [3]. The learning process is controlled by a computer programme utilizing a multi-layered neural network. In this case, learning from the input data is carried out using a variety of weights while a large number of neurons are active. Deep learning has several uses, including natural language processing and the analysis of multimedia data, including audio, video and picture. The proposed study aims to develop an optimised deep neural network that can

accurately predict customer churn rate in the banking industry.

II. REVIEW OF LITERATURE

The various research work that has already been conducted on predicting customer churn using various machine learning models in the banking, telecom, and retail industries is analysed in this part. The paper proposes a decision tree-based prediction models for consumer churn analysis in the banking domain [4]. One of the key benefits of employing a decision tree is the ease with which the if then rules can be built since consumer attrition data is of the categorical and numerical data types. To create the prediction model, the CRISP method is applied. Using the ROC curve, the models are verified. The features are picked using backward exclusion and forward selection strategies. This model obtains an accuracy of 85% when used with the limited amount of data. The work's flaw, though, is the vast amount of data. To anticipate customer attrition, a combination of Randomforest and Support Vector Machine approaches is applied. Hyper planes are used by the supervised learning method known as the support vector machine to distinguish between several classes. In order to discriminate between the classes in high-dimensional space, hyperplanes are employed to produce the biggest margin possible. The categorization flaw is less when the margin is larger. Instead of building just one model, the random forest ensembling technique builds several, and then integrates the accuracy of each model to build a more accurate prediction [5].

The banking industry's client turnover is predicted using an artificial neural network [6]. The Bank data set is collected from a UCI repository. Rapid Miner Simulator is employed in the specified job. Three hidden layers, two output neurons, and nine input neurons made up the neural network. This model has an overall accuracy of 78.18%. The banking industry's client attrition is predicted using the Alyuda Neuro Intelligence simulator in [7]. The goal of the model is to forecast client churn depending on how many goods a consumer chooses. The term "churner" refers to a consumer who utilises no more than three products from the bank. Young people are more likely to be bank churners than older people, according to research, if they have less than three bank products.

Utilizing a variety of the methods described in [8], the customer churn study is conducted. The JRip technique and K means clustering are two of the analysis's methodologies. Here, the WEKA simulator is used. The dataset of riverbanks of Niegeria were employed to collect the data. The data collection included roughly one lakh client records, each of which had eleven distinctive characteristics. The weka environment was used to perform data set cleaning and modification. Training examples were used to train the models, which were then tested on the dataset. After submitting the data set to the k means clusters and resulting in a variety of clusters, the JRIP classifier is used to assess the data. The telecommunications industry uses the idea of churn analysis as well. The various machine learning techniques are used in the paper [9] to analyse customer turnover. The dataset was used from five telecoms firms. With the fast miner tool, the model is created using machine learning methods. Pre-processing is done on the data before it is converted into numerical values for the predictive analysis. The relationship between the values is ascertained using the FP Growth methods before the data is fed into the machine learning model. To do a comparison analysis, the data set is run through five different models. The model's optimality is established in light of the result. In [10], a churn forecasting model for the telecommunications industry is developed using deep learning model. These deep learning models are fed the two different sets of telecom data. A convolution neural network and feed forward neural network are used to achieve an accuracy of 72%. Deep learning models are found as effective as random forest methods, according to experimental findings.

In order to calculate client attrition in the retail industry, [11] uses deep learning-based algorithms. Limited Boltzmann machines are used to create the deep learning models. The client turnover at the supermarket is estimated using POS transactional data sets based on consumer buying habits. ETL will be utilised in conjunction with the input dataset. Prior to being split into both training and testing, outliers are eliminated from the dataset. After the outliers have been eliminated, the data set is randomly split into train and test split portions 80:20. The restricted Boltzmann machine model and the convolution neural network both receive the data set as input. For the CNN testing and training datasets, class labels and features are organised into two distinct matrices. The training set was iterated a total of 30 times. The approach achieved an accuracy rate of 76.9%.

Customer attrition is predicted using a gradient boosting technique in [12]. All transactional data related to payments and subscriptions are seen as input data sets. Cross validation, training data set, and test data set are the three categories into which the data set is divided. The logloss model is employed to evaluate the model's effectiveness. There are a total of 208 attributes in the dataset. The model's accurate characteristics are preserved, while superfluous features are removed. To implement the XGBoost library, Python is utilised. The firefly algorithm is used in [13] to provide the churn forecast for the telecommunication sector. Depending on how bright the light is, each firefly is matched to every other firefly. The ability of firefly algorithms to select the best solution stems from their metaheuristic algorithm. The French telecom company runs an experiment on the data it collects from orange dataset. By mimicking annealing, the firefly's intensity may be controlled. The biggest issue with this method is the enormous amount of computing needed to compare every firefly throughout each cycle to every other firefly. In [14] PSO-based feature selection and simulated annealing are used to create churn projections for telecom data. Several machine learning techniques are used to evaluate the prediction efficacy of the recommended method. In order to address the issues and challenges involved in the banking sector, big data analytics-based techniques are implemented in [20] to the dataset connected to the banking domain.

The present machine learning-based algorithms for predicting customer churn are computationally costly due to the requirement to repeatedly loop over the sizeable training range of data until the model adapts [15]. Another drawback of these models is their poor performance with highly dimensional customer data, which results in a bias in favour of the classes with lower incomes. They frequently forecast data from the majority as a result of this rather than the

minority. To address these concerns, the proposed work is done with the following goals in mind.

- Apply the One Hot Encoding, Tukey Outliers and Label Encoding approaches to each individual column to perform adaptive pre-processing and feature extraction while preserving the column's original modified features that were related to the encoding procedure.
- Greedy forward selection will be used to choose just the fraction of the binary features that are encoded after training the proposed approach on the initial converted columns.
- Hyper parameter fitting and tuning must be done in order to select the optimal features for the model training procedure using effective optimization approaches, such as Adam optimizer, to get the maximum accuracy out of a machine learning model

III. PROPOSED METHODOLOGY

A model based on an adaptive deep neural network (ADNN) is implemented in the proposed study to forecast client churn in the banking industry. The UCI machine learning repository was used to obtain the customer churn data set. There are a total of 14 features in the 10,000 rows that make up the data collection. The Exit variable or class label determines customer turnover. A binary value of 0 indicates that a client stays with the banking organization, whereas a value of 1 indicates that the customer leaves.

Figure 1 below shows the suggested approach's architectural layout. The proposed Model is given the customer churn data collection as input during the first step. Outliers are located and eliminated during the pre-processing stage using Tukey's outlier identification model [16]. One of the important advantage of data pre-processing is that it will improve the prediction accuracy by removing the inconsistent data values resulted due to human or computer error.

The interquartile range (IQR) is used by the Tukey Outlier model to identify outliers. This technique, which doesn't rely on a statistical mean or standard deviation, can handle the large range of values in the presented dataset. The input data set includes independent factors, such as the customer id, last name, and row number, which are not taken into account because they have no bearing on the dependent variable. Due to the fact that the model requires inputs in numerical data, the features are encoded using a label encoding and hot encoding approach. This strategy has already been applied. The data were fitted, transformed, and encoded in numerical form using the label encoder class. The label encoding is applied on one hot encoded data, which divides the column into several columns. The digits 1 and 0 have been used in place of the numbers in the respective columns. Exploratory data analysis employs pre-processed data sets that are graphically portrayed in order to analyse and summarise the aspects of the data. Following data exploration, data normalisation and feature augmentation are applied to the data sets. Since the magnitudes, ranges, and types of the features in the real-world dataset vary so widely, normalisation must be conducted to scale the features that are insignificant or misleading. When features are scaled, the Euclidean distance among sample locations are determined and then normalised. In the proposed research work min max normalization method is utilized to normalize the data. This crucial phase may be used to scale up the characteristics for quicker calculation and standardise the data [17].After feature scaling and normalisation of the supplied data set, the input, hidden, and output layers of the keras library are imported to create a deep learning model. In the example provided, there are eleven independent variables in the input layer, which is made up of a number of nodes, and one dependent variable in the output layer. Utilizing the average of the output layer and input layer, the six hidden layers are created. Using the kernel initializer, the Weights are equally initialised. Without the requirement for an activation function, which has limited capacity and frequently performs badly, a neural network behaves similarly to a standard linear regression model.

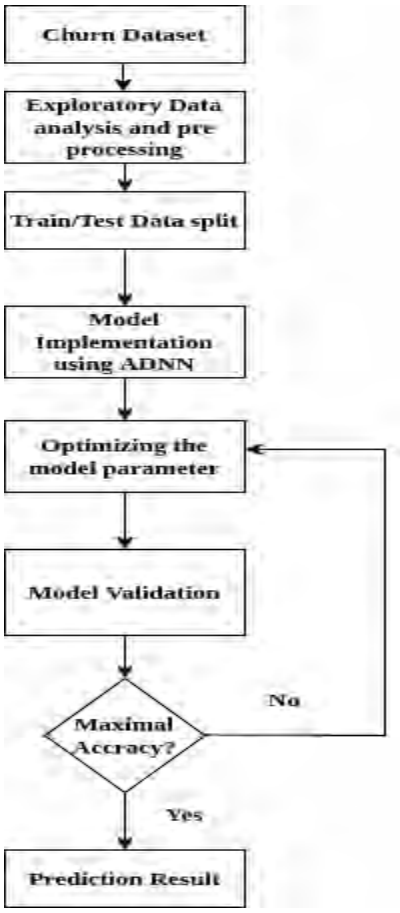


Fig 1: Flowchart of the Proposed Approach

The activation function is employed as a result, and by using relu and softmax for the hidden layers which is ideal [18]. It resolves the vanishing gradient problem that may affect activation functions like tanh. The model's performance is assessed using two subsets of the whole data set. The output layer's activation function is sigmoid, which offers the best performance and a range of [0,1]. About 80 percent of the data are in the training set, while the remaining 20 percent

are in the test set. This is primarily motivated by the need to test the model on new data after it has been trained on training samples to prevent overfitting issues. Adam optimization method is used in applications of deep learning and is seen as an expansion of stochastic gradient. Deep learning models can be optimised via the Adam algorithm. In comparison to AdaGrad and RMSProp the Adam is more effective [19]. Adam is an optimization strategy that is considered when a network in a deep learning application, as opposed to the traditional stochastic gradient descent method. The advantages of this method are its efficiency in terms of processing power and the resulting lower memory needs. It works effectively for issues with bigger amounts of data and/or settings. For issues involving noisy or sparse gradient data, even the Adam optimization can be applied. Little hyper-parameter adjustment is necessary. Adam optimizations change the learning rate based on the network weight, whereas stochastic gradient descent keeps the learning rate constant [20].
The term "log loss model" also refers to the binary cross entropy loss. It is employed to assess how well a categorization model is doing. The binary cross entropy generates probabilities between 0 and 1. Loss will grow if the system probability differs from the actual label. Validation datasets are utilised throughout the model training phase to ensure unbiased results. Each of the 100 epochs, which represent a forward or backward traversal of the training examples, is carried out in batches of 50. The term "Batch size" refers to all of the training examples utilised for a single forward pass and a single backward pass. In proportion to the RAM it consumes, the batch grows in size. The total number of times the instances were passed was indicated by the term "iteration." The predicted values are only taken into consideration as true if the forecast is larger than 0.5; otherwise, they are taken into consideration as false.

The steps below describe how to train a model using an adaptable deep learning network model.

In the first step, the input layer of the neural network node is filled with the attributes of the input data set.

Step 1: We arrive at N[0]=P, where N denotes the neural network's nodes or units and P denotes the properties of features sent among neurons' nodes.

Step 2: The neural network's nodes have evenly distributed weights, which are then multiplied by the appropriate features as illustrated below.

The weights used for each of the features are represented by N[i]= Wf[i]* P, Wt stands for the applied weights for each feature.

Step 3: The intersection of the features of the input layer and the edge weights between hidden layer and the input layer is determined as

input _hdden_layer= Wf[i]* P

Step 4: The hidden layers' neurons are stimulated from left to right using forward propagation. With Softmax acting as the activation function, the weights are distributed via ReLu.

Activation_hiddenlayer=ActivationFunction(inputl_hidden_layer)

Step 5: The activation function is once more used to predict the outcome after activating the hidden layer and carrying out the linear transformation.

input_layer_output = (activation_hidden_layers * Wf[i])

Step 6: The actual outcome is contrasted with the prediction, and any inaccuracies that were predicted are quantified.

Error=Eoutput-Ooutput

Where Eoutput is Actual output and Ooutput predcited output.

Step 7: Adam optimization is adaptive weighting technique that updates the weights in accordance with the cognitive rate of each fold

Step 8: Steps 1 to 5 are carried out after each observation, and Adam optimizer automatically updates the weights. This action is akin to reinforcement learning.

Step 9: The deep neural network is given the whole training set, and until the accuracy is perfect, the epochs are repeated.

IV. RESULT AND DISCUSSION

The data set is split into training and test sets in a 80 - 20 ratio in order to assess the algorithm's performance. Eighty percent of the training examples are used to train the model. The models are fed with the test sets. To determine if the suggested algorithm is accurate, the predicted outcomes are contrasted with the actual labels. The performance of the proposed model is validated using the various parameter such as Precision, Recall, F measure, ROC and confusion matrix .The recall level of the proposed model is computed using following formula.

Recall = True Positives / (True Positives + False Negatives)

The precision rate of the prediction result is computed as

Precision = True Positives / (True Positives + False Positives)

The true positive rate is calculated using the following expression.

True Positive rate= True Positives / (True Positives + False Negative).

The F measure rate was computed using the expression below.

F-Measure = (2 * Precision * Recall) / (Precision + Recall)

A. Decision Tree

The decision tree (DT) model recorded an accuracy rate of 78.4%. The confusion matrix and ROC curve are used to assess the model. Figure 2 below shows the confusion matrix for the decision tree. The confusion matrix yields 1379 true negative predictions, 196 true positive predictions, 201 erroneous negative predictions, and 224 false positive predictions in total.

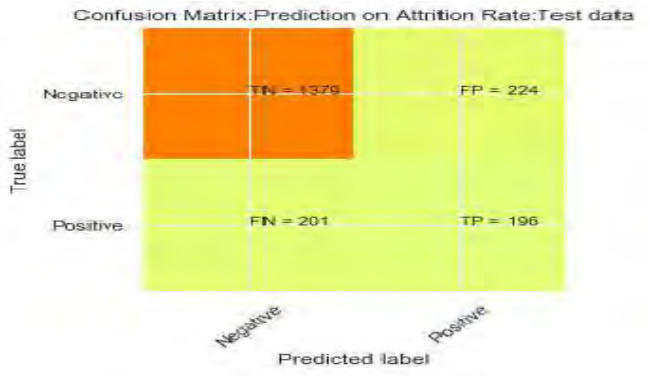


Fig. 2. Confusion matrix with decision tree

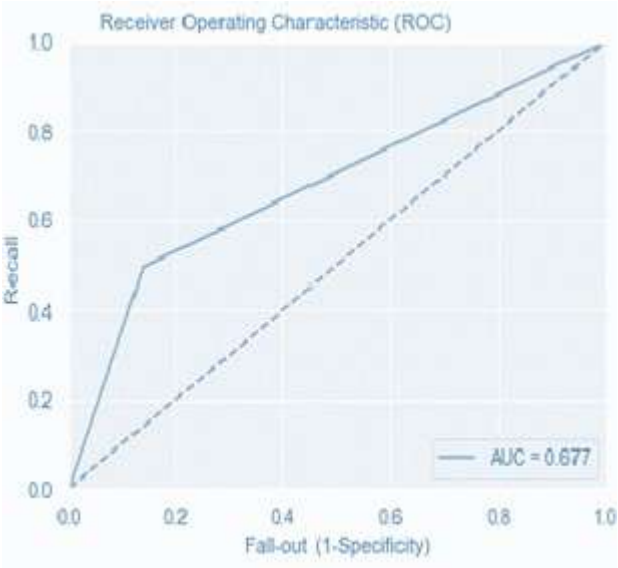


Fig. 3. ROC curve with decision Tree

Figure 3 depicts the ROC curve generated with decision tree when imposed on the decision tree model. The performance of the multi classification problem can be determined using ROC (Receiver Operating Characteristics) curve. The false positive and true positive rates found in the provided model are used to build the ROC curve. The degree or measure of separability is represented by AUC, and ROC is a probability curve. By analogy, a model's ability to identify between patients who have the disease and those who do not is improved by a higher AUC. The ROC curve is represented with true positive rate against the false positive rate on X axis and Y axis respectively. For the provided data set, the decision tree yields an AUC rate of 0.67.

B. Logistic regression

The logistic regression (LR) algorithm's accuracy with the given churn data set was 80.15%. The confusion matrix created using logistic regression is shown in Figure 4. Using this approach, there are 363 false negative predictions and 34 false positive predictions, resulted in 1561 true negative predictions and 42 true positive forecasts. The model is validated in a manner similar to that of the decision tree. The Figure 5 below shows the ROC curve obtained with the logistic regression model.

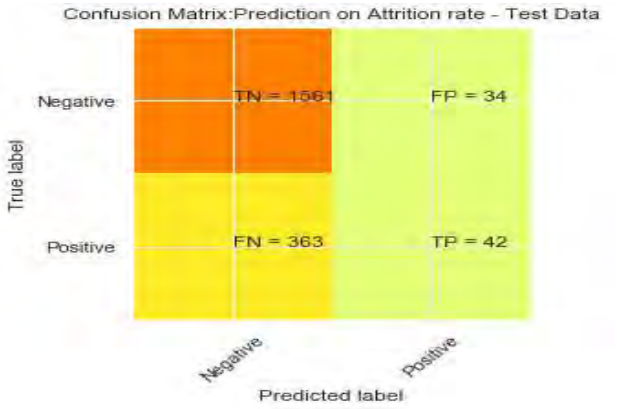


Fig. 4. Confusion matrix plot with logistic regression

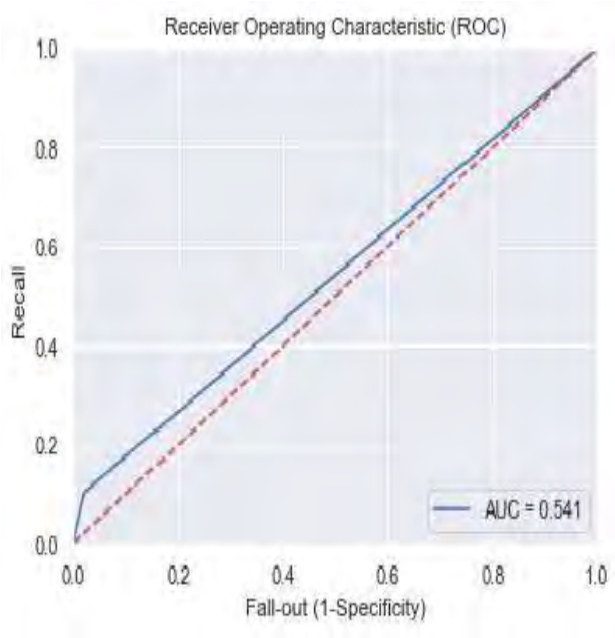


Fig. 5. ROC curve plot with regression model

The logistic regression model yields an AUC rate of 0.541 for the provided data set. It has been found that Logistic regression has higher accuracy than Decision tree.

C. Gaussian Naïve bayes

Since the original data used in the proposed study are actual, valuable inputs, the Gaussian Nave Bayes (GNB) model can be used. By calculating the mean and standard deviation, the Gaussian Naive Bayes method determines or compresses the distributions for each class using the real valued points as input. Thus, Gaussian Naive Bayes is used to generate the predictions. This method outperforms all baseline models in terms of accuracy when estimating customer attrition, coming in at 81.35%. The confusion matrix produced using Gaussian Naive Bayes is shown in figure 6 below. The Gaussian naive Bayes model produced 1535 true negative, 75 true positive, and 330 false negative cases.

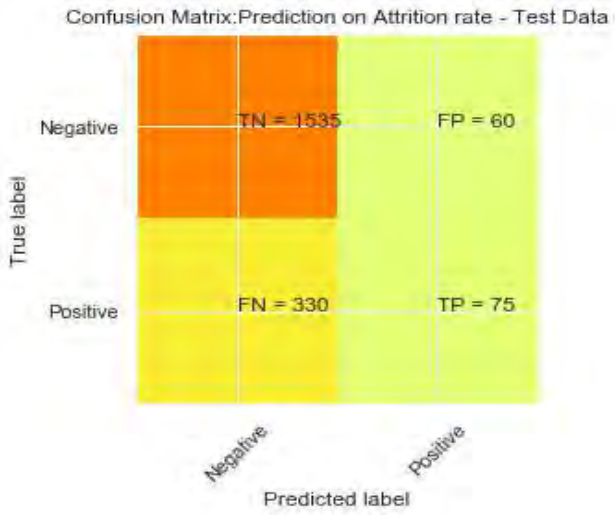


Fig. 6. Gaussian naïve bayes's Confusion matrix plot

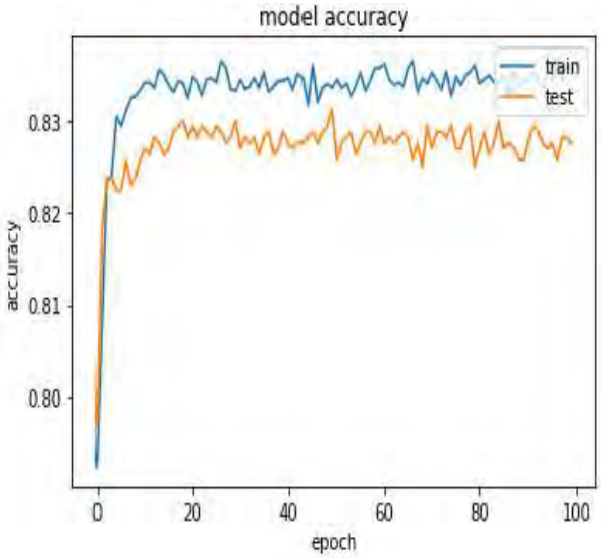


Fig. 8. Neural Network Model's Accuracy Vs Epoch plot

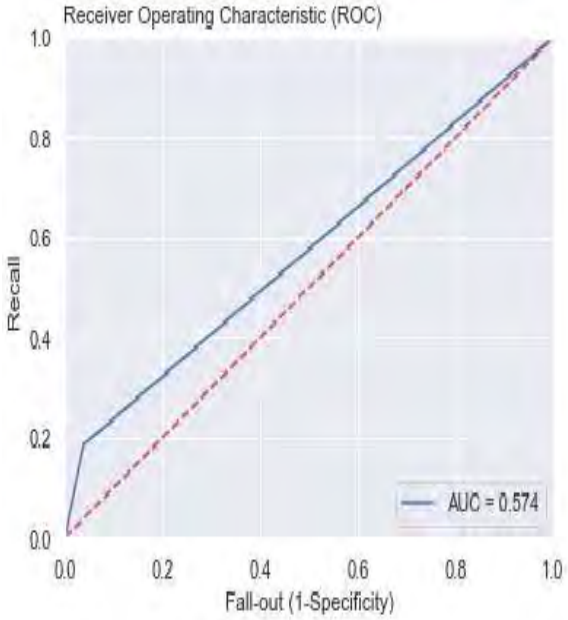


Fig. 7. Gaussian Naïve Bayes 's ROC curve plot

Figure 7 shown above depicts the model's ROC curve. The model achieved AUC rate of 0.574 using naive bayes model.

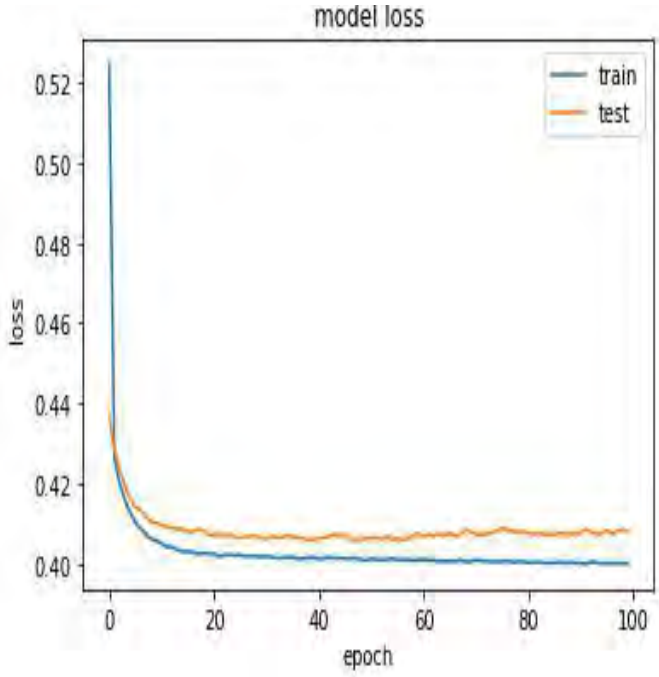


Fig. 9. Neural Network Model's loss Vs Epoch plot

The Figure 9 demonstrates that when more epochs are carried through the training instances, the model loss reduces, leading to an improvement in accuracy with each epoch. Figure 10 shows the confusion matrix that was discovered using ANN. The neural network's confusion matrix produces 1553 true negative, 129 true positive, 276 false negative, and 42 false positive predictions.

D. Artificial Neural Network

The experiment is improved by adopting the basic input- and hidden-layer architecture of an artificial neural network (ANN). The input layer employs the ReLU as the activation function, whereas the output layer utilises the sigmoid. Using the Adam optimizer, the feature weights were changed. Using binary crossentropy or logloss, the model is verified. 100 training example epochs are carried out during the model training phase. The training set results in an accuracy of 84%. Figures 8 and 9 below depict the neural network model's model accuracy and model loss graphs. Figure 8 demonstrates how a neural network model might obtain a test accuracy of 82.9% and a training accuracy of 83.59%.

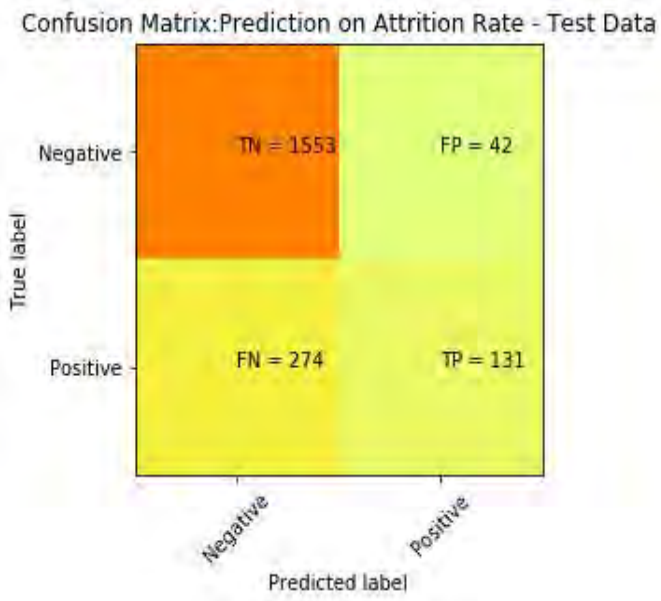


Fig. 10. Confusion matrix of Artificial Neural Network

The ROC curves for the neural network are shown in figure 11. This model achieves an AUC rate of 0.79. This rate has the highest AUC value when compared to all other baseline models.

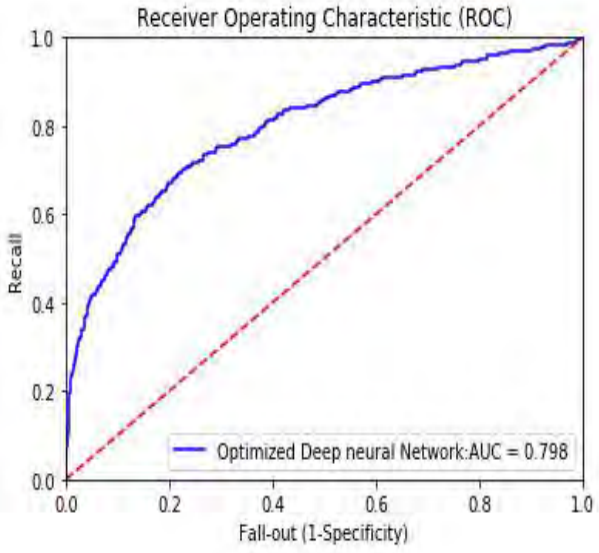


Fig. 11. ROC curve of neural network, indicate that AUC rate of 0.798.

E. Proposed Adaptive Neural Network

The proposed adaptive deep neural network (ADNN), which includes five hidden layers, is used to create the prediction model. Six nodes are found in each layer. The number of nodes in each layer depends on the training set's size. The model receives the customer churn data set as input. Each of the three groups—test, training, and validation—has ten batches of training samples and is comprised of the same data set. The training example performed 100 epochs of iteration. Figures 12 and 13 show the model loss and model accuracy graphs produced using the proposed approach.

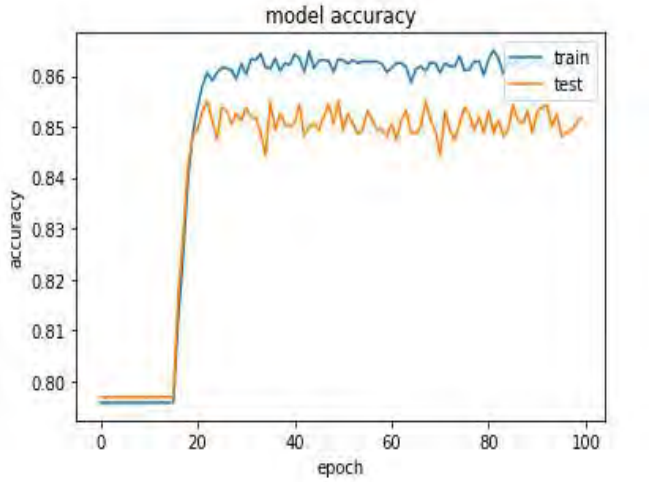


Fig. 12. ADNN 's Model Accuracy Vs Epoch Curve

According to the accuracy graph of the ADNN shown in Figure 12, the accuracy with training data was 86.23% and the accuracy with test data was 85.9%. As depicted in figure 13, the model loss lowers as training cases are run through more epochs, increasing accuracy with each one. The training set of the proposed model has 100 epochs altogether. When using training data, the model loss at the 100th epoch is less than 5%, and when using test data, it is less than 10%.

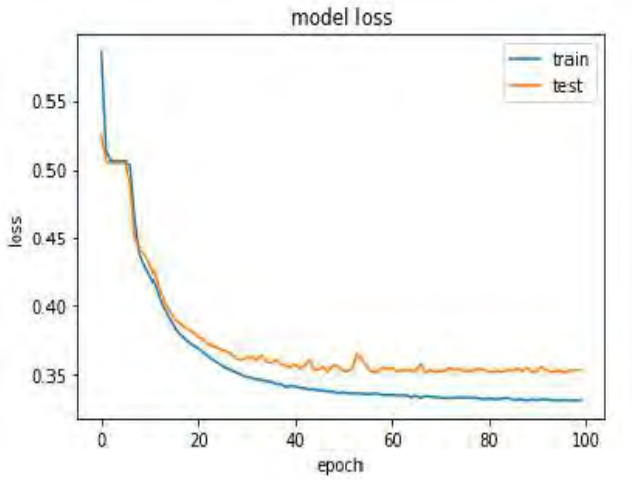


Fig. 13. ADNN 's Model loss Vs Epoch Curev

Figure 14 shows the confusion matrix that the proposed ADNN model generated. Based on the true negative, true positive, and false negative, false positive rates of the proposed model, the confusion matrix is produced. The proposed ADNN's confusion matrix revealed 215 False negative predictions for every 1528 True negative predictions.

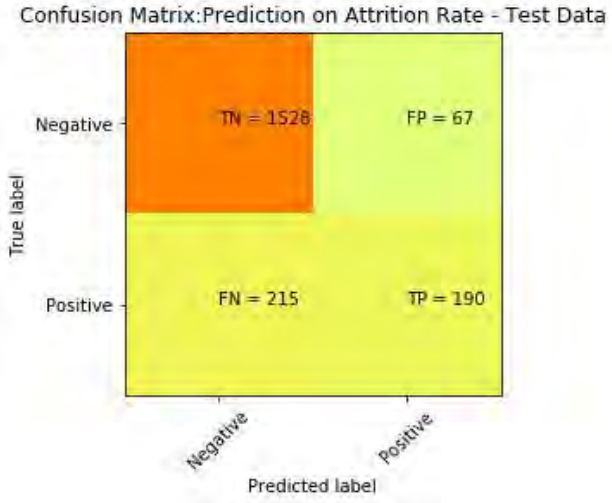


Fig. 14. Confusion matrix of the proposed ADNN

As depicted in figure 13, the model loss lowers as training cases are run through more epochs, increasing accuracy with each one. The training set of the proposed model has 100 epochs altogether. When using training data, the model loss at the 100th epoch is less than 5%, and when using test data, it is less than 10%. Figure 14 shows the confusion matrix that the proposed ADNN model generated. Based on the true negative, true positive, and false negative, false positive rates of the proposed model, the confusion matrix is produced. 215 False negative predictions were found in the proposed ADNN's confusion matrix for every 1528 True negative predictions. The figure 15 above represents the ROC plots achieved through proposed ADNN. The graph illustrates that an AUC rate of 0.862 is obtained with proposed approach when imposed on the churn dataset.

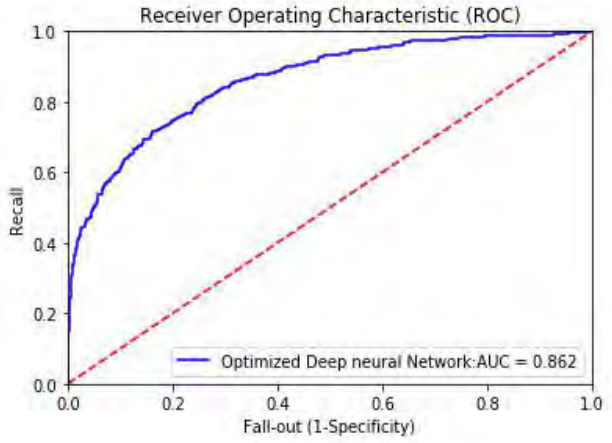


Fig. 15. ROC curve plot with ADNN

TABLE I: ACCURACY WITH DIFFERENT ML MODEL

ML Model	Test Data Accuracy	Recall	Precision	Rate of AUC	F measure
LR	80.15%	0.79	0.75	0.541	0.64
DT	78.4%	0.74	0.73	0.677	0.68
ANN	82.6%	0.81	0.83	0.798	0.77
ADNN	85.9%	0.86	0.88	0.864	0.88
GNB	80.3%	0.78	0.81	0.574	0.75

Table I compares the validation results obtained with the various machine learning algorithms. The table above illustrates that the proposed adaptive deep neural network model outperforms the traditional machine learning algorithm in terms of accuracy ,recall ,precision ,F measure and AUC rate. Figure 16 shows the bar charts that contrast the AUC rate and algorithm accuracy of several machine learning methods. The bar plot illustrates that the proposed adaptive deep neural network outperforms all existing standard machine learning methods in terms of AUC and accuracy rate.

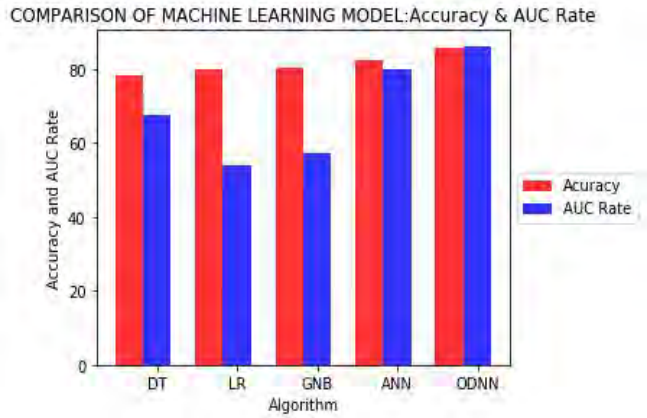


Fig. 16. Accuracy and AUC rate plot with ML algorithm

V. CONCLUSION AND FUTURE WORK

The proposed work uses a variety of machine learning techniques to forecast the customer churn rate in the banking industry, including deep neural networks, decision trees, artificial neural networks, and the proposed adaptive deep neural network model. With test data, the proposed adaptive deep neural network model produced 85.9% accuracy and an AUC rate of 0.864. This demonstrates that the proposed approach performs better than current machine learning methods in terms of accuracy and AUC rate. As part of the future work the proposed adaptive deep neural network can be applied in the telecom sector and health care sectors to analyze the churn rate.

REFERENCES

[1] Machowska, Dominika. "Investigating the role of customer churn in the optimal allocation of offensive and defensive advertising: the case of the competitive growing market." *Economics and Business Review* 4.2 (2018): 3-23

[2] Hassani, Hossein, Xu Huang, and Emmanuel Silva. "Digitalisation and big data mining in banking." *Big Data and Cognitive Computing* 2.3 (2018): 18.

[3] Hatcher, William Grant, and Wei Yu. "A Survey of Deep Learning: Platforms, Applications and Emerging Research Trends." IEEE Access 6 (2018): 24411-24432.

[4] Keramati, Abbas, Hajar Ghaneei, and Seyed Mohammad Mirmohammadi. "Developing a prediction model for customer churn from electronic banking services using data mining." *Financial Innovation* 2.1 (2016): 10.

[5] Mahajan, Deepika, and Rakesh Gangwar. "Improved Customer Churn Behaviour By Using SVM." (2017).

[6] Kanmani, W. S., and B. Jayapradha. "Prediction of Default Customer in Banking Sector using Artificial Neural Network." *International Journal on Recent and Innovation Trends in Computing and Communication* 5.7 (2017): 293-296.

[7] Bilal Zorić, Alisa. "Predicting customer churn in banking industry using neural networks." *Interdisciplinary Description of Complex Systems: INDECS* 14.2 (2016): 116-124.

[8] Oyeniyi, A. O., et al. "Customer churn analysis in banking sector using data mining techniques." *Afr J Comput ICT* 8.3 (2015): 165-174.

[9] P.K.D.M.Alwis, B.T.G.S.Kumara, Hapurachchi,"Customer Churn Analysis and Prediction in Telecommunication for Decision Making" *International Conference on Buisness Innovation*,25-26 August 2018,NSBM,Colombo,Srilanka.

[10] Umayaparvathi, V., and K. Iyakutti. "Automated Feature Selection and Churn Prediction using Deep Learning Models." *International Research Journal of Engineering and Technology* 4.3 (2017): 1846-1846-1854.

[11] Dingli, Alexiei, Vincent Marmara, and Nicole Sant Fournier, "Comparison of Deep Learning Algorithms to Predict Customer Churn within a Local Retail Industry."J. Wang, "Fundamentals of erbium-doped fiber amplifiers arrays (Periodical style—Submitted for publication)," *International Journal of Machine Learning and Computing, Vol. 7, No. 5, October 2017*.

[12] Gregory, Bryan. "Predicting Customer Churn: Extreme Gradient Boosting with Temporal Data." *arXiv preprint arXiv:1802.03396* (2018).

[13] Ahmed, Ammar AQ, and D. Maheswari. "Churn prediction on huge telecom data using hybrid firefly based classification." *Egyptian Informatics Journal* 18.3 (2017): 215-220.

[14] Vijaya, J., and E. Sivasankar. "An efficient system for customer churn prediction through particle swarm optimization based feature selection model with simulated annealing." *Cluster Computing* (2017): 1-12.

[15] Zhang, Qizhi, et al. "Large scale classification in deep neural network with Label Mapping." arXiv preprint arXiv:1806.02507 (2018).

[16] Teh, Thomas, et al. "Generalised Structural CNNs (SCNNs) for time series data with arbitrary graph-toplogies." arXiv preprint arXiv:1803.05419 (2018).

[17] Seymen, Omer Faruk, Onur Dogan, and Abdulkadir Hiziroglu. "Customer churn prediction using deep learning." International Conference on Soft Computing and Pattern Recognition. Springer, Cham, 2020.

[18] Ashia C., et al. "The marginal value of adaptive gradient methods in machine learning." Advances in Neural Information Processing Systems. 2017.

[19] Jones, Pete R. "A note on detecting statistical outliers in psychophysical data." bioRxiv (2016): 074591.

[20] Shakya, Subarna, and Smys Smys. "Big data analytics for improved risk management and customer segregation in banking applications." Journal of ISMAC 3.03 (2021): 235-249.

Proceedings of the 5th International Conference on Smart Systems and Inventive Technology (ICSSIT 2023)
IEEE Xplore Part Number: CFP23P17-ART; ISBN: 978-1-6654-7467-2

Symmetrized Feature Selection with Stacked Generalization based Machine Learning Algorithm for the Early Diagnosis of Chronic Diseases

Dr.Sandeep Kumar Hegde
Assistant Professor
Department of Computer Science and Engineering
NMAM Institute of Technology, Nitte (Deemed to be University)
Karkala Taluk, Udupi District, India
sandeep.hegdey@gmail.com

Rajalaxmi Hegde
Assistant Professor
Department of Computer Science and Engineering
NMAM Institute of Technology, Nitte (Deemed to be University)
Karkala Taluk, Udupi District, India
laxmi123.prabhu@gmail.com
(Corresponding Author)

Veeranna Hombalimath
Associate Professor,
Department of Biotechnology,
KLE Technological University
Hubballi, Karnataka, India.
drvshombalimath@kletech.ac.in

D Palanikkumar
Professor
Department of Computer Science and Engineering
Dr NGP Institute of Technology
Coimbatore
Tamilnadu, India.
palanikkumard@gmail.com

Neha Patwari
Assistant Professor
Department of Information Technology
Thakur College of Engineering and Technology,
Mumbai, Maharashtra, India.
maheshwarineha0411@gmail.com

Dankan Gowda V
Assistant Professor
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management, Bangalore, Karnataka, India.
dankan.v@bmsit.in

2023 5th International Conference on Smart Systems and Inventive Technology (ICSSIT) | 978-1-6654-7467-2/23/$31.00 ©2023 IEEE | DOI: 10.1109/ICSSIT55814.2023.10061062

***Abstract*— Many healthcare sectors generate huge amount of healthcare data every day. To determine the risk of various chronic conditions, the collected data can be correctly analyzed using machine learning algorithms. Using machine learning algorithms, Health care analytics predicts the presence or absence of disease and detects various diseases. Health care analytics will improve patient care and medical practitioners' decision-making process by performing early disease detection process. Stacking is a meta-learning-based transfer learning approach that combines various machine learning algorithms. The advantage of layered generalization is that it allows to use the predictive capability of different machine learning algorithms to develop intermediate solutions, which improves the overall performance of the model. The proposed method utilizes a symmetric feature selection algorithm to extract the appropriate features from the chronic disease dataset. The proposed algorithm determines the merit of the feature by considering the entropy and information gain of the individual features. Metaheuristic algorithms can solve a wide range of complex real-world issues with high quality. A stacked generalization-based meta heuristic algorithm is used in the proposed approach to predict the possibility of chronic diseases. The experiment makes use of chronic disease datasets from the kaggle repository. The experimental findings reveal that the proposed symmetrized feature selection algorithm in combination with Stacked generalization-based Metaheuristics technique has obtained a better accuracy than the existing approaches.**

Keywords*— *Stacked generalization, Chronic Diseases, Genetic Algorithm, Metaheuristic, Machine Learning

I. INTRODUCTION

Chronic disease is now regarded as one of the world's most serious challenges in this modern age. Because the US spends nearly $3 trillion on health care to cure this disease [1]. As per a new, these diseases would impact 67% of the population by 2025. Chronic illnesses are ones that kill people over a long period of time. Heart disease. Cancer, renal disease, diabetes, asthma, and mental illness are examples of chronic diseases [2]. With 35 million diabetics, India already has the world's highest diabetes population. One solution could be to forecast the likelihood of chronic disease early on so that it can be treated as soon as possible. A key aspect of the machine learning algorithm is feature selection process. The accuracy with which the feature selection technique is used has an effect on the overall performance of the machine learning model. Wrapping, Filtering and embedding are considered as examples for feature selection techniques [3]. In the wrapper-based feature selection approach, a machine learning classifier is utilized to choose features [4]. The mechanism for selecting embedded features takes use of both filter and wrapper techniques. These strategies must be able to eliminate noise, outliers, and other non-predictive characteristics [5]. The risk of chronic diseases are early predicted by combining symmetrized feature selection technique with a genetic algorithm based on stacked generalization approach. The experiment is carried out using chronic disease datasets. The method is built on the concepts of redundancy and importance. The suggested feature selection technique takes the chronic disease datasets as input. Every feature will be subjected to a significance and redundancy assessment. The particular features are maintained only if their relevance exceeds the chosen threshold value and they are not redundant. As the application of meta-heuristic algorithms alone is insufficient to solve all difficulties, are not fast, thus in the proposed approach combination of filter and wrapper method has been applied. In the proposed work metaheuristic based is Genetic algorithm is used to perform optimization task.

To maximize the predictive ability of several machine learning algorithms such as Naive Bayes, Support Vector

Machine, Decision Tree, KNN, Random Forest, and Logistic Regression, the concept of stacked generalization based genetic algorithm (SG-GA) is used as a classifier. It combines the prediction results obtained with each approach using a voting-based strategy. Stacked generalization refers to the process of training a new model to include the predictions of two or more previously trained models. Stacking is also known as mixing since it incorporates predictions from previous models or sub models. With simple averaging or voting, the predictions of several sub models can be combined into an overall weighted sum. The models must be able to solve the problem, but they don't need to be the finest models conceivable. Existing research on stacked Ensemble methods is still going strong, however there are a few difficulties to address: (1) There has not been much work done to understand the weight of categories for classification selection. (2) There has not been enough research into the relationship between paired label correlation and multi-label classification performance. To overcome these challenges, proposed approach implemented a stacked generalization in combination with meta-heuristic algorithm known as genetic algorithm.

The following are the objectives of the proposed approach:

- Combining the finest machine learning architectures by voting strategy.
- Chronic illness prediction using a hybrid machine learning technique.
- Significantly improving the predictive accuracy of the machine learning model.
- Utilize machine learning models in conjunction with one another.
- Increasing the precision and reducing error rate in comparison to individual core models.

The K fold cross validation process eliminates the issue of model underfitting and overfitting. The precision, ROC, F measure, confusion matrix and recall are used to validate the results obtained using the proposed approach. According to the findings of the experiments, the suggested method outperforms an existing approach in terms of accuracy. The paper is structured in this manner. The second section focuses on existing literature work. The methodology is proposed in Section III. Section IV discusses the results of the experiments and follows with a conclusion section

.

II. LITERATURE REVIEW

This section discusses the various literary works that have previously been published on this subject.

To predict diabetic disease, a neuro-fuzzy interface model is deployed [6]. The data was gathered from National Institute of Diabetes to conduct the experiment. A Genetic Algorithm is used to choose the features. When compared to the RNN algorithm, the proposed methodology achieved better performance .In[7] decision tree-based algorithm is sued to forecast diabetes illness. The approach was accurate to within a 72 percent margin of error. The proposed technique omits the usage of a feature selection algorithm for the purpose of removing duplicate features. In [8] the use of an adaptive SVM-based system was demonstrated to identify the existence of breast cancer. The technique produced a 76 percent accuracy rate. However, the suggested approach has an overfitting problem. In[9], an LS-SVM approach was used to predict knee joint disease. The characteristics were selected using a genetic algorithm in conjunction with an apriori algorithm[10]. The proposed approach has a 94% accuracy rate. A feature selection strategy based on relief was paired with a linear SVM-based machine algorithm to anticipate heart disease[11]. The Weka data mining tool is used with clustering approaches to anticipate heart disease[12]. The proposed method showed a higher level of accuracy when compared to earlier strategies. In[13] to predict Heart Disease, a Nave Bayes classifier is used in combination with genetic algorithm . The proposed method had a 75 percent accuracy rate. It is suggested in[14] that infinite latent feature selection be used to choose a subset of Heart disease characteristics. The SVM algorithm was used to perform classification task in combination with the feature selection. Six hundred and ninety-nine instances were included in the study, each with thirteen unique characteristics. The best features for the classification job were found using proposed feature selection technique. In[15] the optimal features from the Breast cancer data set were extracted using a sub-optimal-based feature selection approach. In the feature selection procedure, ANN, SVM, and Bayes Network were applied. When compared to previous techniques, it was determined that the new methodology was more accurate. In[16] strategy for categorizing anti-cancer therapeutic responses based on autoencoders was proposed. To choose the subset of features, the autoencoder was combined with a Boruta-based technique. A random forest strategy was used to tackle the classification problem. When compared to earlier procedures, the new procedure proved to be more efficient. The study's primary goal is to develop an accurate machine learning model for predicting chronic disease. These diseases affect more than 67% of India's population. In [17] application of machine learning was analyzed to cancer prognosis and prediction. Additionally, the study provided an overview of the several techniques used to predict cancer problems. A system for forecasting renal disease was described in[18] that used a number of algorithms, including the J48, AD Tree, Random Forest, and Decision Tree algorithms. The random forest algorithm outperformed the other methods, according to the approach. In[19], numerous machine learning techniques were applied to predict renal illness, including Logistic Regression, Artificial Neural Networks, and Decision Trees. The findings of the experiment indicated that the Artificial Neural Network algorithm performed better at predicting disease. In[20], a support vector machine and a technique based on artificial neural networks were utilised to predict renal disease. The goal of this technique was to compare the algorithm's ability to forecast disease. The experimental findings illustrated that ANN algorithm outperformed better in predicting the disease. In [21] a neural network based model was proposed to early predict the numerous disorder through human nail. The nail pattern collected from the various patients were used to predict the disease. In the proposed work the photo of nail having certain condition was considered as dataset which was constructed using open cv tool. The color feature of the nail was used to discover the disease. The experiment was conducted by implementing neural network using input layer , two hidden layers and output layers where transfer function namely T ANSIG and PURELIN was implemented at the hidden layers .The experimental results illustrated that the result was 60% accurate in early predicting the disease.

III. PROPOSED METHODOLOGY

The proposed hybrid feature selection method has been used to predict chronic disease using the Ensemble learning approach (which combines several machine learning algorithms to achieve a common goal). The suggested method combines statistical analysis methods, machine learning algorithms, and a genetic algorithm to pick an optimal subset from the complete features space while taking advantage of both filtering and wrapper methods. The proposed system takes the disease dataset related to Breast Cancer, Heart Disease, Diabetes and Lung Cancer as input .The dataset to conduct the experiment was obtained from UCI repository. Data preprocessing will be applied to the dataset, removing any noise and outliers that may exist. The preprocessed data is fed into the Symmetrized Feature Ranking to select the relevant features from the dataset. The UCI set in terms of performance.

Algorithm: Symmetrized Feature Ranking

Consider Sm as a symmetry matrix
En denotes the entropy values of feature subsets
Fs indicates the Feature subset
Determine the entropy using

$$Entropy(f) = \sum_{i=1}^{c} - p_{fi} \log_2 p_{fi}$$

Obtain the merit of each of the feature traits as

$$M_{fs} = \frac{nm_{cf}}{\sqrt{n+n(n-1)m_{ff}}}$$

Determine the information gain InfoGn(FC|FV)
FC represent the class label of the considered feature
FV represents the feature values
for each $FC_i \in FC$ do
  Calculate Pf(FC[i])
    Fi= Fi +Pf(FC[i])*$\log_2$(P(Fc[i])
    F←Fi
end for
for each feature as ∈ FA
  calculate Pf(f[k])
  cunt = Fi+Pf(f[k])*$\log_2$(P(f[k]))
  Fi←count
end for
for every FCi do
  for every fj do
   dtermine Pf(FC[i]|f[k])
   F_{Sub}= Fs+Pf(BC[i]|a[k]) * $\log_2$ Pf(BC[i]|a[k])
   F← F_{Sub}
  end for
end for
Determine the symmetry matrix as

$$Symmtery(f_{i1,}f_{i2}) = 2 \cdot \frac{inforGain(f_{i1}|f_{i2})}{entrpy(f_{i1})+entrpy(f_{i2})}$$

Figure 1. Symmetrized Feature Ranking algorithm

The proposed symmetrized feature extraction algorithm employs the concepts of entropy, information gain, and merit of feature subset. The Feature selection process, removes unnecessary data, improve learning accuracy, and improve comprehensibility. The figure 1 shows the algorithm of proposed symmetrized feature ranking algorithm.

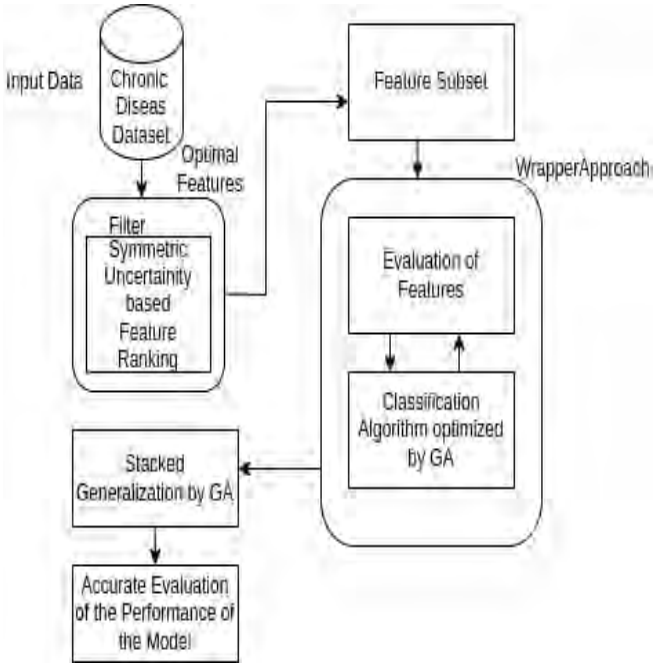


Figure 2 :Proposed Architecture Diagram

Algorithm: Stacked Generalization

Input: Data set Dt = {(Fi, c1), (F2, c2), ..., (Fm, cm) };
Initial level algorithms IL1, ..., CT;
Next level algorithm NL.
for l= 1, ... n:
hi = Li(Dt) % train using initial-level individual learner h_i
end;
Dt' = θ; % Create a new example set
for l = 1, ... , k:
for i = 1, ... ,n:
zli = hi(Fi)
end;
D' = D' U{(zil, Zi2, ..., ZiT) ,yi)}
end;
hr' = L(Dr'). % Apply the Next-level training algorithm to train hr'
Output: H (F) = hr' (h_1 (F), ... ,hT (F))

Figure 3:Stacked Generalization Algorithm

The mutual information principle is used by the symmetrized feature selection method to extract the best features from a set of features. The reciprocal data shows the relationship between each attribute and the influence it has on the target class. As a consequence, the usefulness of each feature is decided by assessing its relationship to the target class concept. To quantify the gain of each attribute, the concept of information gain is applied.

The frequency of occurrences of each class label will be determined using entropy. The procedure of selecting

features based on symmetrized feature selection algorithm is illustrated in Figure 1. To begin, equation 1 is used to compute the entropy of each feature. Purity is determined for every feature by utilizing equation 2 to calculate the entropy value. Each subset of the feature is rated for quality, and equation 4 is used to create a symmetrical link between the feature set. The feature selection system will choose the most relevant features from the input dataset in order to impact the prediction output.

Algorithm: Genetic
Initialize a baseline population
For (gen =1 to N_{genrn}) Do
  For(j=1 to N_{pouln})Do
    Determine the fitness of each individual j:f_i(gen):
  End for
Save optimistic individual for population gen +1
  For(j=2 to N_{popln}) Do
    Select the pair of individuals
    Create crossover of pair individual
    Mutate the selected individual
    Move individuals into population gen + 1
  End for
End for

Figure 4:Genetic Algorithm

Equation 1 below can be used to determine the entropy of a feature subset.

$$Entropy(Fs) = \sum_{j=1}^{n} -pf_s \log_2 pf_s \qquad (1)$$

In the equation (1) above Fs indicate the feature set pf_s indicates the proportion of the samples belongs to class n. Entropy always range between 0 and 1.

The information gain among two features f1 and f2 can be expressed as the difference between $entrpy(f_1)$ and $entrpy(f_2)$.

$$InfGain(f_1|f_2) = entrpy(f_1) - entrpy(f_2) \qquad (2)$$

Equation (3) below can be used to compute the merit of each feature subset.

$$Mf = \frac{nm_{clf}}{\sqrt{n+n(n-1)mf_sf_s}} \qquad (3)$$

In the equation above Mf represent the merit of the feature, n represents the number of features in the dataset, m_{clf} represents the average value of feature class correlation and mf_sf_s represents the average value of feature to feature correlation.

The equation (4) is used to determine the symmetry between the two features

$$Symmtery(f_{i1,}f_{i2}) = 2 \cdot \frac{inforGain(f_{i1}|f_{i2})}{entrpy(f_{i1})+entrpy(f_{i2})} \qquad (4)$$

The proposed architecture diagram is depicted in figure 2 above. The suggested framework is first supplied with the chronic illness dataset. To filter out the useful features from the provided dataset, a symmetrized feature subset algorithm is applied. The extracted features will undergo wrapper-based feature subset evaluation where the importance of the features is further evaluated by applying the various classification algorithms in combination with genetic algorithm. The genetic algorithm is used in combination with the various classifier to further optimize the predictive ability of the machine learning classifier. The algorithm of stacked generalization is shown in figure 3 above. Stacked generalization is an ensemble strategy in which a new model is trained to integrate the predictions of two or more previously learnt models or your dataset. Stacking is used to explore the space of alternative solutions to a particular problem. The stacking generalization is achieved in the proposed approach by means of various algorithms such as SVM, Decision Tree, Random Forest, Naïve Byes Classifier, Multi-Layer Perceptron (MLP). In the presented methodology stacked generalization algorithm is applied in combination with the genetic algorithm on considered dataset and the performance of the model is determined using the concept of voting. The genetic algorithm's pseudocode is displayed in figure 4 above. The benefit of utilizing a genetic algorithm with stacked generalization is that the genetic algorithm has the potential to explore the solution space in each generation. The stacked generalized model is optimized through genetic algorithm (GA) by evaluating the fitness of the solution. To further optimize the prediction result generated through stacked generalized model, the GA parameter like size of population , crossover and mutation rate are imposed on the implemented model which inadvertently improves the solution space of the model ,accordingly prediction result obtained through the algorithm will be improved. Overfitting and underfitting are no longer a problem using the K fold cross validation approach. The performance of the implemented model is accurately validated using the parameter such as F measure, precision, ROC area, confusion matrix and recall. The proposed method outperforms the baseline algorithms in terms of accuracy, according to the experimental results.

IV. Result and Discussion

The experiment was carried out utilising chronic disease datasets from the Kaggle repository[22].The diabetes dataset was originally gathered from National Institute of Diabetes .The dataset had detail about 488 diabetes patients which was defined in 9 columns. The lung cancer dataset contained description about 60 patients defined in 7 dimensions originally belongs to the Oxford University Hospitals NHS trust. The heart disease dataset was originally collected from Cleveland database had description about 1026 patients defined in 76 dimensions. The considered breast cancer dataset contained 570 patient detail defined in 32 dimension which originally extracted from Regional Cancer Registries.As input, the Datasets are fed into the Proposed system. To reduce noise and outliers, the collected data will be pre-processed.One of the important advantage of data pre-processing is that it will improve the prediction accuracy by removing the inconsistent data values resulted due to human or computer error. In the proposed approach noise are analysed by applying Synthetic minority over sampling technique and outliers are removed by applying Interquartile range analysis(IQR) .The proposed symmetrized feature selection technique is used to filter the best features from the datasets. The predictions are made using a stacked generalisation model based on a genetic algorithm.

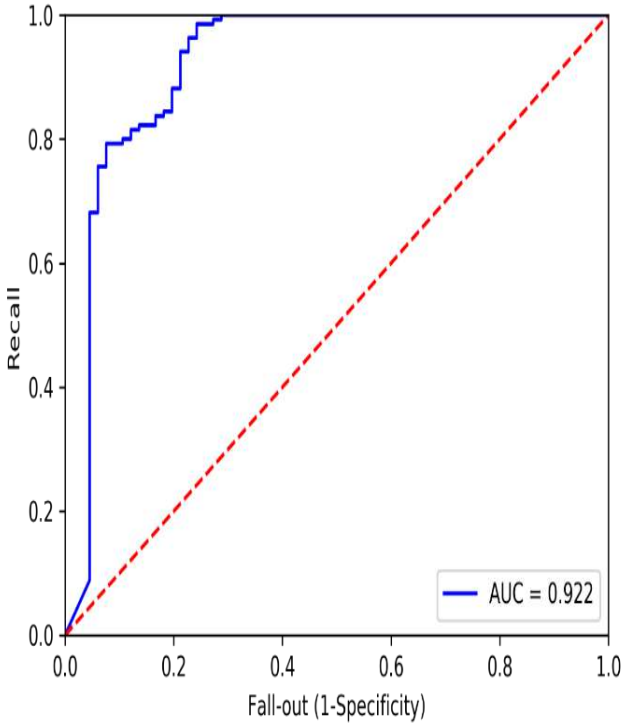


Figure 5: ROC Curve using Lung Cancer Data

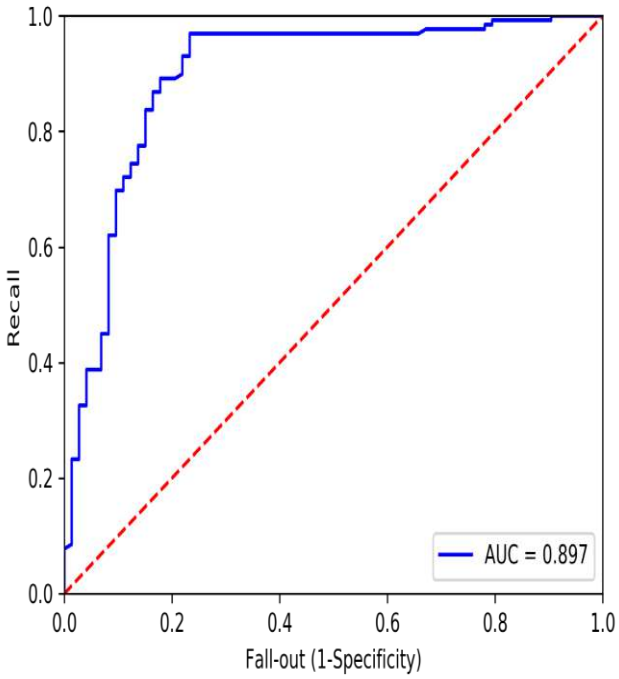


Figure 6: ROC curves using Diabetes Data

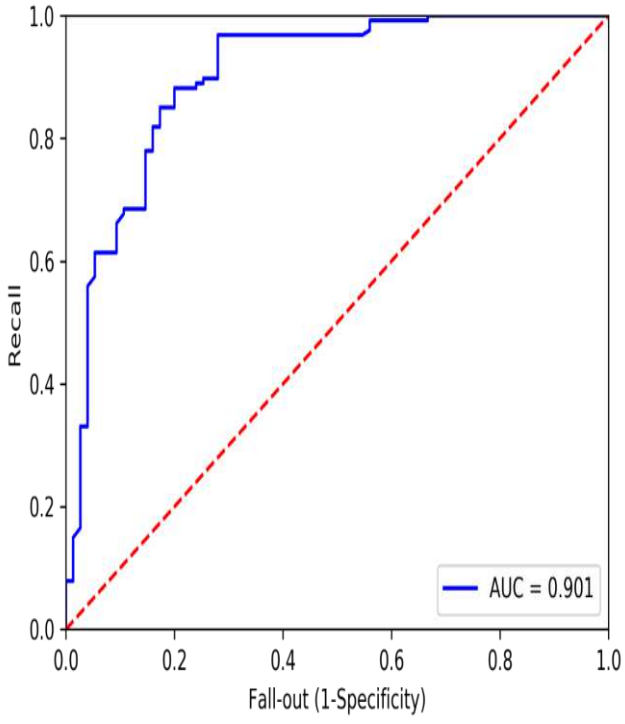


Figure 7: ROC curves using Breast Cancer Data

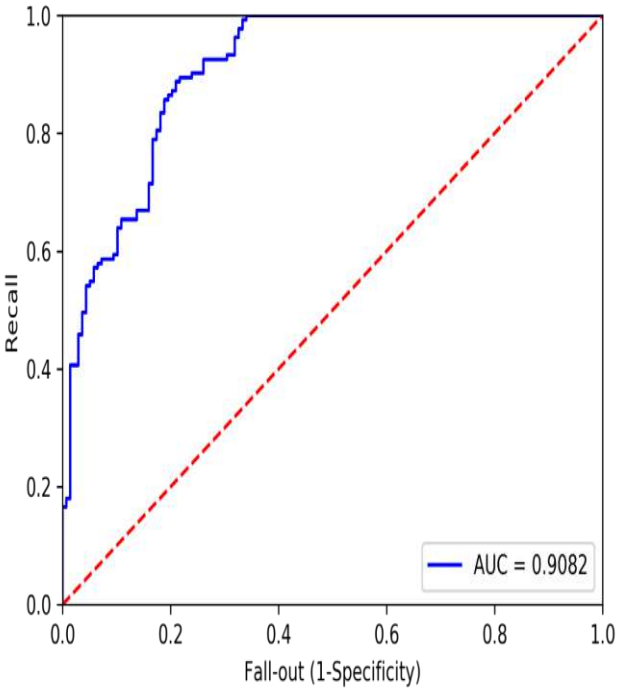


Figure 8: ROC curves using Heart Disease Data

To avoid the biased model and issue of data underfitting and overfitting, the datasets will be subjected to a K fold cross-validation process. To determine the performance of the constructed model the various measures like Precision, Root mean square, ROC, True positive rate (TPR), Recall, True negative rate (TNR) and F-measure are used. One of the most critical machine learning performance indicators is (ROC). The Area Under Curve (AUC) is a technique used to visualize the ROC.

Figures 5 ,6, 7 and 8 illustrate the ROC curve derived using the proposed method. Lung Cancer data and Diabetes had ROC rates of 0.922 and 0.897, respectively, whereas Breast Cancer data and Heart Disease have ROC rates of 9082 and 0.901 respectively.

Recall = True Positives / (True Positives + False Negatives)

The precision rate of the prediction result is computed as

Precision = True Positives / (True Positives + False Positives)

The true positive rate is calculated using the following expression.

True Positive rate= True Positives / True Positives + False Negative.

The F measure rate was computed using the expression below.

F-Measure = (2 * Precision * Recall) / (Precision + Recall)

TABLE 1: PERFORMANCE PARAMETER I ANALYSIS OF SG-GA

Data Set	**Pre**	**Rec**	**F M**	**TPR**
Diabetes	0.881	0.873	0.892	0.901
Lung Cancer	0.924	0.915	0.934	0.918
Breast cancer	0.901	0.892	0.912	0.912
Heart Disease	0.921	0.913	0.901	0.933

TABLE 2: PERFORMANCE PARAMETER II ANALYSIS OF SG-GA

Disease Data Set	**RA**	**KS**	**RMS**	**MAE**
Lung Cancer	0.922	0.223	0.285	0.115
Diabetes	0.891	0.331	0.215	0.216
Heart Disease	0.908	0.378	0.227	0.229
Breast cancer	0.901	0.262	0.229	0.247

TABLE 3: CONFUSION MATRIX ON DIABETES DATASET USING SG-GA

True Label	TNR	416	FPR	44
	FNR	31	TPR	277
	Predicted Label			

The performance statistics provided in table 1 are acquired when health care datasets are fed into the proposed system. The predictive ability of the proposed approach is further validated using simulation parameters such as kappa statistics (KS), mean absolute error (MAE), ROC area (RA), root mean square (RMS). The result obtained on these parameter using the proposed approach is shown in table 2.

The table 3 shows the confusion matrix obtained using the Diabetes data. It shows that out of 768 cases, 693 are accurately categorized. The confusion matrix generated with proposed approach on Breast Cancer dataset is shown in table 4. As per the confusion matrix, the Breast Cancer Data correctly predicts 261 out of 286 cases.

TABLE 4: CONFUSION MATRIX ON BREAST CANCER DATA USING SG-GA

True Label	TNR	125	FPR	13
	FNR	12	TPR	136
	Predicted Label			

TABLE5: CONFUSION MATRIX ON LUNG CANCER DATA USING SG-GA

True Label	TNR	197	FPR	13
	FNR	11	TPR	99
	Predicted Label			

Table 5 shows the confusion matrix obtained with Lung Cancer dataset using the proposed approach. As illustrated out of 320 test set occurrences, 296 instances are successfully predicted and only 24 are incorrectly forecasted**.** The confusion matrix generated for heart disease data using the proposed approach is shown in table 6 below. Given 270 test instances as input to the proposed model, approach predicted 245 instances correctly.

TABLE 6: CONFUSION MATRIX ON HEART DISEASE DATA USING SG-GA

True Label	TNR	166	FPR	15
	FNR	10	TPR	79
	Predicted Label			

TABLE 7: RESULT ON CHRONIC DISEASE DATA USING WRAPPER METHOD

Approach	Accuracy			
	Diabetes	Breast Cancer	Lung Cancer	Heart Disease

Decision Tree-GA	82.78	79.20	80.79	81.92
Naïve Bayes-GA	85.82	81.59	86.01	83.05
Random forest-GA	86.21	83.26	88.54	84.82
Logistic Regression-GA	84.91	86.63	86.49	85.05
SVM-GA	83.77	88.51	82.62	85.06
MLP-GA	81.42	84.63	83.03	78.92
Stacked Generalization-GA	90.21	91.25	92.50	90.74

As illustrated in the table 7 , the proposed Stacked Generalization model in combination with Genetic Algorithm have achieved 90.21% accuracy with diabetes dataset , 91.25% accuracy on Breast Cancer Dataset ,92.50% of accuracy with Lung Cancer dataset and 90.74% accuracy with heart disease dataset. Hence the accuracy achieved with proposed system is comparatively higher than the baseline version of algorithm which are applied in combination with Genetic algorithm.

V. CONCLUSION AND FUTURE SCOPE

In the proposed work novel stacked generalization based genetic algorithm is proposed to early forecast the chronic diseases such as lung cancer, diabetes, heart disease, and breast cancer. The experiment is undertaken by gathering the disease dataset from the UCI. Initially, best features from a chronic disease dataset are extracted using symmetrized feature selection algorithm. The proposed approach achieves stacking generalization through the use of a variety of algorithms, including SVM, Decision Tree, Random Forest, Naive Byes Classifier, and Multilayer Perceptron. The stacked generalization model is further optimized by genetic algorithm to boost the predictive ability of the model. Further result obtained using the proposed technique is compared with other wrapper based approach which are in turn optimized using genetic algorithm. The statistical findings of the analysis results indicates that proposed method outperforms than the baseline algorithms in terms of accuracy and various validation parameter. As part of the future work, the stacked ensemble model can be implemented by combining various machine learning algorithms which can further optimized using various evolutionary algorithms such as particle swarm optimization to further improve the accuracy of the machine learning model. These combination can be further applied on other chronic disease dataset such as breast cancer dataset, asthma dataset, Alzheimer disease dataset to predict the disease at the pre mature stage itself.

REFERENCES

[1] Jain, Divya, and Vijendra Singh. "Feature selection and classification systems for chronic disease prediction: A review." *Egyptian Informatics Journal* 19.3 (2018): 179-189.

[2] Gharibdousti, Maryam Soltanpour, et al. "Prediction of chronic kidney disease using data mining techniques." *IIE Annual Conference. Proceedings*. Institute of Industrial and Systems Engineers (IISE), 2017.

[3] Reddy, N. Satish Chandra, et al. "Classification and Feature Selection Approaches by Machine Learning Techniques: Heart Disease Prediction." *International Journal of Innovative Computing* 9.1 (2019).

[4] Heidari, Ali Asghar, et al. "Ant Lion Optimizer: Theory, Literature Review, and Application in Multi-layer Perceptron Neural Networks." *Nature-Inspired Optimizers*. Springer, Cham, 2020. 23-46.

[5] He, Jinbo, and Xitao Fan. "Evaluating the Performance of the K-fold Cross-Validation Approach for Model Selection in Growth Mixture Modeling." *Structural Equation Modeling: A Multidisciplinary Journal* 26.1 (2019): 66-79.

[6] Alby, S., and B. L. Shivakumar. "A prediction model for type 2 diabetes using adaptive neuro-fuzzy interface system." *Biomedical Research (0970-938X)* (2018).

[7] Saxena, Kanak, and Richa Sharma. "Diabetes mellitus prediction system evaluation using c4. 5 rules and partial tree." *2015 4th International Conference on Reliability, Infocom Technologies and Optimization (ICRITO)(Trends and Future Directions)*. IEEE, 2015.

[8] Gürbüz, Emre, and Erdal Kılıç. "A new adaptive support vector machine for diagnosis of diseases." *Expert Systems* 31.5 (2014): 389-397.

[9] Jac Fredo, Agastinose Ronickom, et al. "Classification of normal and knee joint disorder vibroarthrographic signals using multifractals and support vector machines." *Biomedical Engineering: Applications, Basis and Communications* 29.03 (2017): 1750016.

[10] Wu, Yunfeng, et al. "Classification of knee joint vibration signals using bivariate feature distribution estimation and maximal posterior probability decision criterion." *Entropy* 15.4 (2013): 1375-1387.

[11] Mirmozaffari, Mirpouya, Alireza Alinezhad, and Azadeh Gilanpour. "Heart disease prediction with data mining clustering algorithms." *Int'l Journal of Computing, Communications & Instrumentation Engg (IJCCIE), ISSN* (2017): 2349-1469.

[12] Guyon, Isabelle, et al. "Gene selection for cancer classification using support vector machines." *Machine learning* 46.1-3 (2002): 389-422.

[13] Le, Hung Minh, Toan Dinh Tran, and L. A. N. G. Van Tran. "Automatic Heart Disease Prediction Using Feature Selection And Data Mining Technique." *Journal of Computer Science and Cybernetics* 34.1 (2018): 33-48.

[14] Roffo, Giorgio, et al. "Infinite latent feature selection: A probabilistic latent graph-based ranking approach." *Proceedings of the IEEE International Conference on Computer Vision*. 2017.

[15] Xue, Bing, Mengjie Zhang, and Will N. Browne. "Particle swarm optimisation for feature selection in classification: Novel initialisation and updating mechanisms." *Applied soft computing* 18 (2014): 261-276.

[16] Lu, Xiaolu, et al. "Autoencoder based Feature Selection Method for Classification of Anticancer Drug Response." *Frontiers in Genetics* 10 (2019): 233.

[17] Konstantina Kourou et.al, "Machine learning applications in cancer prognosis and prediction" Computational and structural bitechnology Journal, Elsevier.

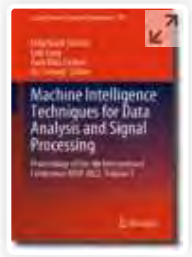
[18] P.Swathi Baby, T. Panduranga Vital ,"Statistical Analysis and Predicting Kidney Diseases using Machine Learning Algorithms" International Journal of Engineering Research & Technology (IJERT) ISSN: 2278-018, Vol. 4 Issue 07, July-2015,206-210.

[19] K.R.Lakshmi1, Y.Nagesh2 and M.VeeraKrishna3, "Performance Comparision of Three Data Mining Techniques for Predicting Kidney Dialysis Survivability", International Journal of Advances in Engineering & Technology, Mar. 2014, Vol. 7, Issue 1, pp. 242-254.

[20] AndrewKusiak, Bradley Dixonb, Shital Shaha,"Predicting survival time for kidney dialysis patients: a data mining approach", Elsevier Publication, Computers in Biology and Medicine 35, 2005, page no 311–327

[21] Venkataranganathan, A. V., R. J. Hariharan, and M. Roopa. "A Non-invasive Diagnosis of Early Stage Diseases through human nail using Neural Networks." Journal of Artificial Intelligence and Capsule Networks 4.1 (2022): 80-86.

[22] https://www.kaggle.com/datasets/cdc/chronic-disease



Machine Intelligence Techniques for Data Analysis and Signal Processing pp 187–199

Performance Analysis of Parametric and Non-parametric Classifier Models for Predicting the Liver Disease

Vijayalakshmi G. V. Mahesh, M. Mohan Kumar & Alex Noel Joseph Raj

Conference paper | First Online: 31 May 2023

Part of the Lecture Notes in Electrical Engineering book series (LNEE,volume 997)

Abstract

Liver plays a major role in maintaining the overall health of the body, so it is important to know the functionality or well-being of the liver. Various tests conducted on the blood samples reveal the health status of the liver; then, an appropriate treatment will

be suggested if tests indicate affliction. This paper presents a framework that uses the measurements obtained from blood samples after a series of tests and applies machine learning algorithms to predict the liver disease. The proposed work utilized the Indian Liver Patient Dataset with 583 samples with parametric and non-parametric machine learning algorithms for evaluating the performance of the methodology. The simulation results analyzed with the performance metric: sensitivity, specificity, precision, negative predicted value (NPV), false positive rate (FPR), false discovery rate (FDR), false negative rate (FNR), accuracy (ACC), F1-score, Matthews correlation coefficient (MCC), chi-square value, Kappa, Cramer's V, and confidence intervals (CI) of 95% indicated the better performance of the non-parametric classifier with reduced type I and type II errors.

Keywords

Parametric classifier **Non-parametric classifier**

Liver disease **Machine learning**

Performance evaluation **Confusion matrix**

References

1. Tanwar N, Rahman KF (2021) Machine learning in liver disease diagnosis: current progress and future opportunities. IOP Conf Ser Mater Sci Eng 1022(1). IOP Publishing

2. Wu C-C et al (2019) Prediction of fatty liver disease using machine learning algorithms. Comput Methods Prog Biomed 170:23–29

3. Sartakhti JS, Zangooei MH, Mozafari K (2012) Hepatitis disease diagnosis using a novel hybrid method based on support vector machine and

simulated annealing (SVM-SA). Comput Methods Prog Biomed 108(2):570–579

4. Singh J, Bagga S, Kaur R (2020) Software-based prediction of liver disease with feature selection and classification techniques. Procedia Comput Sci 167:1970–1980

5. Liu Y, Meric G, Havulinna AS, Teo SM, Ruuskanen M, Sanders J, Zhu Q et al (2020) Early prediction of liver disease using conventional risk factors and gut microbiome-augmented gradient boosting. medRxiv 2020

6. Naseem R, Khan B, Shah MA, Wakil K, Khan A, Alosaimi W, Uddin MI, Alouffi B (2020) Performance assessment of classification algorithms on early detection of liver syndrome. J Healthcare Eng

7. Singh AS, Irfan M, Chowdhury A (2018) Prediction of liver disease using classification algorithms. In: 2018 4th international conference on computing communication and automation (ICCCA). IEEE, pp 1–3

8. Naveau S, Raynard B, Ratziu V, Abella A, Imbert–bismuth F, Messous D, Beuzen F et al (2005) Biomarkers for the prediction of liver fibrosis in patients with chronic alcoholic liver disease. Clin Gastroenterol Hepatol 3(2):167–174

9. Lichman M (2013) UCI machine learning repository. University of California, School of Information and Computer Science, Irvine, CA

10. Dougherty G (2012) Pattern recognition and classification: an introduction. Springer Science & Business Media

11. Larose DT, Larose CD (2014) Discovering knowledge in data: an introduction to data mining, vol 4. Wiley

12. Han J, Pei J, Kamber M (2011) Data mining: concepts and techniques. Elsevier

13. Howell DC (2011) Chi-square test: analysis of contingency tables, pp 250–252

14. Kateri M (2014) Contingency table analysis. In: Methods and implementation using R, 1st ed. Editorial Advisory Board, Aachen, Germany

Author information

Authors and Affiliations

Department of ECE, BMS Institute of Technology and Management, Bangalore, India

Vijayalakshmi G. V. Mahesh

Department of ECE, Yenepoya Institute of Technology, Moodbidri, India

M. Mohan Kumar

Department of Electronic Engineering, Shantou University, Shantou, China

Alex Noel Joseph Raj

Corresponding author

Correspondence to M. Mohan Kumar .

Editor information

Editors and Affiliations

Department of Computer Science and Engineering, National Institute of Technology Raipur, Raipur, Chhattisgarh, India

Dilip Singh Sisodia

Department of Information and Communication Technology, University of Malta, Msida, Malta

Lalit Garg

Department of Electrical Engineering, Indian Institute of Technology Indore, Indore, India

Ram Bilas Pachori

Department of Mathematics, Indian Institute of Technology Indore, Indore, India

M. Tanveer

Rights and permissions

Copyright information

About this paper

Cite this paper

Mahesh, V.G.V., Kumar, M.M., Raj, A.N.J. (2023). Performance Analysis of Parametric and Non-parametric Classifier Models for Predicting the Liver Disease . In: Sisodia, D.S., Garg, L., Pachori, R.B., Tanveer, M. (eds) Machine Intelligence Techniques for Data Analysis and Signal Processing. Lecture Notes in Electrical Engineering, vol 997. Springer, Singapore. https://doi.org/10.1007/978-981-99-0085-5_16

.RIS .ENW .BIB

DOI	Published	Publisher Name
https://doi.org/10.1007/978-981-99-0085-5_16	31 May 2023	Springer, Singapore

Print ISBN	Online ISBN	eBook Packages
978-981-99-0084-8	978-981-99-0085-5	Intelligent Technologies and Robotics Intelligent Technologies and Robotics (R0)