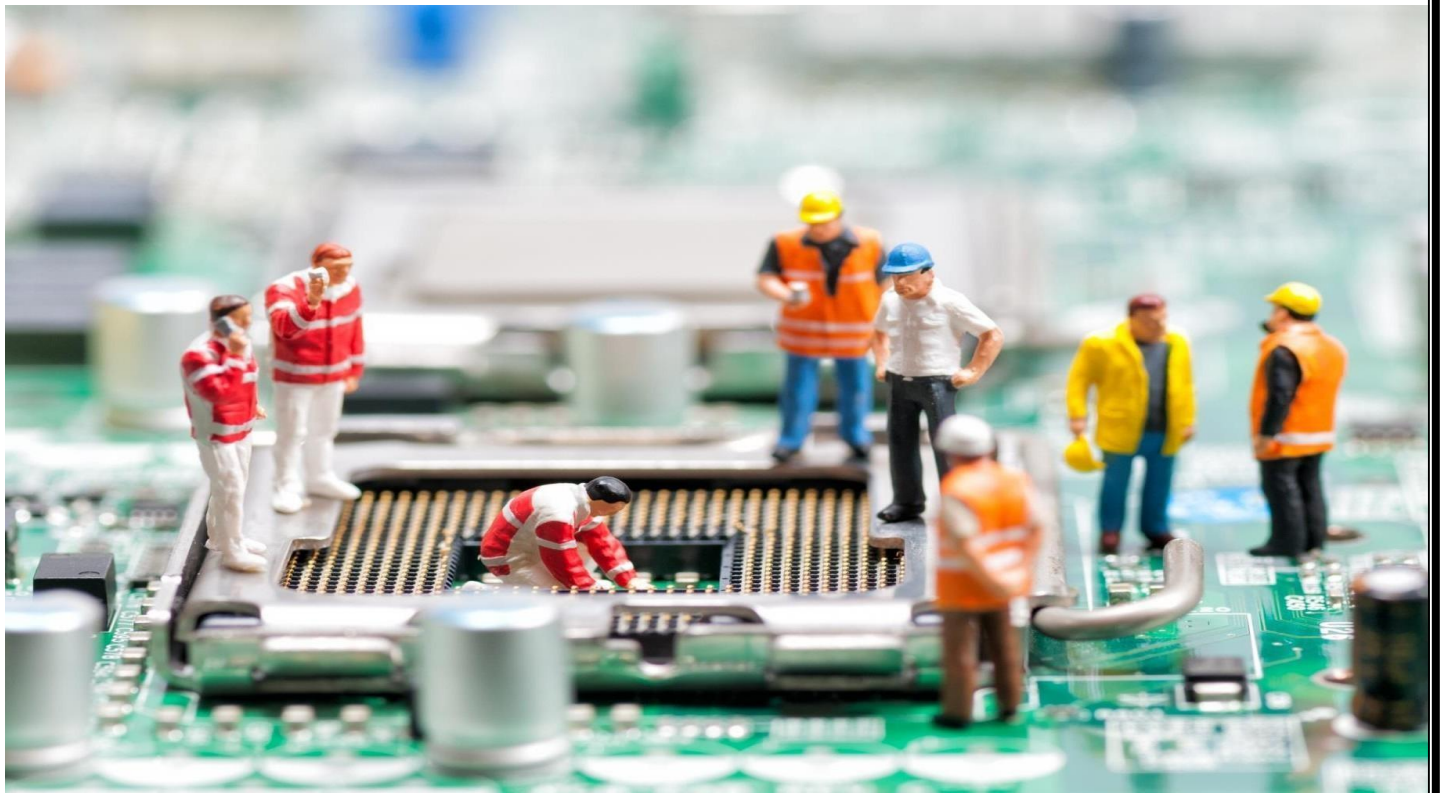# BMS
INSTITUTE OF TECHNOLOGY AND MANAGEMENT

Avalahalli, Doddaballapur Main Road, Bengaluru – 560064

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

RESEARCH COMPENDIUM

YEAR: 2020-21

Study and Analysis of BTED Error Correction Codes for Cryptography Applications

Ramesha M [1], Jeevan K M [2], Dankan Gowda V [3], Bharathi Gururaj [4], Sridhara S B [5]

[1] Assistant Professor, GITAM School of Technology, GITAM University, Bengaluru, India,
rameshmalur037@gmail.com

[2] Assistant Professor, GITAM School of Technology, GITAM University, Bengaluru, India,
jeevanjeevan77@gmail.com

[3] Assistant Professor, B.M.S. Institute of Technology and Management, Bengaluru, India, dankan.v@bmsit.in

[4] Assistant Professor, ACS College of Engineering, Bengaluru, India, bharathigururaj@gmail.com

[5] Professor, Vijaya Vittala Institute of Technology, Bengaluru, India, sridharasb1947@gmail.com

ABSTRACT

In recent days securing of transmitted data is an essential part of any communication system. There are several approaches with respect secure data like the utilization of cryptographic algorithms and other techniques. In this paper, Bit Transition Encoder and Decoder (BTED) error correction codes are analyzed for securing user data with the utilization of error correction codes (ECC) operations, such as point addition, point addition, point doubling, and point negative operation. Bit Transition Encoder and Decoder (BTED) error correction codes better than conventional error correction codes with several aspects such as area overhead in hardware, decoding time, and efficiency. The proposed techniques can be effectively utilized for memory applications. The error-correcting capability of these codes is also included for efficient implementation. This paper analyses the BTED Error Correction Codes for Cryptography Applications and compared the error correction capability of various error correction codes.

Key words: Cryptography, Error Correction codes, Data Security, Encoding, Decoding

1. INTRODUCTION

Encoding and decoding of data play a vital role in any wireless communication system. There are several encoding mechanisms are utilizing for the encoding of the data. The encoding process imposes redundant bits to secure the data in several formats. The imposing of redundant bits creates ambiguity for the decryption of data and also the possibility of error in the received data. By using a suitable error correction algorithm the ambiguity of recovering data can be minimized with the help of error correction codes operations. The commonly used error correction operations are point addition,

point doubling, and point negative operation. In point addition operation, the bits adding a point along an elliptic curve to itself repeatedly to obtain the desired results. Certain cryptographic algorithms are implemented based on the point addition mechanism. Point doubling is the operations specified for points of the elliptic curve to obtain accurate results from the encoded data. It should be the same as if we wanted to sum not two distinct but rather two equal points. Point negative operation is performed on the bits to accommodate a particular interval for the processing bit [1]. Figure 1 shows a generalized block diagram of encryption and decryption. As an instance, the input to encryption block is considered as plain text and key. The key is generated by the use of the data set. The data set consists of binary digits that pertain to the input but modified through by performing point addition, point addition, point doubling, and point negative operation. The received encrypted data is decrypted in the decrypted block. The decryption block consists of cipher text transferred data and key that is generated from the error correction operations. The once the algorithm extracts the plain text it will be compared with the original data for authentication purposes.
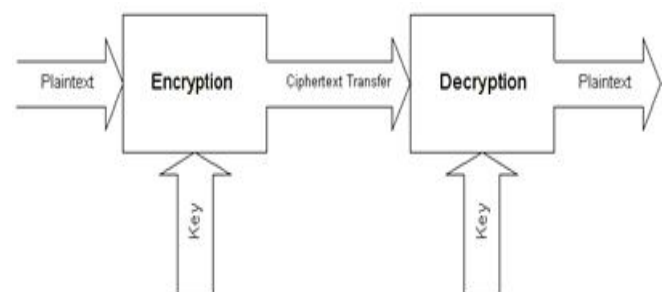

Figure1: Block diagram of encryption and Decryption

The overall organization of the paper as follows. In the second section, the ECC operations are explained with mathematical expressions and their corresponding curves. In the third section, Scalar multiplications of points are discussed. In the

fourth section, the BTED algorithm explained, and finally, the results and conclusion of the paper is described.

2. ECC POINT OPERATIONS
In this section, the computations of curve points for encryption and decryption operations are analyzed in detail and also how to compute the points for a negative number is discussed.

A.Point Addition:
In point addition operation, the bits adding a point along an elliptic curve to itself repeatedly to obtain the desired results [2]. Given 2 points on an elliptic curve, $J(x_1, y_1)$ and $K(x_2, y_2)$, then the addition of these points results in $L(x_3, y_3)$, which lies on the curve as depicted in Figure 2.

$$\lambda = [(y_2-y_1)/(x_2-x_1)] \text{ (mod p)} \qquad (1)$$
$$x_3 = [\lambda-x_1-x_2] \text{ (mod p)} \qquad (2)$$
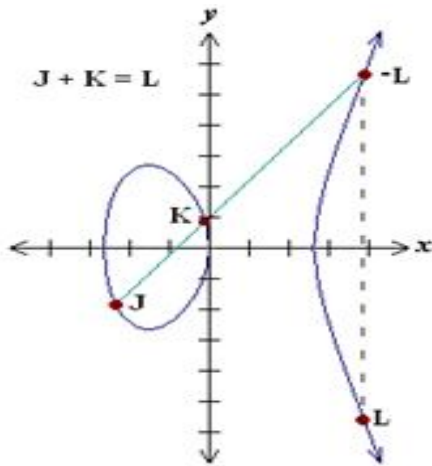$$y_3 = [\lambda(x_1-x_3)-y_2] \text{ (mod p)} \qquad (3)$$

Figure.2: Point addition operation on elliptic curve.

B.Point Doubling

Point doubling is the operations specified for points of the elliptic curve to obtain accurate results from the encoded data. Given a point $J(x_1, y_1)$ on an elliptic curve, point doubling i.e., $J(x_1, y_1) + J(x_1, y_1)$ yields $L(x_3, y_3)$, which lies on that curve as outlined in Figure.3.

$$\lambda = [(3x_1^2 + a) / (2y_1)] \text{ (mod p)} \qquad (4)$$
$$x_3 = [\lambda-2x_1] \text{(mod p)} \qquad (5)$$
$$y_3 = [\lambda(x_1-x_3)-y_2] \text{ (mod p)} \qquad (6)$$
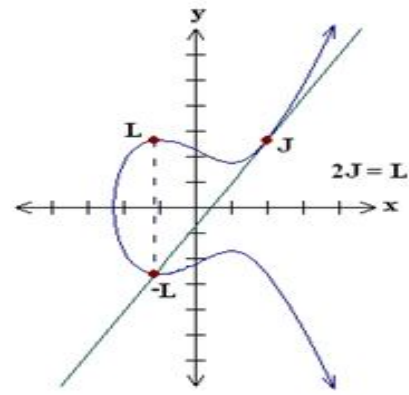

Figure.3: Point doubling operation on elliptic curve

C.Point Negative
Point negative operation is performed on the bits to accommodate a particular interval for the processing bit. Given a point $J(x_1, y_1)$ on an elliptic curve, to find $-J(x_1, y_1)$ is given by $J(x_1, p- y_1)$, as illustrated in Figure.4.
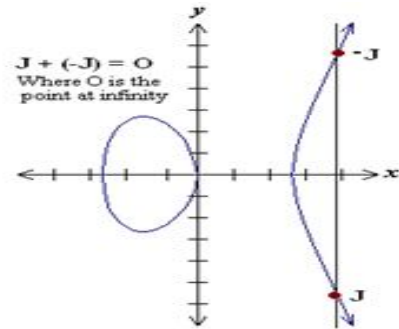

Figure.4: Point Negative operation on elliptic curve

3. SCALAR MULTIPLICATION OF POINTS
Given a point $P(x1, y1)$ on an elliptic curve, to calculate $Q(x2, y2) = k P(x1, y1)$, where k is any number, it requires frequent point additions and point doublings. The calculated points are stored in read-only memory for write and read operation. The volume of each memory location is 32 bits, and the total number of memory locations is 256 [3]. All 256 points are in the form of 16x16 matrixes, and each point size is 32 bits. The input information is replaced with the S-box value deployed on x and y coordinate value, as depicted in the below instance. Assume the input data is 4512, the primary digit in x-coordinate i.e., 4 and the last number in y-coordinate i.e., 2. The value 4512 is substituted by the fourth column and second-row values from S-box. The outcome of the S-box value is encrypted; employing BTED and encrypted data is broadcasted through wireless communication [4]. The received information at the receiver is decrypted by utilizing the syndrome calculator and error detector circuit. The S-box value is plotted as points on an elliptic curve. Depending on the input values, the S-Box values are substituted for subsequent encryption and decoding using BTED [5]. The figure 5 shows the hierarchical demonstration of elliptic-curve cryptography (ECC).
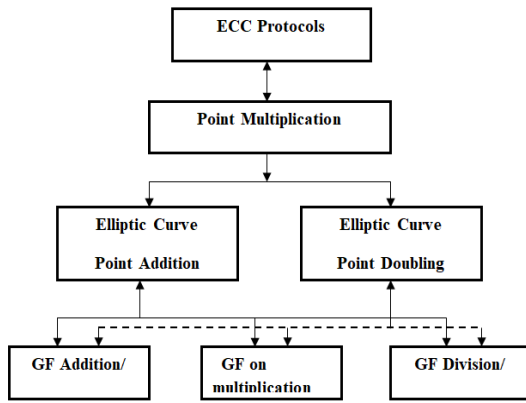
Figure.5: The hierarchical demonstration of elliptic-curve cryptography (ECC)

4. BTED ENCRYPTION AND DECRYPTION

The decoding is necessary to identify and correct errors in the word obtained from encryption output. Assume that original data bits are B3, B2, B1, and B0 and the redundant bits are C0 and C1. The bits C0 and C1 are obtained via the binary formula (XOR operation).

$C0 = B0$ xor $B2 = 1$ xor0 $=1$
$C1 = B1$ xor $B3 = 0$ xor1 $= 1$

Then assume now that Multiple Bit Upsets (MCUs) occur in bits $B3$, $B2$, and $B0$, the received redundant bits $C0$ and $C1$ are computed.

$$C_0' = B_0' \, xor \, B_2' = 0 \, xor \, 1 = 1$$
$$C_1' = B_1' \, xor \, B_3' = 0 \, xor \, 0 = 0$$

In order to detect these errors, the syndrome bits $S0$ and $S1$ are obtained as follows.

$$S_0 = C_0' \, xor \, C_0 = 1 \, xor \, 1 = 0$$
$$S_1 = C_1' \, xor \, C_1 = 0 \, xor \, 1 = 1$$

These results suggest that the error bits B2 and B0 are wrongly observed as the initial bits thus, these 2 errors bits not corrected. This instance describes that for a direct paired task, the number of even bit errors cannot be recognized. In the beginning, from data bits 'D' the obtained repetitive bits H4, H3, H2, H1, H0, and V01-V31 are formed. Using equation 13 and 14, the level condition bits H4, H3, H2, H1, H0 and the vertical error bits S3-S0 are assumed as follows:

$$\Delta H_4 H_3 H_2 H_1 H_0^1 = H_4 H_3 H_2 H_1 H_0^1 - H_4 H_3 H_2 H_1 H_0 \qquad (7)$$

$$S_3 - S_0 = V_0^1 \oplus V_0 \qquad (8)$$

Similar computations are executed for the rest of the vertical error bits. Here "−" represents the decimal integer subtraction. When S3 – S0 are equated to zero, then the stored on code word has only original data bits in the symbols, and there are no errors. When $\Delta H_4 H_3 H_2 H_1 H_0^1$ and S3 – S0are nonzero, then there is an error. Induced errors are identified in symbol 0. These errors are corrected by using the below equation (9).

$$D_{0 correct} = D_0 \oplus S_0 \qquad (9)$$

Earlier it was determined by considering error location; a parallel calculation can be used to differentiate a limited number of errors. In any instance, when these decimal computations are employed to distinguish errors, these errors can be determined in such a way that unraveling error is kept aside. The positioning technique of decimal error detection using the proposed structure is specified in the above equations and also shown in decrypted data. Initially, the extra horizontal bits H4H3H2H1H0 are obtained from the primary data bits as follows.

$$H_4 H_3 H_2 H_1 H_0 = D_3 D_2 D_1 D_0 + D_{11} D_{10} D_9 D_3 = 1010 + 1100 = 10110$$
$$H_4 H_3 H_2 H_1 H_0' = D_3 D_2 D_1 D_0 + D_{11} D_{10} D_9 D_{10} = 0111 + 1111 = 10110$$

Then, the flat syndrome bits *H4H3H2H1H0* can be obtained using decimal integer subtraction as follows.

$$\Delta H_4 H_3 H_2 H_1 H_0 = H_4 H_3 H_2 H_1 H_0' - H_4 H_3 H_2 H_1 H_0$$
$$= 10110 - 10010 = 00100$$

The decimal approximation of H4H3H2H1H0 is not equal to zero, and the errors are known and represented as symbol 0 or symbol 2. Thus, these flopped bits are placed in a precise area to utilize the vertical disorder bits S3 − S0 or S11 − S8. Thus, in decimal computation, the proposed system has higher flexibility for ensuring memory against M. Consequently, this is feasible for single and decimal slips and also for various types of errors per line may be studied for the proposed strategy irrespective of errors. In decimal one of the basics for these bits, H4H3H2H1H0 is zero. The 7-bit slips occur in symbol 0 and 2 simultaneously; the disentangling error can be declined in the sequence 1, 2, and 3, and are essential properties of the BTED algorithm. In this algorithm, all single-error and multiple errors are resolved in two sequential images. Phenomena of sequence 4 and 5 conferred are redressed, and various errors per column are identified by the even error bits.

Thus exhibit memory recollections from substantial Multiple Bit Upsets (MCUs). Then again, impacts of type 4 and 5, is important to identify this decoding slip based on these essential variables are achieved [6]. Data bits in symbols with decimal integer 0 and 2 are equated 2m – 1 and therefore the error computation in symbol 0 and 2 equal to 2m – 1 and the error computation in symbol 0 and 2. E.g., when m = 4, then the probability of the cryptography errors is shown in the below equation.

$$X_{\Delta H=0} = 4x \left(\frac{1}{2^4}\right)^2 x X_{BWMC} \approx 0.001 \qquad (10)$$

Similarly, If m=8, then the probability of the cryptography errors is obtained in using the below equation.

$$X_{\Delta H=0} = 4x \left(\frac{1}{2^8}\right)^2 x X_{BWMC} \approx 0.0000011 \qquad (11)$$

The plot of eight errors in data is shown in Figure.6. From this figure, it is assumed that memory as a regime that

contains few errors, and in the middle of these errors, there are no more than 3 bits. Thus it is not a major problem. In a comparative method the errors in symbol1 and symbol 2 are differentiated and rectified [7].
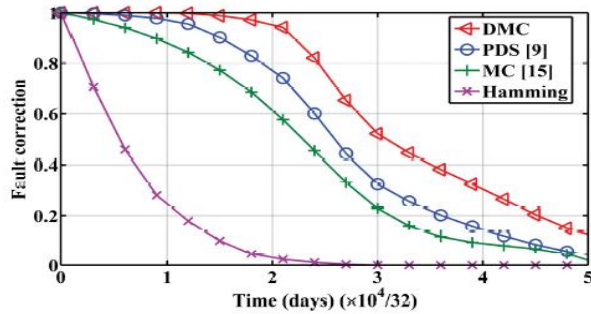

Figure.6: Time Difference security code plot for eight errors in data

5. RESULTS AND DISCUSSION

The total power of the Passive tag design is about 5mW, which incorporates a Digital baseband processor, memory, and peripheral devices. The area is smaller because the BTED method is included in the design. In program design, target specifications like input and output are considered. After software design, the Verilog HDL code is compiled and synthesized to verify the design syntax. Once the compilation and synthesis are completed, the Verilog HDL code is simulated to study the performance. The simulation was executed using Simulink software. The entire design is copied on to the FPGA Artix-7 FPGA Kit to examine its practicality once the simulation is completed.

There are several other parameters are taken to considerations for comparison purposes namely, total power, delay, and device utilization summary are indicated in Table 1 for different ECC's. This proposed research work has satisfactory performance in terms of power and speed compared to other ECC's.

Table 1: Error Correction capability BTED Technique

Error Correction Code used	Number of Information Bits used	Redundant Bits	Coding efficiency (β)%	Error Correction capability
Proposed work	32	36	52.9	05
Decimal Matrix Code	32	36	52.9	05
Punctured distinction set (PDS) codes	32	19	37.3	--
Matrix Code	32	28	46.7	02
Hamming code	32	07	19.9	01

6. CONCLUSION

Error rectification codes are employed to boost memory assurance and make the memory error-free. Various ECC's are employed to differentiate the event of an error and also corrected and the distinguished ones. On the other hand, the error detection ability and, therefore, the overheads differ given the codes used. The projected BTED system accomplishes recognition and correcting the errors effectively using the bitwise matrix code algorithm technique. Comparing with earlier works, the improved BTED can correct errors up to 5 bits in a symbol. The projected system demonstrates that it has a higher protection level compared to the large multiple cell upsets in the memory cells.

REFERENCES

1. S.Raza, **Secure communication for the Internet of Things - a comparison of link-layer security and IP sec for 6LoWPAN**, *Security and Communication Networks*, vol. 7, no. 12, pp. 2654–2668.
2. O. Srinivasa Rao, **Efficient Mapping Methods for Elliptic Curve Cryptosystems**, *International Journal of Engineering Science and Technology*, Vol. 2(8), 2010, ISSN: 0975-5462, 3651-3656.
3. Sanghyeon Baeg, Pedro Reviriego, Juan Antonio Maestro, Shijie and Richard Wong (2011), **Analysis of a Multiple Cell Upset Failure Model for Memories**. *ACM Transactions on Design Automation of Electronic Systems*, Vol. 16, No. 4, Article 45, 1084-4309.
4. B, Sridhara & M, Ramesha & Patil, Veeresh. **Adaptive Scheduling Design for Time Slotted Channel Hopping Enabled Mobile Adhoc Network.** *International Journal of Advanced Computer Science and Applications*.11. 10.14569/IJACSA.2020.0110333.
5. M. Ramesha, S. B. Sridhara, A. B. Anne Gowda, N. Anughna, and B. Gururaj, **Design and development of low power BTED cryptography algorithm on FPGA,** *International Journal of Advanced Trends in Computer Science and Engineering*, vol. 9, no. 4, pp. 4359–4362, 2020, doi: 10.30534/ijatcse/2020/27942020.
6. V. D. Gowda, S. B. Sridhara, K. B. Naveen, M. Ramesha, and G. N. Pai, **Internet of things: Internet revolution, impact, technology road map and features**, *Advances in Mathematics: Scientific Journal*, vol. 9, no. 7, pp. 4405–4414, 2020, doi: 10.37418/amsj.9.7.11.
7. Ramesha M, **FPGA Implementation of Low Power High Speed BTED Algorithm for 8 Bit Error Correction in Cryptography System**, *International Journal of Emerging Trends in Engineering Research*, vol. 8, no. 7, pp. 3893–3897, 2020, doi: 10.30534/ijeter/2020/158872020.
8. V. Dankan Gowda, A. C. Ramachandra, M. N. Thippeswamy, C. Pandurangappa, and P. Ramesh Naidu, **Modelling and performance evaluation of**

anti-lock braking system, *J. Eng. Sci. Technol.*, vol. 14, no. 5, pp. 3028–3045, 2019.

9. Dankan. V. Kishore, D. V. Gowda, 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.

Intelligent Traffic Light Control Using Image Processing- A Review

Dr. Ambika R[1], Dr. C S Mala[2], Shashikala J[3]

[1,2] *BMS Institute of Technology and Management, Bangalore, India*

Abstract - **Urban traffic congestion is increasing day by day. There is a need for the introduction of advanced technology and equipment to improve the state-of-the-art of traffic control. Traffic problems nowadays are increasing because of the growing number of vehicles and the limited resources provided by current infrastructures. The simplest way for controlling a traffic light uses timer for each phase. Another way is to use electronic sensors in order to detect vehicles and produce signal that cycles. We propose a system for controlling the traffic light by image processing. The system will detect vehicles through images instead of using electronic sensors embedded in the pavement.**

Keywords: Traffic Density, Edge Detection, Image processing, Image acquisition,

Image enhancement

1. Introduction

Traffic monitoring is important for road usage and management. Traffic information need to be collected and informed to the users. Various sensors to be deployed at appropriate places to find traffic parameters for updating traffic information. Image processing plays a vital role in capturing the density of traffic and the data may be considered for finding the density of traffic according to various time of the day. Traffic lights also play a major role in traffic control and regulation of the traffic on a daily basis in India. The traffic lights are of three colours: Red indicates stop, Yellow indicates wait and Green indicates go. Users are made to wait for the signal to change from red to yellow and then from yellow to green. The time that a commuter has to wait for is decided by the traffic signals. They are pre-programmed to wait for a fixed duration of time after every change in signal. It is independent of vehicle density on the road and remains constant. Sometimes particular road may be more crowded than other roads. During such situations, more time may be allocated to that road comparatively. This reduced traffic congestion and the system may be improved. There is a need for designing such intelligent system which allocated the time based on the vehicle density. In this paper, some of the methods

used for monitoring and controlling the traffic based on image acquisition and processing have been reviewed and compared.

2. Review of Related Work

PALLAVI CHOUDEKAR et. al presented us the use of image processing in traffic management [1]. This paper aims to ddistinguish the presence and absence of vehicles in road images, Signal the traffic light to go red if the road is empty, Signal the traffic light to go red if the maximum time for the green light has elapsed even if there are still vehicles present on the road. Authors have designed a hardware module for demonstration.

RAHISHET et. al proposed an intelligent traffic light control [2]. Authors have used image processing technique to capture the images and give more delay for the green light wherever required. In this system, basically, the waiting time for the vehicles on road with more density is reduced. Images for each lane are taken and processed simultaneously and a decision is passed as to which lane should be given how much amount of time and which should be the highest priority.

OMKAR RAMDAS GAIKWAD, et. al have proposed image processing-based traffic light control. They would be drawing lines of green color at an arbitrary distance. The reason for plotting only green colored lines or stripes is that we have to detect these strips through camera and there are very less number of vehicles that are painted in green. Camera would be placed at the signal which would be focusing on the lines drawn.

Singam Pardhasaradhi Reddy et. al., demonstrates that image processing is a far more efficient method of traffic control as compared to traditional techniques [6]. The use of our technique removes the need for extra hardware such as sound sensors. The increased response time for these vehicles is crucial for the prevention of loss of life. Major advantage is the variation in signal time which control appropriate traffic density using Image matching. The accuracy in calculation of time due to single moving camera depends on the registration position while facing road every time. Output is clearly indicated some expected results. It showed matching in almost every interval that were decided as boundaries like 10%, 35%, 68% etc.,

Aditya Kamath and Abhijeet Khanna presented a LabVIEW using image processing and particle analysis. This is to indicate increase in vehicular density in traffic-prone roads and alert same instantly [7]. This system focusses on the vehicular density which sets specific thresholds according to the time of the day. It also displays an application of computer vision for traffic flow management and road traffic analysis.

Some of the methodologies are discussed in the section.

3. Methodology

3.1 Image Acquisition

The image is captured by a webcam. It is then transferred to the computer via a USB cable. The image acquisition and further processing is done by using MATLAB.

Fig 1. Image Acquisition Flow Diagram

3.2 Image Processing

The image is captured by using a webcam placed at the road junction. It has the capability of taking images of all the roads meeting at the junction. The webcam is mounted on the DC motor. The speed of rotation of the camera is designed to be such that it is greater than the click-to-capture time of the camera. But the rotation of DC motor is controlled in such a way that only the images of one lane are captured. After whole process is done and the duration of the traffic light is set for that lane. The Dc Motor is rotated to an extent that the images of the next lane can be captured. This process is repeated for all four lanes. The acquired image is converted to grayscale image for further processing. The grey scale image is then converted to a binary image that contains only two colours, black and white. This image is known as the threshold image. The main purpose of thresholding the image is a radical reduction of information in order to simplify further processing. The thresholded image is then complemented for further image processing.

3.3 Image Enhancement

In this process the images are adjusted in such a way that the results are more suitable for further processing. In this, the obtained image is converted into a grayscale image.

3.4 Thresholding

Thresholding is transforming the grayscale image into black and white image (binary: white=1, black=0). The main purpose of thresholding is a radical reduction of information in order to simplify further processing. White colour is assigned to all the pixels that have luminosity greater than the threshold level and the others as black.

3.5 Edge Detection

Edge detection refers to the process of identifying and locating sharp discontinuities in an image. The discontinuities are abrupt changes in pixel intensities which characterize the boundaries of objects in an image.It filters out useless information, while keeping the important structural properties of an image. In this proposed system, canny edge detection technique is used. The boundaries of each image are found and the number of objects is calculated.

3.6 Object Counting

To count the objects present in the image, the close boundaries of the objects are identified. The exterior boundaries of the objects as well as the boundaries of holes inside these objects in the binary image are counted.

4. Changing the Duration of Traffic Light

The duration of the traffic signal is monitored as follows:

The camera controlled by Arduino microcontroller rotates in a clockwise direction and stops to take pictures of one lane. The clicked pictures are sent to MATLAB for image processing. The images are processed in MATLAB and the priority of giving green light for that lane is decided as per its traffic density. If there are more number of vehicles i.e very high density of vehicles, then green light is given for more duration say 90s, else if traffic density is moderate, green light duration is lower around 30s-50s,,and if traffic

density is very less then red light is given for that lane. Then the camera rotates to the next lane and same procedure is repeated. he same procedure is repeated for all four lanes. After all the lanes have given their green signals, the traffic light completes its one cycle of traffic monitoring and congestion control. This process repeats and the duration of green signal given to each lane keeps adjusting itself after every rotation of the camera.

5. Conclusion

This technique can be effective to combat the growing pressure of traffic on Indian roads. It uses image processing to estimate the density of vehicles on roads and regulates the traffic at fixed intervals of time. It is cost efficient and does not require the installation of complex machinery to monitor the traffic. Deploying this system will not only save the time consumed in waiting at traffic junctions, but will also conserve a lot of resources that are otherwise wasted.

References

1.   Pallavi Choudekar, Sayantibanerjee, M.K. Muju "Real Time Traffic Light Control Using Image Processing" (Ijcse),Vol. 2 No.1, Issn : 0976-5166, 2012.

2.   RAHISHET, APARAJITA SAHOO, APARNA INDORE, VAIBHAV DESHMUKH, PUSHPA US "INTELLIGENT TRAFFIC LIGHT CONTROL USING IMAGE PROCESSING" Proceedings Of 21st IRF International Conference, 8th March 2015, Pune, India, ISBN: 978-93-82702-75-7.

3.   Omkar Ramdas Gaikwad, Anil Vishwasrao, Prof. Kanchan Pujari, Tejas Talathi "Image Processing Based Traffic Light Control", (Ijsetr)Volume 3, Issue 4, April 2014.

4.   Prashant Borkar Et.Al., Predictive Traffic Light Control System: Existing Systems And Proposed Plan For Next Intersection Predition, International Technology Research Letters, Volume-1 Issue-1 2012.

5.   Ben Ahmed Mohamed, et. Al., A Statistical Multiplexing Method for Traffic Signal Timing Optimization in Smart Cities, TELKOMNIKA Indonesian Journal of Electrical Engineering Vol. 15, No. 1, July 2015, pp. 107 ~ 113 DOI: 10.11591/telkomnika.v15i1.8092

6.   Singam Pardhasaradhi Reddy et. al., "Design of Traffic Control based on Digital Image by Intelligent Systems", International Journal of Innovative Research

in Science, Engineering and Technology, Vol. 6, Issue 4, April 2017.

7.   Aditya Kamath and Abhijeet Khanna, "Image Processing and Particle Analysis for Road Traffic Detection" International Journal of Computer Applications (0975 – 8887) Volume 55– No.2, October 2012.

8.   Ambika R et.al., " Design and Development of Real Time Vehicular Crash Detection System", Grenze International Journal of Engineering and Technology, Special Issue, PP. 6-10, 2018.

COMPARATIVE ANALYSIS AND STUDY OF SELF DEFENCE MODULES FOR SAFETY OF WOMEN

Ambika R[1]

[1]BMS Institute of Technology and Management

[1]ambikar@bmsit.in

Abstract: *Women all over the world are facing and even subjected to unethical physical harassment. Security for women is still a major issue as the number of crimes and harassment over women and girls is increasing day-by-day. In this age of technology, mobile phone is one of the gadgets that almost everyone like and uses to keep in touch with family and friends. All they need is a device that can be carried everywhere easily. This proposed work deals with a quick responding, cost protection system for an individual and especially for women using which a woman in anguish can call for help just with the press of a button on this smart gadget. Self Defence module for women safety is like a Smart Watch for Women safety. It has the ability and potential to help women with technologies that are embedded into a compact device. It is specially designed for women safety and protection. It has a control button that will be used by women to inform nearby police when they are in distress.*

Keywords : Android, Safety, GSM, GPS

1. INTRODUCTION

Even in this modern era, women are feeling insecure to step out of their house because of increasing crimes in our country. The corporate and IT sector are currently in boom. Many women are working in corporate even in night shifts. There is a feeling of insecurity among the working women. The proposed device is more like a safety system in case of emergency. This device can be fitted in a jacket (similar to a blazer for women). It is an easy to carry device with more features and functions. The emergency push button is held to one of the buttons of the jacket. The main purpose of this device is to intimate the parents and police about the current location of the women. A GPS system is used to trace the current position of the victim and a GSM modem is used to send the message to the predefined numbers. There are several applications that reduce the risk of sexual abuse by sending SMS but in our model we also provide an audio circuit which is more useful for physically challenged people.

2. EXISTING WORK

Moser, c. and c. mcilwaine discussed a voice keyword recognizing app to recognize the user and activate the app functionality even when the mobile keypad is locked [1]. The GPS module tracks the longitude and latitude to trace an exact location of a user and sends the prestored emergency message including location to the registered contact numbers. The Audio Recording module starts the recording of the conversation for five minutes and stored as evidence. The message goes in queue if any network problem and send when network gets available. A notification is generated for successful delivery of the message. Also, user can select contact through voice-based contact list and make a call.

Richard Hill et. Al., [2] proposes a Spy Camera Identification and Women Attack Rescue System (SCIWARS) app which consist of two modules. The first module act as an intelligent alert system which detects the infrared rays coming from every Night-vision hidden cameras placed in changing rooms, hotels rooms etc., and also informed the user about unsafe place through message. Now it is the user responsibility whether to register a complaint or not by forwarding the notification with the location to legal authorities such as Police. The second module will get activated by pressing any key continuously which will provide the help to the victim from physical attack in unsafe situation. It sends the emergency message containing location to register contacts. It also records the voice and captures the images of the surrounding for 45 seconds. This information also stored in secret location of mobile for future evidences.

An android app to provide security at two different situations is proposed in reference [3] by Mughal, e. and k. krause. The First module provides security to Women at Emergency Situations propose a Save Our Souls (SOS) app to provide the security on a single click of SOS button for the women travelling at night or alone. No need to unlock the screen, instead by just pressing the power button it directly triggers the application to run at the background, to send the emergency message including the location in the form of latitude and longitude to the registered contacts. The second module proposes an android based home security system that provides security of house belongings and Senior Citizen in the user absence. Since the security of senior citizen is always a concern with increasing number of robbery incidents. This app informs the user about an attempt of intrusion activity at home through a message and a feedback SMS triggers an alarm in the house. The minimum requirement is the android mobile, a hardware circuit embedded with a switch and GSM modem that are connected to the door. When an intruder tries to open the door, the switch triggers an interrupt for the microcontroller to activate the GSM modem to send warning SMS to store registered number in the modem. At the receiver end, the application pop up the menu frequently for user attention. If the user fails to acknowledge in the defined time interval, then the automatic positive acknowledgement message will be sent to the remote GSM modem which in turn interrupt the microcontroller for an alarm.

Rathmell proposed a portable device as a belt which is automatically activated base on the pressure difference crosses the threshold in an unsafe situation[4]. A GPS module tracks the location and sends the emergency messages to three emergency contacts every two minutes with updated location through GSM. The system also activates the screaming alarm that uses a siren, to call out for help and also generates an electric shock to harm the attacker for self-defense which may help the victim to escape. The device mainly consists of micro controller on the ATMega328 board which programmed using the ARDUINO programming language.

Paper [5] proposes the women security device called as "Suraksha" which is easy to operate the device. This device can be activated through- voice command, Press a switch key and shock (i.e. when the device is thrown with force, a force sensor used to activate the device). In an emergency situation, it will send the message including instant location to the police, via the transmitter module and registered numbers via a GSM module. Currently, the work is under process to embed it in jewellery, mobile or another carrier like belt etc. It can play a major role in the propose projects where all the police stations are connected and share the criminal records, crime investigating cases etc.

An extended vehicle tracking system to track the vehicle based on GPS with that it also provides the safety through an emergency button kept under the vehicle seat using GSM. As the increasing economic growth rate of a country, many companies are establishing their setup in the nearby region of the cities. Since the security of women employees' inside the private transportation is the companies' responsibility. In the unsafe situation, an employee needs to press the emergency button to activate the device Teltonika-FM1100. It, in turn, enables simultaneously the Android device used to capture the images inside the vehicle and the GPS system which track the vehicle position in the form of

latitude and longitude. An alert message including the location is sent a card to the company special team and nearby police station through GSM SIM. After that, it is the responsibility of police squad and company team to handle the situation.

Paper [7] proposed a women safety system which used a microcontroller ATMEGA 328P. Here the MEMS sensor is used to sense any mishappening with women according to the extraordinary movement of the body. If in any case, MEMS sensor is unable to sense the mishappening then the switch in the watch can be pressed manually to indicate any mishappening. As soon as any mishappening is detected by the sensor the same is indicated to the controller. Upon receiving the signal the controller starts generating shock waves through shock wave circuit and at the same time a message containing the location of the victim obtained through GPS is transmitted to the relative or friend whose number is already in the program.

3. IMPLEMENTATION

Entire work is divided in to three main modules, namely GSM, GPS and Panic Sensor. As presented in the block diagram, the proposed module will have emergency switch, which can be pressed when the user is in danger. Emergency switch is interfaced with GPS and GSM modules to send a message of the location of the user to the registered numbers. Bluetooth module will also be activated which produces a call for help in case the user is not able to call for help. In case the user is not able to voluntarily press the emergency button, there is also a panic sensor present in the device which can be used. The panic sensor consists of a temperature sensor and an LDR. In case of fear or panic, the temperature sensor can sense the varying body temperature due to production of hormones when in fear. The LDR helps in detecting if a person comes too close to the user. This device calls for help and sends a message only after LDR is continuously sensed for a certain number of times. Hence it is easier to find and save the user when in danger.
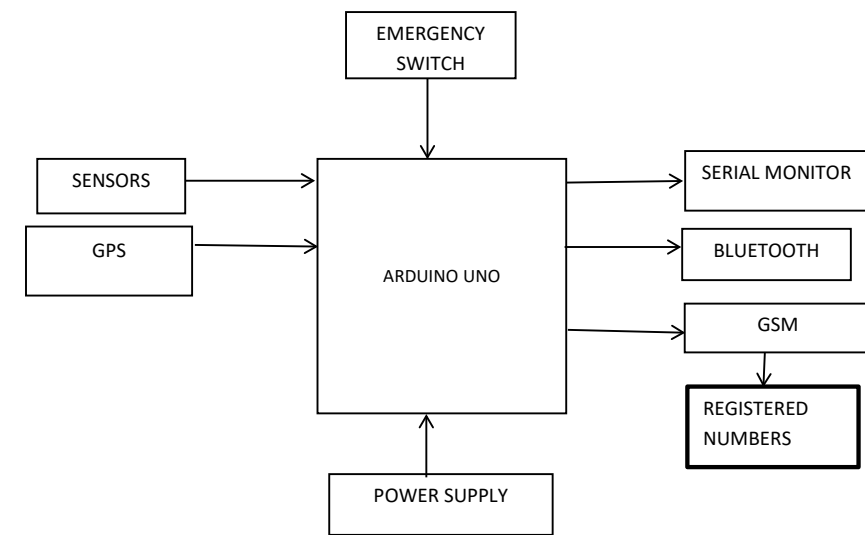

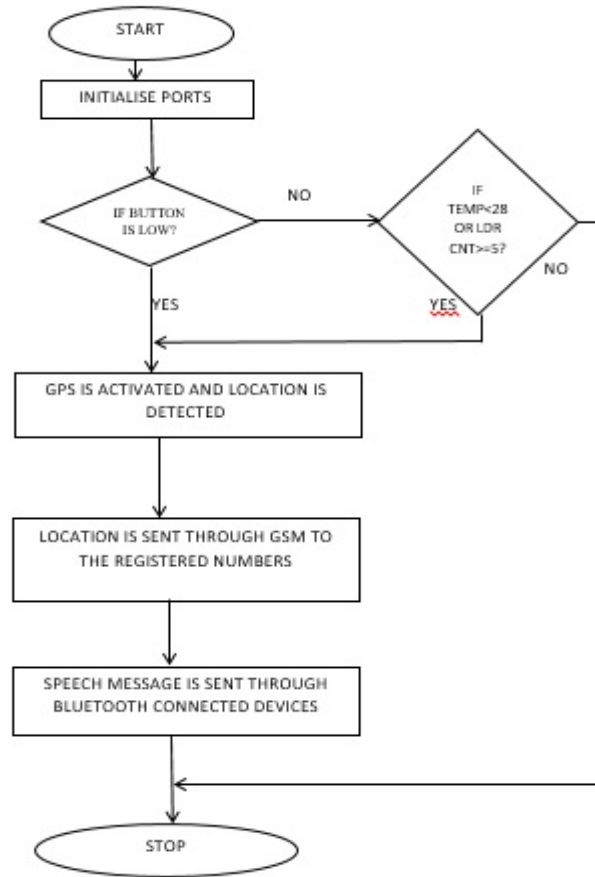**Figure 1. Block Diagram of the Proposed Module**

Figure 2. Flowchart of the Proposed Module

This device directly gets connected to the satellite through GPS when activated. Then the location is transferred through the GSM, it also contains a shock mechanism to produce non-lethal electric shock in emergency situations to deter the attacker. Being safe and secure is the demand of the day. Our effort behind this project is to design and fabricate a gadget which is so compact in itself that provide advantage of personal security system. This design will deal with most of the critical issues faced by women and will help them to be secure. Women's security is a critical issue in current situation. These crimes can be brought to an end with the help of real time implementation of our proposed system.

4. APPLICATIONS AND ADVANTAGES

Proposed module may be used in Security appliances, Safety of women and used as a legal evidence of crime with exact location information for prosecution.

Main advantages of the proposed module may be Sophisticated security, Monitors all hazards and threats, alert message to mobile phone for remote information, Mobile number can be changed at any time, can be used to prevent incidents.

CONCLUSION

Our effort behind this project is to design and fabricate a gadget which is so compact in itself that provide advantage of personal security system the emergency response system which is helpful for women in the incidents of crime. It is low cost system which can store the data of the members in the particular locality and provide immediate alert in case of crime against women. This provides women security. Being safe and secure is the demand of the day.

REFERENCES

[1] *Moser, c. and c. mcilwaine (2006), "Latin American urban Violence as a development concern: towards a framework for Violence reduction", World Development, Vol. 34, no. 1, pp.89-112.*

[2] *Hill, r., J. temin and L. Pacholek (2007), "Building Security where there is no Security", Journal of Peacebuilding and Development, Vol. no. 2, p. 38-51.*

[3] *Mughal, e. and k. krause (2009), "closing the gap Between Peace operations and Post Conflict insecurity: towards a Violence reduction agenda", International Peacekeeping, Vol. 16, no. 1, pp.136-150.*

[4] *Rathmell, a. (2009), "Security and Justice development – what next?", Journal of Security Sector Management, Vol. 7, p no. 2.*

[5] *Charlotte Bunch and Roxanna Carillo, "Global Violence against Women: The Challenge to Human Rights and Development" in Michael Klare and YogeshChandrani (eds.), World Security: Challenges for a New Century, third edition (New York: St. Martin"s Press, 1998), p. 230.*

[6] *Beth Woroniuk, "Women"s Empowerment in the context of Human Security", Bangkok, Thailand, December 7-8 1999.*

[7] *Reardon, op. cit., "Feminist Concepts of Peace and Security," p. 139.*

[8] *Susan McKay, "Gender Justice and Reconciliation," Women"s Studies International Forum, vol.23, No. 5, Sept. 2000.*

Technical Article

Nipple Localization in Automated Whole Breast Ultrasound Coronal Scans Using Ensemble Learning

Alex Noel Joseph Raj[1] (iD), Ruban Nersisson[2],
Vijayalakshmi G. V. Mahesh[3], and Zhemin Zhuang[1]

Abstract

Nipple is a vital landmark in the breast lesion diagnosis. Although there are advanced computer-aided detection (CADe) systems for nipple detection in breast mediolateral oblique (MLO) views of mammogram images, few academic works address the coronal views of breast ultrasound (BUS) images. This paper addresses a novel CADe system to locate the Nipple Shadow Area (NSA) in ultrasound images. Here the Hu Moments and Gray-level Co-occurrence Matrix (GLCM) were calculated through an iterative sliding window for the extraction of shape and texture features. These features are then concatenated and fed into an Artificial Neural Network (ANN) to obtain probable NSA's. Later, contour features, such as shape complexity through fractal dimension, edge distance from the periphery and contour area, were computed and passed into a Support Vector Machine (SVM) to identify the accurate NSA in each case. The coronal plane BUS dataset is built upon our own, which consists of 64 images from 13 patients. The test results show that the proposed CADe system achieves 91.99% accuracy, 97.55% specificity, 82.46% sensitivity and 88% F-score on our dataset.

Introduction

Amongst the different forms of cancers that affect women, breast cancer is the most commonly diagnosed around the world. It contributes to 11.8% of cancer deaths globally.[1] Since the availability of healthcare services and awareness of cancer is not state-of-the art everywhere, in many cases cancer is not detected at an early stage. Thus, patients undergoing treatments normally suffer from advanced stages of cancer, which greatly reduces their survival rate.[2] Early diagnosis through regular screening has brought down the mortality rates by 40% to 50%.[3] Various screening methods such as thermography, magnetic resonance imaging, mammograms, and ultrasound scans are available for cancer diagnosis. Amongst these, mammograms are considered as the standard due to their high sensitivity. However, in the case of patients with radiologically dense tissues in their breasts, the sensitivity of 2D mammograms has been observed to drop down to 48%, thus leading to a risk of false diagnosis of malignant tumors.[4] Multimodal scans that use a combination of various scanning methods for the accurate detection of malignant tumors reduce this risk significantly.[3] Automated Whole Breast Ultrasound (AWBUS) is a popular technology used for early detection

of cancer. It is either used as an independent screening method or as a part of multimodal scans.

Kotsianos-Hermle et al. compared and analyzed the detection ability of handheld ultrasound scan, mammography and Automated Whole Breast Ultrasound (AWBUS) of 107 breast lesions. The authors reported that AWBUS has better efficiency and can be considered as an adjunct for mammography. A study based on the size of the lesion reported a strong correlation between mammographic study and AWBUS study, and it was reported that the sensitivity of ultrasound is much better than the mammographic studies.[5] The advantages of having an AWBUS scan are: (a) it provides higher reproduction of sonographic reports; (b) it introduces less operator-related variability and (c) it

[1]Shantou University, Shantou, Guangdong Province, China
[2]Vellore Institute of Technology, Vellore, Tamil Nadu, India
[3]BMS Institute of Technology and Management, Bangalore, Karnataka

Corresponding Author:
Alex Noel Joseph Raj, Key Laboratory of Digital Signal and Image Processing of Guangdong Province, Department of Electronic Engineering, College of Engineering, Shantou University, Shantou, Guangdong 515063, China.
Email: jalexnoel@stu.edu.cn

performs faster image acquisition as compared to Hand-Held Ultrasound (HHUS).[6]

In AWBUS, coronal, sagittal, and transverse are the three major views for effective analysis of breast lesions. Amongst them, the coronal view gives better reliability and short reading time compared to a transverse view. Likewise Chae et al.[7] reported that coronal views present more information of breast anatomy with less number of images and are better suited for lesion detection in automated breast ultrasound (BUS) images.

With the advent of the machine and deep learning techniques, the automated algorithms have evolved for the segmentation of lesions in BUS images. Yuan Xu et al., proposed a CNN based ultrasound breast image segmentation algorithm, where two CNN networks are implemented to provide segmentation and classification. The 3D images obtained from three orthogonal planes were used by CNN-I to segment the lesion followed CNN-II which provided a classification of the segmented regions either as skin, gland, mass or fat. The authors reported a high score of more than 80% for performance metrics such as accuracy, precision recall and Fscore.[8] Lipismita et al., presented a novel clustering method known as Multi-scale Gaussian Kernel induced Fuzzy C-means (MsGKFCM) where the lesion margin is accurately detected by Multi-scale Vector Field Convolution (MsVFC). The overall segmentation accuracy is improved by avoiding local solutions. It is achieved using multi scale clustering process obtained by modifying the objective function of the conventional GKFCM.[9]

A recent survey by Mie Xian et al., presented the challenges faced in the segmentation of breast lesions due to speckle noise followed by the discussion of the various evaluation metrics adopted for segmentation. The authors concluded that machine learning approaches have a large potential in ultrasound image segmentation which can aid accurate diagnosis.[10]

Localization and segmentation of Nipple Shadow Region (NSA) is an important process since (a) by determining the location, the nipple region serves as a reference point to understand the growth characteristics of a suspected lesion during its examination over a period of time, and (b) it denotes a point or a region that can be documented across different modalities (MRI, CT, X-ray) to refer to a specific lesion, thus facilitating the radiologist in determining its malignancy.[11] The increase in the false positive rate in the localization of NSA is the major challenge in improving the sensitivity of the nipple detection in the BUS images. To tackle this challenge, many Computer-Aided Detection (CADe) tools are employed.[12,13]

Mediolateral Oblique (MLO) and craniocaudal view (CC) are **two** standard mammographic projections. In these two views, MLO is used to depict most of the muscles associated with the breast including the pectoral muscle and is mainly used to capture the upper outer quadrant of the breast. CC presents the top-bottom view of the breast which is used to visualize the medial part and the external lateral portion. Mendez et al., developed an algorithm with combined features to detect the nipple from both MLO and CC views of the mammographic images. The features consist of the maximum height of the breast border, maximum gradient, and maximum second derivative of the gray levels across the median-top section of the breast.[14] Jas et al.[15] proposed a heuristic approach to search the NSA in mammogram images on both MLO and CC views. Petroudi and Brady[16] addressed a way to segment the "fat-band" location through curvature changes, which helped in localization of the nipple region. Furthermore, Mustra et al., have defined the importance of craniocaudal projections in nipple detection in mammogram images.[17] Popli et al.[18] have discussed the best positions for breast scans to get an accurate diagnosis and Sweeney et al.[19] reviewed various quality criteria for craniocaudal projection.

Most of the methods suggested in the literature concentrate on identifying the nipple location from MLO or CC views of mammograms and few studies have listed the advantages of automated ultrasound for efficient localization of NSA. Wang et al., proposed a method to determine the nipple position in a 3D breast ultrasound (BUS) images using a Hessian-based filter. The method identifies the central line through the NSA which is the dark region formed below the nipple and area along the coronal plane of the BUS image. This method requires all the slices of the 3D scan in order to detect the NSA in the form of a tube-like structure.[11] Chae et al., proposed a method for segmenting the nipple in Digital Breast Tomosynthesis (DBT) images using the shape and location of the nipple on the breast boundary. This method takes advantage of the protruding nature of the nipple and uses a square mask to detect the position of the visible nipple. These methods return a single coordinate as the nipple point for a 2D slice and do not identify the boundary that separates the nipple from the rest of the breast.[20] Recently Zhuang et al.[21] employed a modified UNet model referred to as the GRA-UNet to segment the nipple region in AWBUS images.

In this paper, we address a fully automated CADe model to accurately detect the NSA from the coronal scans of the breast ultrasound images. Initially, we extract shape features by employing an iterative sliding window across the BUS images using Hu Invariant Moments and later texture features through Gray-level Co-occurrence Matrix (GLCM). Then the features are concatenated and presented to an Artificial Neural Network (ANN) to infer the probable NSA regions. Later, we extract the contour features from the probable NSA regions which include fractal dimensions, edge distance, and contour area and feed them to an SVM classifier to predict the true nipple region.

The rest of the paper is organized as follows: Section II discusses the methodology and provides detailed information about Pre-Processing, which includes the need for plane transformation from sagittal to coronal view, the use of
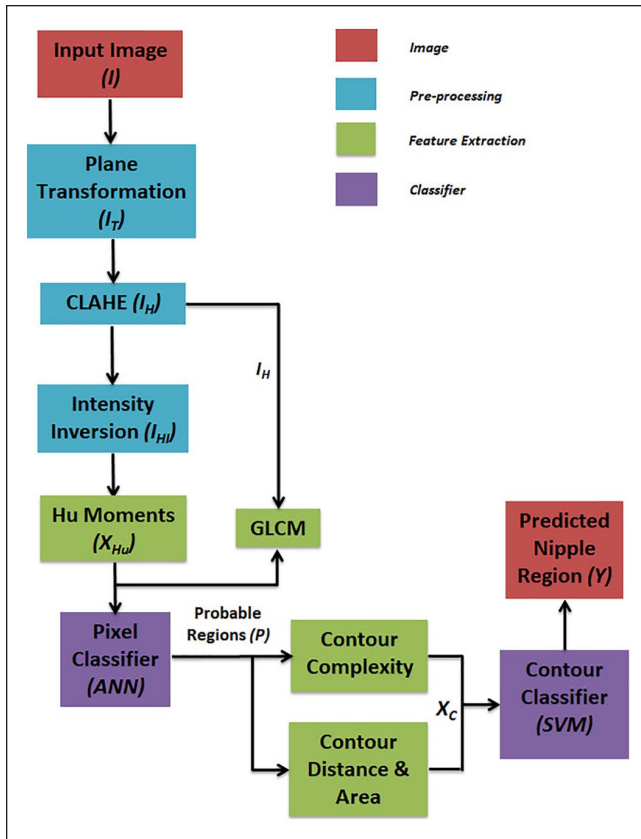
Figure 1. The block diagram of the system.

derived from the sagittal scans of 13 different patients. The sagittal scans are stacked together and sliced along the XY plane at different depths to obtain the coronal views of the breast as shown in Figure 2.

The images from the dataset are pre-processed to transform them suitable for extracting efficient shape and texture features. These features are then concatenated and given to the ANN classifier to predict probable NSAs. Further, the contour features are selected to represent the probable NSAs and later to train SVM classifier to obtain the true NSA regions. A detailed description of the proposed methodology is presented below.

Image Pre-processing

The image pre-processing is carried out in three steps: plane transformation, contrast limited adaptive histogram equalization (CLAHE), and intensity inversion

Plane transformation. Normally AWBUS scans provide Ultrasound (US) image slices (I) along the sagittal plane (XZ axis). Each scan contains N number of slices, where N is directly related to the resolution of the AWBUS scanner and area of the breast under examination. Figure 3 illustrates a few US slices where the area marked in red shows the NSA as a dark tube-like structure seen horizontally across the image. Although on an average each AWBUS scan contains 700 to 800 US image slices, the segmentation of NSA along the sagittal plane has the following problems:

- The nipple region is visible only in few slices, which makes it difficult to acquire sufficient data to train the model that can automatically perform segmentation. This is because the NSA appears as a tube along its length, thus appearing only in a few slices acquired along the sagittal plane.
- Due to the cylindrical nature of the NSA region in three-dimensional perspectives, out of the 800 slices acquired along the sagittal plane, the slices that only pass within the diameter of the NSA region will help in localization of the nipple region. Also, processing sagittal slices captured away from the NSA diameter will be redundant as no additional information is present.
- Due to characteristics of the acoustic waves, regions apart from the nipple -such as lesions- cast sonographic shadows in US scans. This leads to the emergence of dark regions that appear similar to the NSA as shown in Figure 3. These similar regions will present false localization when directly subjected to any segmentation algorithm.

CLAHE for denoising, and the intensity inversion for effective extraction of shape and texture features. Section II also addresses the feature extraction process and explains the use of Hu moments and GLCM to obtain shape and texture features respectively. The methodology section also presents the use of ANN in determining the probable NSA regions which is followed by SVM technique. SVM is employed for classifying the true NSA regions from probable NSAs. Section III presents qualitative and quantitative results of the proposed technique, followed by section IV where the proposed technique is compared to other state-of-the-art segmentation techniques. Finally, the conclusion summarizes the contributions.

Methodology

The algorithm to segment NSA from the AWBUS scans has the following stages: Image Pre-Processing, Shape and Texture Features Extraction, prediction of probable NSA, and determination of correct NSA regions using SVM classifier. The workflow of the entire process is illustrated in Figure 1. The AWBUS dataset required for the segmentation process was provided by the physicians of the Department of Ultrasound, First Affiliated Hospital of Shantou University, Shantou, Guangdong, China. The dataset comprised 64 coronal slices

Therefore, it is challenging to perform automated NSA segmentation on the images along the sagittal plane. To enable accurate localization, we stack all the sagittal slices
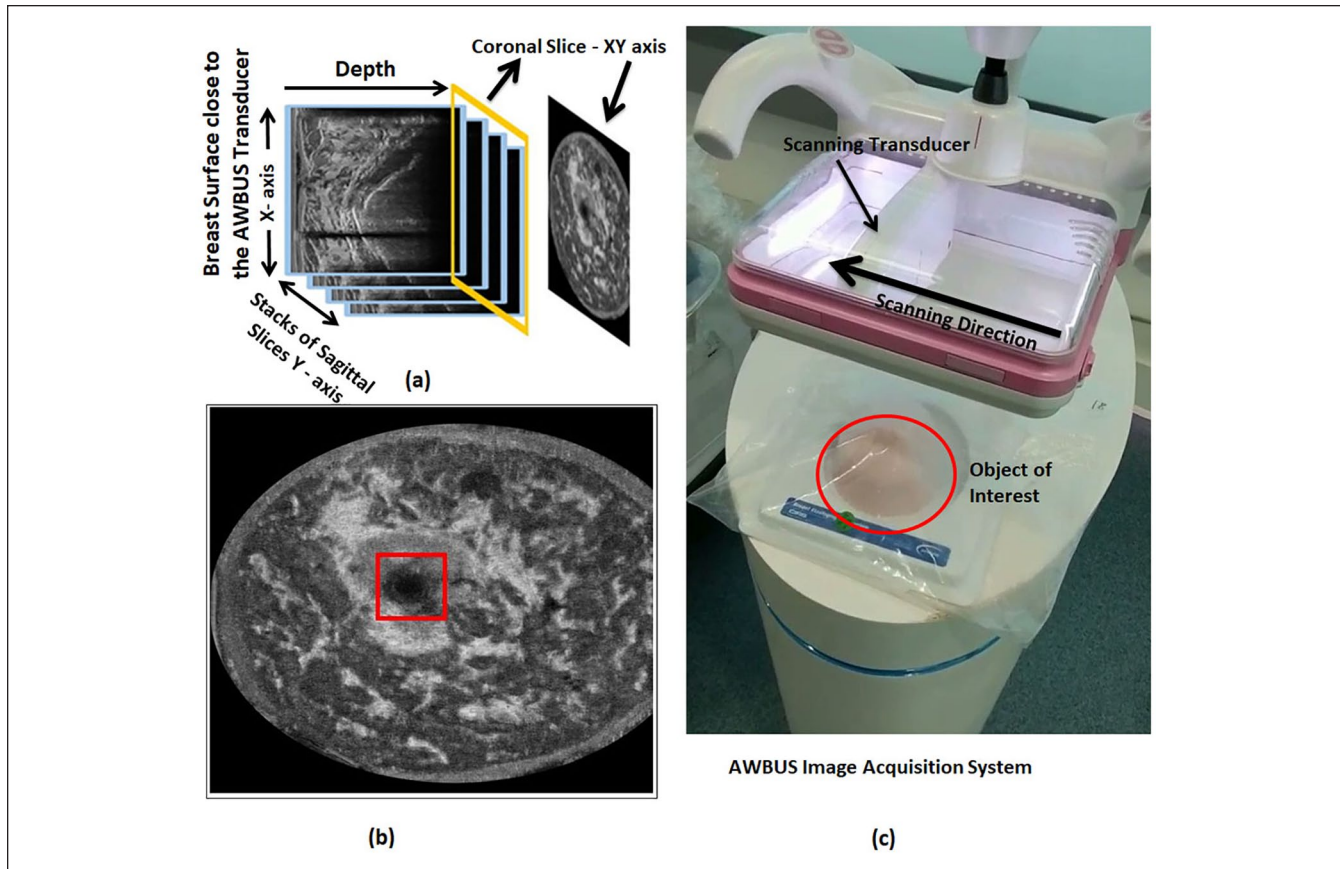
Figure 2. (a) Stacking the slices from the sagittal plane and re-slicing along the coronal plane. (b) NSA is shown as a disk in the coronal view indicated inside a red box along with few examples from the data set. (c) Automated Whole Breast Ultrasound (AWBUS) Image acquisition system; illustrating the transducer, scanning direction and the object of interest.
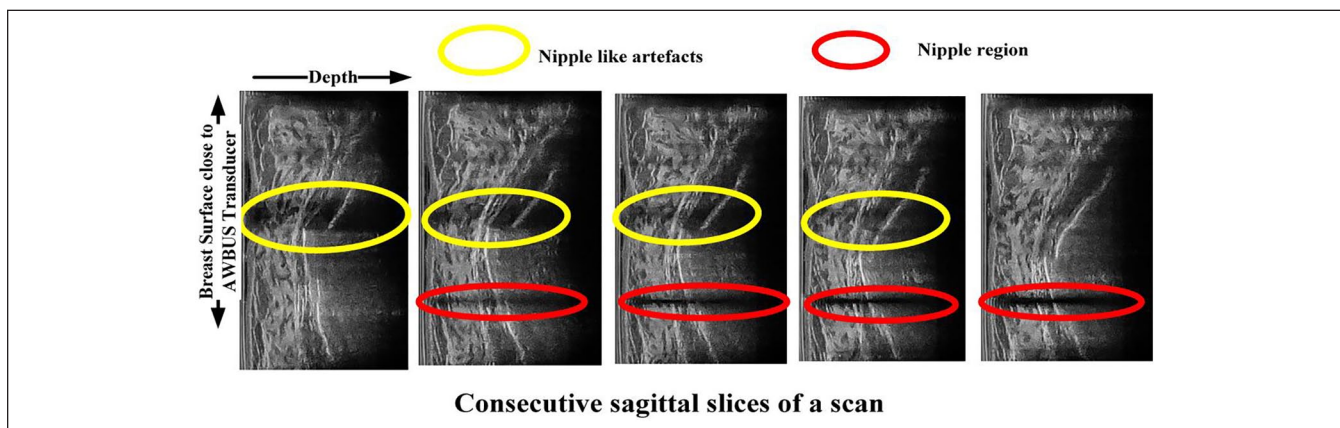

Figure 3. Illustration of NSA visible as a dark tube on the sagittal plane along with other artifacts that look similar to the NSA.

derived from the AWBUS scan of a patient as a three-dimensional matrix and then sliced along to the XY axis at different depths to obtain the coronal view of the breast under examination. Figure 2 presents the AWBUS scanning process and the extraction of the breast slice along the coronal plane.

The resulting coronal slices I_T has many significant advantages over the sagittal slices I which is shown in Figurer 2(a). The distinguishability of the nipple region from other tissues has increased significantly and the nipple region appears as a dark disk-like region marked as red in Figure 2(b) after plane transformation. Also on average, the number
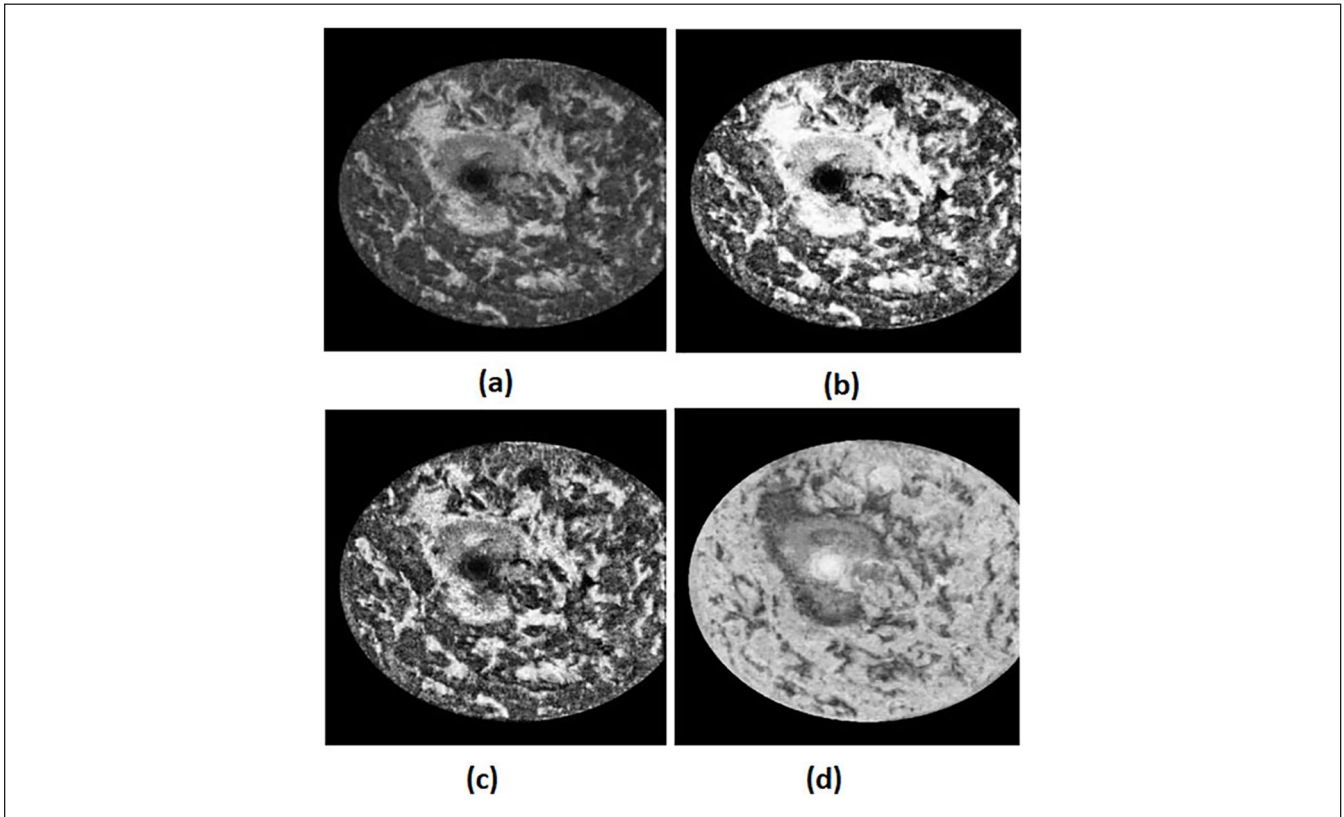
Figure 4. (a) The original image. (b) Image after histogram equalization. (c) Image after CLAHE. (d) Image after intensity inversion.

of slices that contained the nipple region has increased from 3.08% to 7.03%. The total number of slices considered for this study is approximately 700. Thus presenting an increased amount of relevant data by 128%.

Contrast limited adaptive histogram equalization. Ultrasound images are prone to speckle noise, blurred edges, and other artifacts which lead to poor segmentation performance therefore adequate pre-processing is required for accurate segmentation. The coronal slices obtained after plane transformations are qualitatively better, but suffer from low contrast in differentiating the NSA region from the background. Many contrast enrichment algorithms are based on adjusting the histograms. The global contrast enrichment technique is not suited for the images where the local information is necessary and also for the unclear lighting. Histogram equalization is a common technique for contrast enhancement that distributes the intensities over all possible gray values but global histogram equalization generally results in oversaturated as shown in Figure 4(b).

For medical images, as observed by Konyar and Ertürk[22] and Mousania and Karimi,[23] Contrast Limited Adaptive Histogram (CLAHE) is an effective method for improving the contrast of the region of interest without causing oversaturation. It employs a local transformation function that enhances each pixel derived from a neighborhood region by applying the histogram equalization on non-overlapping regions of the images by means of interpolation that is used to correct the variations among the borders.[24] It has two parameters, namely the number of tiles and the clip limit. The clip limit controls the over amplification of noise and tile measure size of the non-overlapping region. Once the histogram is calculated for each tile, they are reordered and ensured that its height is not exceeded above the to the clip limit. Later, the equalization is performed by calculating the cumulative histogram over the chosen sub region. Therefore, we apply CLAHE over a window of 8×8 to obtain the contrast-enhanced image I_H as shown in Figure 4(c). To quantitatively compare the performance improvement, the Histogram Flatness Measure (HFM) and Histogram Spread (HS)[25] are computed.

HFM is the ratio of the geometric mean to the arithmetic mean of the histogram intensities and is computed as shown in equation (1).

$$\text{HFM} = \frac{\left(\prod_{i=1}^{n} x_i \right)^{\frac{1}{n}}}{\frac{1}{n} \sum_{i=1}^{n} x_i} \tag{1}$$

where, x_i is the histogram count of the ith histogram bin and n is the total number of bins.

The arithmetic mean is always greater than or equal to the geometric mean, thus HFM lies between [0,1].

Table 1. Histogram Equalization Score for I_T and I_H.

Measure	I_T	I_H
HFM	0.38926	0.86930
HS	0.21875	0.37051

HS is the ratio of the quartile distance to the histogram range. The quartile distance is computed by taking the difference between 3rd quartile and 1st quartile of the histogram. The range is found by measuring the difference between the maximum and minimum intensities of the image. HS is computed as shown in equation (2).

$$HS = \frac{\left(3^{rd} - 1^{st}\right) quartile\ of\ the\ histogram}{\left(Max - Min\right) Intensity\ of\ the\ image} \qquad (2)$$

Where, 3rd quartile indicates the histogram bins where the cumulative histogram is 75% of the maximum value and the 1st quartile indicates the histogram bins where the cumulative histogram is 25% of the maximum value. For an image with uniform distribution, the maximum value of HS is 0.5.

The measures for I_T and I_H are presented in Table 1. The higher values of HFM and HS for I_H indicate better contrast improvement than the original coronal images.

Intensity inversion. As the final step of pre-processing, we invert the pixel intensities of I_H using equation (3). Inverting the pixel intensities changes the NSA from dark (grayscale values close to zero) to light (grayscale values close to 255) enabling efficient calculation of shape features through Hu moments.[26]

$$I_{HI}\left(x, y\right) = 255 - I_H\left(x, y\right) \qquad (3)$$

The intensity inverted image (I_{HI}) is shown in Figure 4(d). Later we extract the shape from I_{HI} as explained in the following sections.

Shape and Textures Feature Extraction

The target of the proposed method is to classify each pixel within the image as either a nipple or a non-nipple pixel. Using intensity-based descriptors (such as computing the mean or maximum or minimum intensities with an $N \times N$ window) alone to classify the pixels will fail since the AWBUS slices contain artifacts that have intensities similar to NSA as shown in Figure 2(b). Since in most of the coronal slices, NSA resembles a disk-like region, here we adhere to extracting the shape information as one prominent feature to distinguish the NSA region from other similar regions.

Hu Moment Invariants, or simply Hu Moments, can effectively describe the shape of the object within an image.[27] It is a popular feature used for pattern recognition in medical imaging.[28-31]

Being shape descriptors, Hu Moments are traditionally calculated for an $N \times N$ window which is then slid across the image along the rows and columns to obtain the shape descriptors for the entire image. The raw moments are computed for the central pixel of each window using equation (2).

$$M_{ij} = \sum_{x=0}^{r-1}\sum_{y=0}^{c-1} x^i y^j I_{HI}\left(x, y\right) \qquad (4)$$

Where i and j are non-negative integers, x, y are the coordinates of a pixel in the intensity-inverted image I_{HI} and r and c are the numbers of rows and columns of the image respectively. These raw moments M_{ij} are used to calculate the centroids using equations (5) and (6) respectively

$$\underline{x} = \frac{M_{10}}{M_{00}} \qquad (5)$$

$$\underline{y} = \frac{M_{01}}{M_{00}} \qquad (6)$$

Later, the raw moments defined by equation (4) are transformed using the centroid to obtain central moments as presented in equation (7).

$$\mu_{ij} = \sum_{x=0}^{r-1}\sum_{y=0}^{c-1} (x - \underline{x})^i (y - \underline{y})^j I_{HI}(x, y) \qquad (7)$$

The computed central moments are translation invariant and to obtain scale invariance, we normalize the central moments as per equation (8)

$$\eta_{ij} = \frac{\mu_{ij}}{\mu_{00}^{\frac{i+j}{2}+1}} \qquad (8)$$

Here, instead of using the traditional sliding window approach, we employ an iterative sliding window approach to compute more accurate shape descriptors. In the first iteration, an $M \times M$ sliding window, W^a with a stride of length, l pixels is traversed along the row and column of I_{HI} and the Hu-Moment for every $(i,j)^{th}$ pixel within each window is computed and associated with the center pixel of the window as shown in equation (9).

$$Hu_a(P_{i,j}^a) = h_\rho\left(I_{HI}(i, j)\right), P_{i,j}^a \in W_k^a \qquad (9)$$

Where h_ρ represents the Hu moment of inertia function of order ρ, W_k^a represents the kth window, i,j represent the coordinates along the row and column respectively and $P_{i,j}^a$ denotes the center pixel of W_k^a. In the second iteration, another $N \times N$ sliding window, W^b is traversed across Hu_a and the maximum value within W_k^b is associated with the central pixel $P_{i,j}^b$ as shown in equation (10).

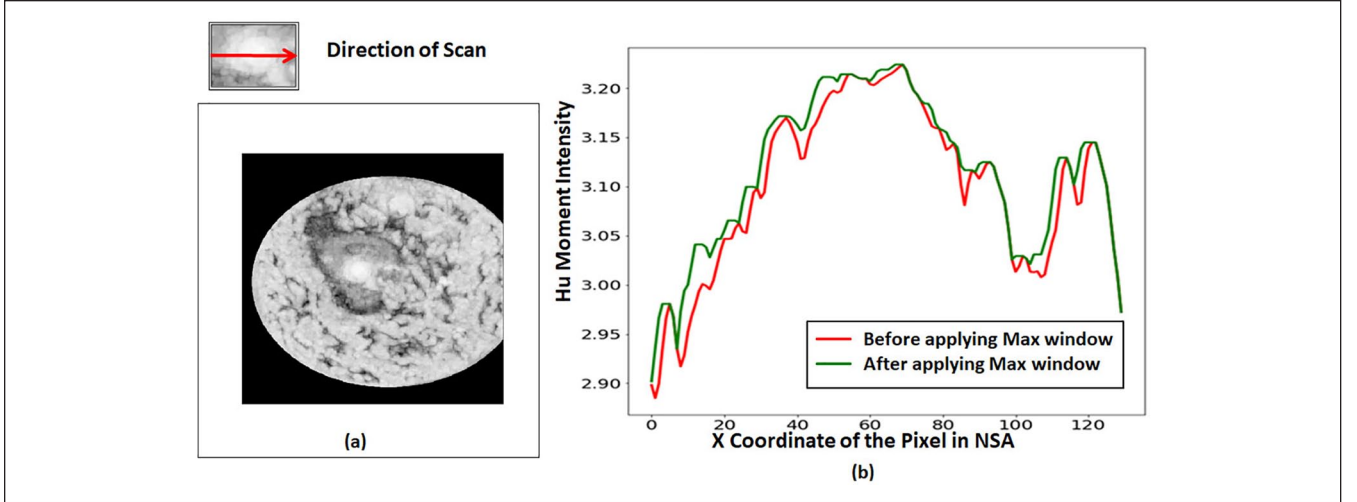$$Hu_b(P_{i,j}^b) = max\{Hu_a(i, j)\}, P_{m,n}^b \in W_k^b \qquad (10)$$

Figure 5. (a) Illustration of nipple region after the second iteration for the first order Hu moment. (b) Comparison of intensities before and after the second iteration for the horizontal profile drawn across the nipple regions.

It was observed that the values of Hu_b for the pixels within the NSA were higher when compared to the rest of the breast and the NSA presented by Hu_b were smoother and more amplified as compared to the traditional sliding window was shown in Figure 5. For our experiments, W^a, W^b and l were set to 5 × 5 pixels and 1 respectively.

In the third and final iteration, we slide a window W_k^c of size N × N pixels with a stride length of one over Hu_b and extract shape feature vector (Hu_{fv}) by concatenating the Hu_b over the neighborhood of the central pixel of W_k^c to the central pixel as shown in equation (11).

$$Hu_{fv}(P_{i,j}^c) = [Hu_b(P_{0,0}),..Hu_b(P_{N \times N-1})], P_{i,j}^c \in W_k^c \qquad (11)$$

Where Hu_{fv} is the Hu moment—based shape feature for k^{th} window and $P_{i,j}^c$ denotes the center pixel. The extraction of shape features through iterative sliding window technique is pictorially illustrated in Figure 6. From Figures 5 and 6, it should be noted that the boundary of the NSA is sharp, smooth and distinguishable when Hu_b is considered.

Although a set of seven Hu moments can be derived from equation (8) by submitting various combinations of (i,j), it was observed (refer Figure 7) that the first order Hu-moment (h_1) computed using equation (12) presented a clear distinction between the NSA and the rest of the breast region as compared to other higher-order Hu-Moments. Also, the first-order moments are computationally least extensive and hence we extract shape description from h_1 for further classification of NSA regions.

$$h_1 = \eta_{20} + \eta_{02} \qquad (12)$$

In our final iteration, we set W^c as 3 × 3, and thus obtain a shape feature vector Hu_{fv} of size 9. The window sizes W_k^a

and W^b for the proposed iterative technique were selected through an exhaustive process, where sliding windows of sizes of 3, 5, and 7 were tested. Amongst these, 5 × 5 window provided smoother and amplified responses that helped in localization of NSA regions more accurately (Please refer to Figures 5(b) and 6). W_k^c was always fixed to 3 x 3 since we are interested in extracting the shape characteristics surrounding the center pixel. Further in all cases the step size was fixed to 1.

After extracting shape descriptors, image texture estimation is done by Gray Level Co-occurrence Matrix (GLCM) that could differentiate NSA and the background regions. GLCM estimates the joint probability density function of the gray level pair of the image. GLCM is widely used in medical images for lesion segmentation. Milosevic et al.[32] used a total of 20 GLCM features to diagnose the abnormal patterns from breast thermography. Hossain et al.[33] extracted the GLCM based textural features from the lung parenchyma region to detect lung cancer. Moreover, Sankaranarayanan made use of energy, contrast, entropy, correlation, homogeneity features of GLCM to classify oral cancer at an earlier stage.[34] Initially, the texture descriptors were calculated for the CLAHE enhanced image I_H image where a sliding window of size 9 × 9 was used to compute the GLCM features across four directions [0, 45, 90, and 135]. Later the features obtained across the four directions were averaged to find the complete textural characteristics of the image. Normally features such as Contrast, Entropy, Correlation, and Energy are calculated from the GLCM matrixes; however, since entropy presents a measure of homogeneity, we employ entropy descriptor G calculated using equation (13) as texture features for successful differentiation of NSA from the background regions. The derived feature map is shown in Figure 7(h).
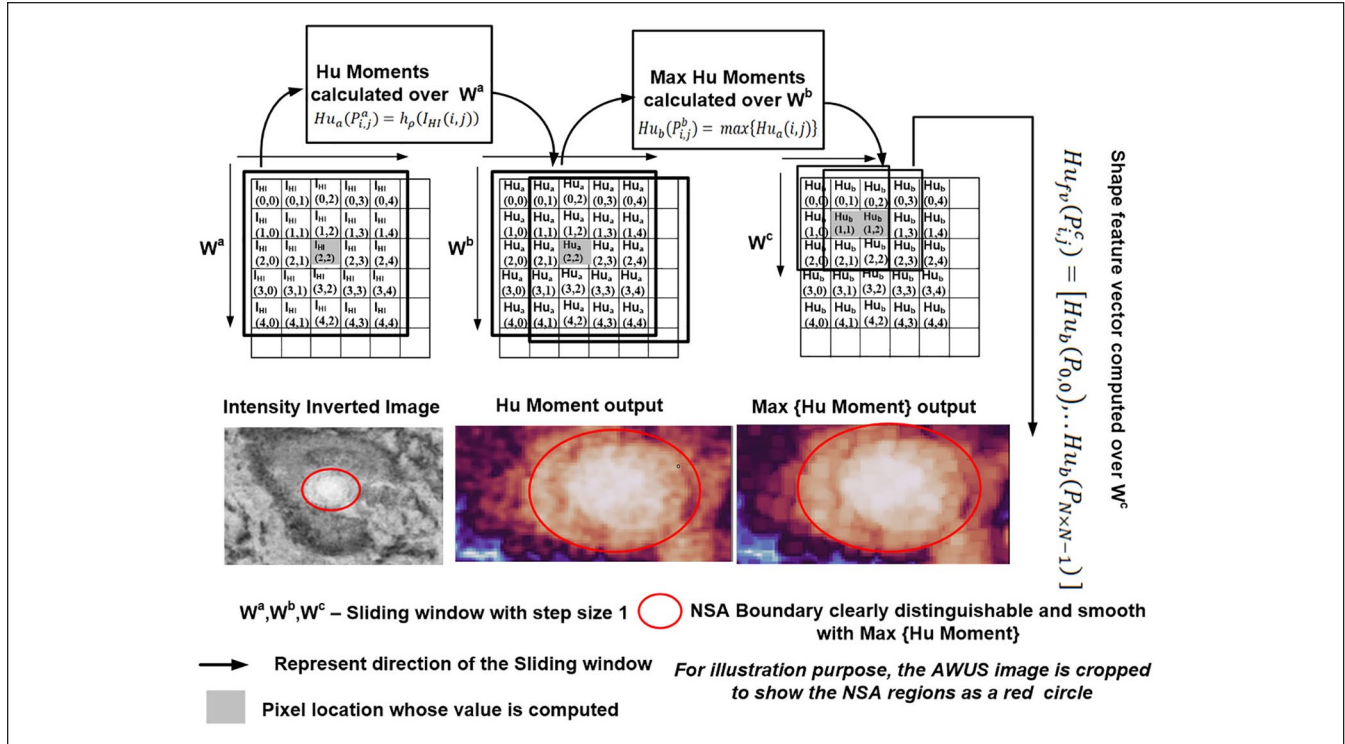
Figure 6. Illustration of extraction of shape features through iterative sliding window technique.

$$G = -\sum_{i}^{g}\sum_{j}^{g}p(i,j)log_e(p(i,j)) \qquad (13)$$

Where, $p(i,j)$ is the probability of occurrence of intensity level in an image.

Similar to the calculation of Hu_{fvc}, 3×3 window is transversed across G and values over the neighborhood are concatenated to obtain the 9 texture feature.

$$G_{fv}(P_{i,j}^d) = [G(P_{0,0}),..G(P_{N \times N-1})], P_{i,j}^d \in W_k^d \qquad (14)$$

Where G_{fv} is the GLCM based shape descriptor for kth window and $P_{i,j}^d$ is the center pixel

Finally, we concatenate the shape and texture features as indicated in equation (15) for each pixel to obtain 18 features which are further normalized as per equation (16) and then fed to the ANN for accurate classification of NSA and background regions.

$$X(P_{i,j})_{l=1.........18} = [Hu_{fv}, G_{fv}] \qquad (15)$$

$$X = \frac{X_l - \mu}{\sigma} \qquad (16)$$

Where X is the normalized feature vector, μ is the mean of the feature and σ is the standard deviation of the feature. The extracted texture and shape features are further annotated for the defined classification task.

Determining Probable NSAs

To determine the probable NSAs from the coronal slices, the extracted shape, and texture features with the annotations are fed to an ANN for determining the probable NSA regions. The architecture of ANN consists of an input layer, 4 hidden layers with *tanh* activation function and the output layer with a log-sigmoid activation function. The input layer accepts the feature vector and the output layer presents a classification score of 0 or 1. Here a score of 1 represents the occurrence of NSA pixel and 0 the background.

To train the ANN and validate the model, two non-overlapping datasets (training and testing) were created manually using a holdout approach. The training dataset had a total of 10599 pixels, of which 4508 pixels were from the NSA and 6091 from the non-NSA regions (background) from 14 breast scan slices of 13 patients. These data were carefully selected so that the selected data will have a balanced representation of NSA and Non-NSA classes. While selecting the data, the NSA and Non-NSA regions are correctly captured even though they appear in opposite way, so that the number of false positives is reduced. The feature vectors, each of length 18 extracted from the selected pixels along with their labels, were fed as inputs to the ANN for training. The features were labelled 0 and 1 based on the ground truths provided by the experts(generally more than1) from the Shantou First Affiliated Hospital, Shantou, who have several years of practical experience in the analysis of ultrasound images. Hence the confident level of the experts in the annotation of the ground truth is high. The label "0"
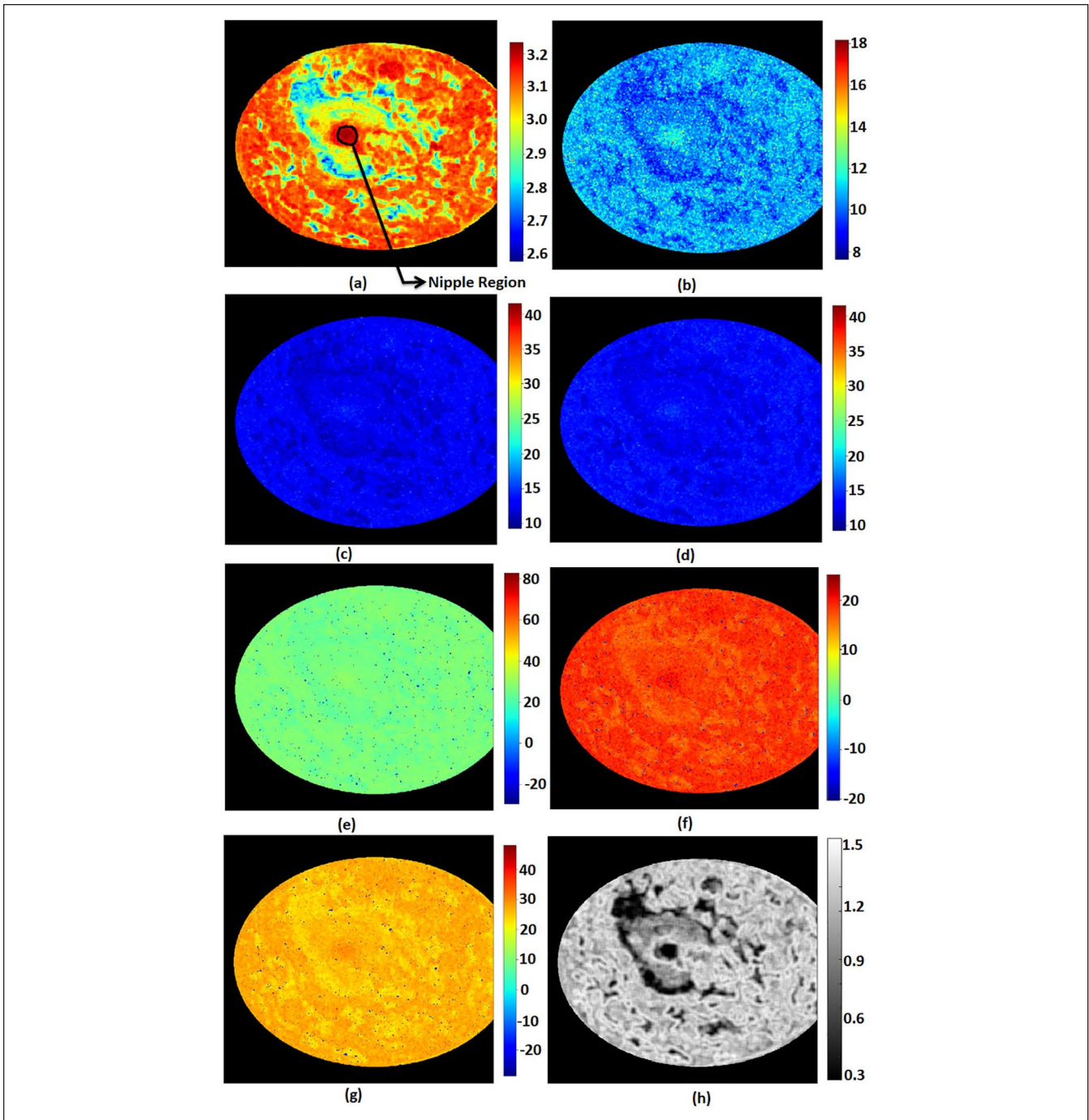
Figure 7. Illustration of first 7 Hu moments (a–g) and GLCM based Entropy features (h). Here the color bar indicates the Hu moments values obtained for values various orders. It also illustrates whether the chosen Hu moment can be readily used to distinguish the NSA region from the background. The gray scale in (h) indicates the entropy value for the GLCM feature. For interpretation of the references to colors in this figure legend, refer to the online version of this article.

indicates NSA and "1" indicates non-NSA. The parameters of the ANN were tuned by setting an epoch limit of 500 with cross-entropy as the loss function. The training accuracy was calculated using equation (17) that provided 92.7%.

The trained ANN was then used to predict the nipple regions in 50 unseen AWBUS images of the dataset. Thus a testing set of 7,08,6797 pixels was considered for the prediction. These testing data set points are excluded from the training set. Since the output of the ANN presented probabilistic values for each pixel, in order to reduce the false-negative rate a low threshold of 0.5 was used for classification. Therefore, the classifier was able to identify most pixels belonging to the NSA in almost all images. Figure 8 shows the contour of the detected region in red and the ground truth in yellow for one of the images.
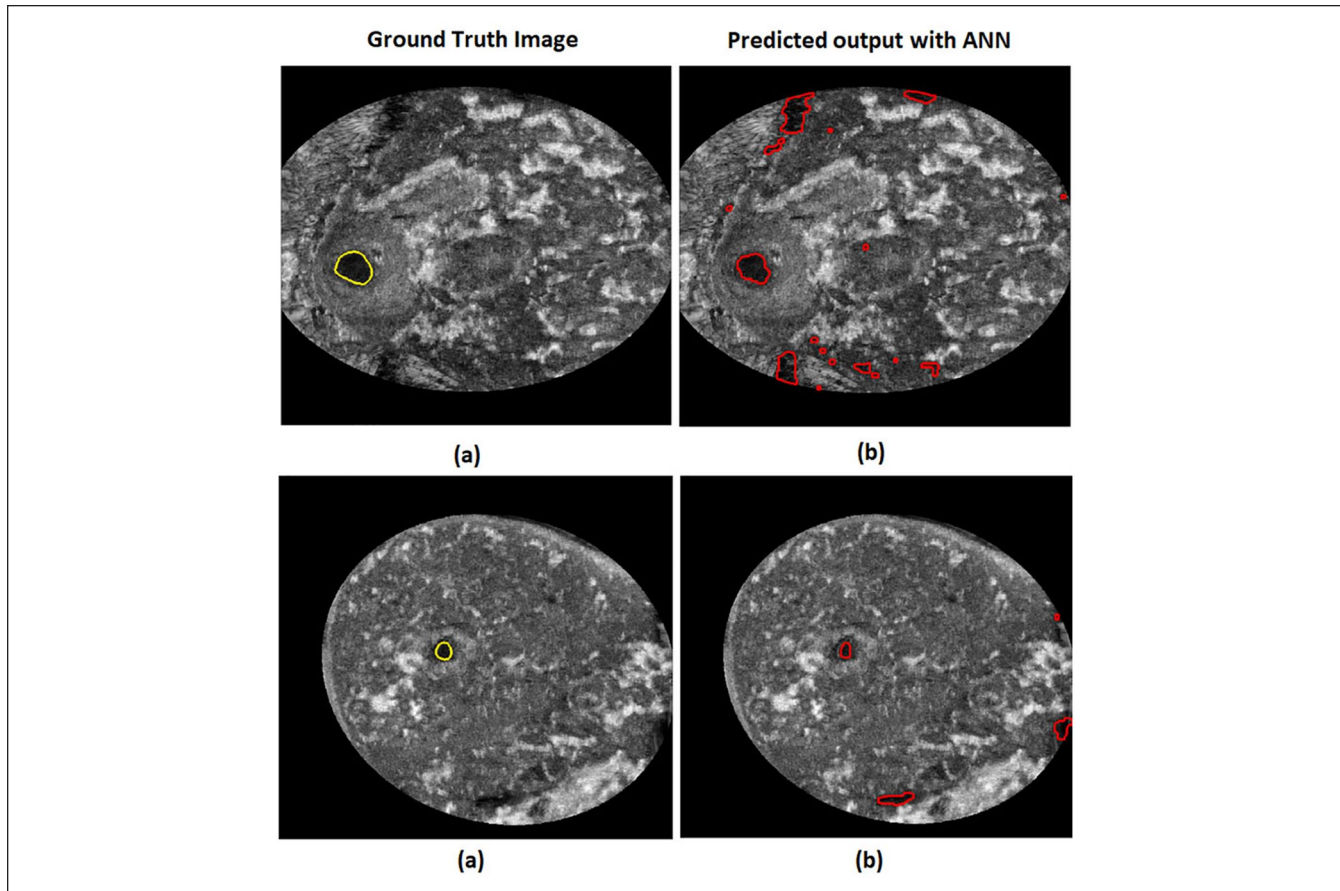
Figure 8. Predicted NSA regions of two different images using ANN after morphological operations. (a) Original image showing the ground truth (marked in yellow) annotated by the experts. (b) Illustrates the predicted outputs (marked red) from the ANN. For interpretation of the references to colors in this figure legend, refer to the online version of this article.

Table 2. Confusion Matrix from ANN.

Before morphological operations				After morphological operations			
Predicted class	non-NSA	6975259 98.42% (TN)	9451 0.13% (FN)	Predicted class	Non-NSA	7002296 98.79% (TN)	12106 0.17% (FN)
	NSA	65264 0.92% (FP)	36823 0.52% (TP)		NSA	38227 0.54% (FP)	35241 0.497% (TP)
		Non-NSA	NSA			Non-NSA	NSA
		Actual class				Actual class	

Unfortunately, some images had various other lesions that appear similar to the NSAs which resulted in false positives. Though we employ morphological opening and closing on the reconstructed predictions to successively fill the holes in larger regions and eliminate smaller false predictions, some larger false positive NSA predictions remained, as shown in Figure 8. The confusion matrix before and after the use of morphological operators is presented in Table 2, as are the performance measures (refer Table 3) such as accuracy, specificity, sensitivity, F score and Intersection of union (IOU, also referred to as Jaccard Similarity index). IOU is a measure of similarity that compares the predictions of the classifier with target labels (annotations) and the range of the index is [0,1]. When the index moves towards 1, the prediction will also move more towards the target label. The performance measures computed using equations 17 to 21 demonstrate the ANN classifier's efficiency in detecting probable NSA regions.

$$accuracy = \frac{TP + TN}{TP + TN + FP + FN} \tag{17}$$

$$specificity = \frac{TN}{TN + FP} \tag{18}$$

Table 3. Performance Measures for Predictions from ANN.

	Before morphological operations	After morphological operations
Accuracy	0.9895	0.9929
Specificity	0.9907	0.9946
Sensitivity	0.7958	0.7384
F score	0.4964	0.5759
IOU	0.3301	0.4043

$$sensitivity = \frac{TP}{TP + FN} \qquad (19)$$

$$F\,score = \frac{2*TP}{2*TP + FP + FN} \qquad (20)$$

$$IOU = \frac{TP}{TP + FN + FP} \qquad (21)$$

As visible from the results of the predictions, although the ANN detects the NSA with decent accuracy, the number of false positives is high. Though morphological opening and closing operations were used to fill the holes in the larger regions to reduce false positives, but it resulted in an increase of false negatives, which further reduced the sensitivity by 0.0574.

From the reconstructed predictions from the ANN, the following observations were made:

- The true NSAs resemble a distorted ellipse while the false NSAs had haphazard shapes and boundary structures as shown in Figure 9(a).
- The true NSAs were farther from the boundaries (refer to the Figure 9(a)) of the breast as compared to false NSAs.
- The true NSAs were larger as compared to the false NSAs.

NSA-Classification

Thus, in order to identify the true NSAs out of the predicted probable NSAs, we manually separate the predicted contours of the nipple (NSA) and non-nipple region (non-NSA) and the contour features were extracted for each of them to train the classifier. They are as follows:

Fractal dimensions. The NSA contours are generally circular or oval-shaped and thus display a lower degree of shape complexity. Therefore, Fractal Dimensions were used to assess the complexity of the contours of the predicted NSA regions. For each predicted NSA contour **P**, the shape complexity was calculated using the method proposed by Mandelbrot.[35] Recently Zhuang et al.[36] have also used Fractal Dimensions for segmentation and classification. To measure the shape complexity of the contour, the radius r in the range 2 to 8 pixels was taken and the circles were drawn along the

periphery of the predicted NSA as depicted in Figure 9(b) to (e). For each **P**, the number of circles, L(r) drawn was calculated and then plotted on a log-log scale with r on the X-axis and L(r) on the Y-axis as shown in Figure 9(f). A least square error best-fit line was then drawn on the resulting points and the slope of the line α was calculated. The r and L(r) values for one of the contours are given in Table 4 and the relation between the radius r and the number of circles drawn is presented below.

$$(L_{i=1}^{P}(r)) = \alpha_{i=1}^{P}(r_{j=2}^{8}) + \beta_{i=1}^{P} \qquad (22)$$

Where **P** denotes the contours of the probable NSA regions predicted by the ANN, *r* represents the different radius of the circles drawn around P, L(r) represents the total number of circles obtained for each r; and α and β signify the fractal dimensions and the fractal length respectively. Also i and j represent the indices denote the indices associated with the contours of the probable NSA regions and radius r respectively. In general, the complex shapes have larger α values as stated by Okubo and Aki.[37]

Since there are many predicted contours for each image, few contours of NSA appear to be contours of non-NSA and vice versa, thus we calculate the mean and standard deviation of fractal dimension, α_{mean} and α_{std} for NSA and non-NSA to reduce the error in discrimination using equations (23) and (24) respectively.

$$\alpha_{mean} = \frac{\sum_{i=1}^{P}\alpha_i}{P} \qquad (23)$$

$$\alpha_{std} = \sqrt{\frac{\sum_{i=1}^{P}(\alpha_i - \alpha_{mean})^2}{P}} \qquad (24)$$

Edge distance. Due to the fact that the true NSAs are farther from the breast boundary as compared to non-NSAs, the centroid of each contour present in NSA and non-NSA is calculated using eqs. (3) and (4). Then the shortest distance δ between the breast boundaries and the centroids is calculated. Similar to α_{mean} and α_{std}, δ_{mean}, and δ_{std} were calculated for the contours of the same image.

Contour area. Like the fractal dimension and edge distance, the contour area is also significant in discriminating between
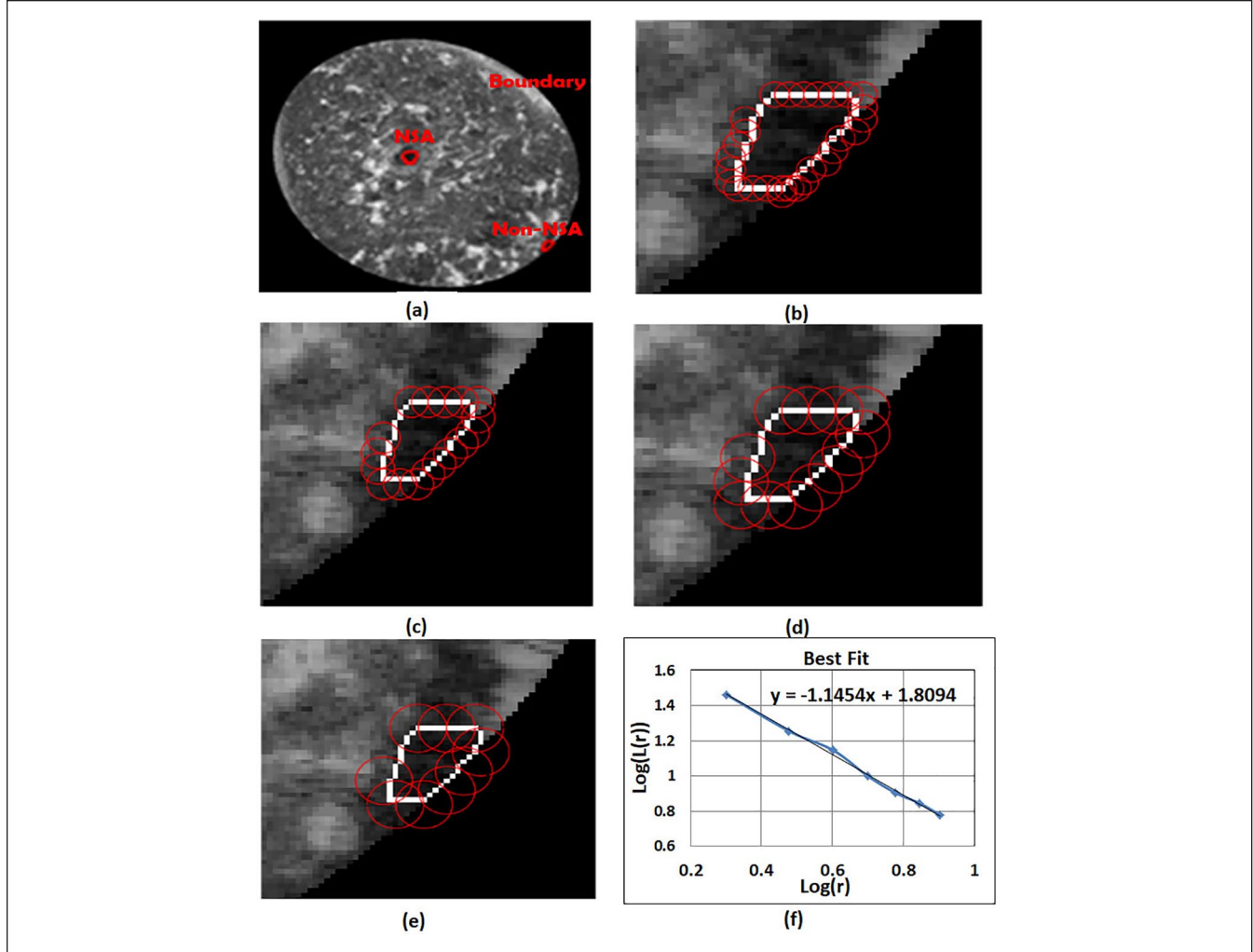
Figure 9. (a) Original image with two contours. (b) Circles with radius $r = 2$. (c) $r = 3$. (d) $r = 4$. (e) $r = 5$. (f) Best fit for data in Table 4 (From the graph $\alpha = -1.1454$).

Table 4. r and L(r) Contour Length Plotted in Figure 9(f).

r	2	3	4	5	6	7	8
L(r)	29	18	14	10	8	7	6

the true NSA and non-NSAs. The area of each contour λ was calculated using Green's formula,[38] by considering the number of pixels enclosed within the contour. Thus from the contour, features are extracted and a feature vector X_{Ci} is created for each contour P_i present in NSA and non-NSAs of the training image.

$$X_{Ci} = [\alpha(P_i), \alpha_{mean}(P_i), \alpha_{std}(P_i), \delta(P_i),$$
$$\delta_{mean}(P_i), \delta_{std}(P_i), \lambda(P_i)] \qquad (25)$$

The feature vector is presented to the Support Vector Machine (SVM) classifier to classify NSA and non-NSA contours. SVM classifier was selected as it (a) performs well on the data even with lesser number of samples considered for the training the model, (b) has a robust regularization characteristics that overcome the problem of overfitting and (c) can model the nonlinear classification by using an appropriate kernel that transforms the data from low dimensional space into to higher-dimensional space.

Since the work considered needs to identify the true NSAs out of the predicted probable NSAs, RBF kernel in SVM was appropriate. The probable NSA regions predicted by the training dataset of ANN were used to train the SVM. From our analysis on an average the 12 NSA regions (includes both true and false) were predicted per BUS image. However, since the number of true NSA were less, data augmentation techniques involving four affine transformations such as shift along vertical axis and horizontal axis, shear transformation and flipping about the horizontal plain were employed to increase the number of true NSA regions to balance the dataset. The number of true NSAs were expanded to a total
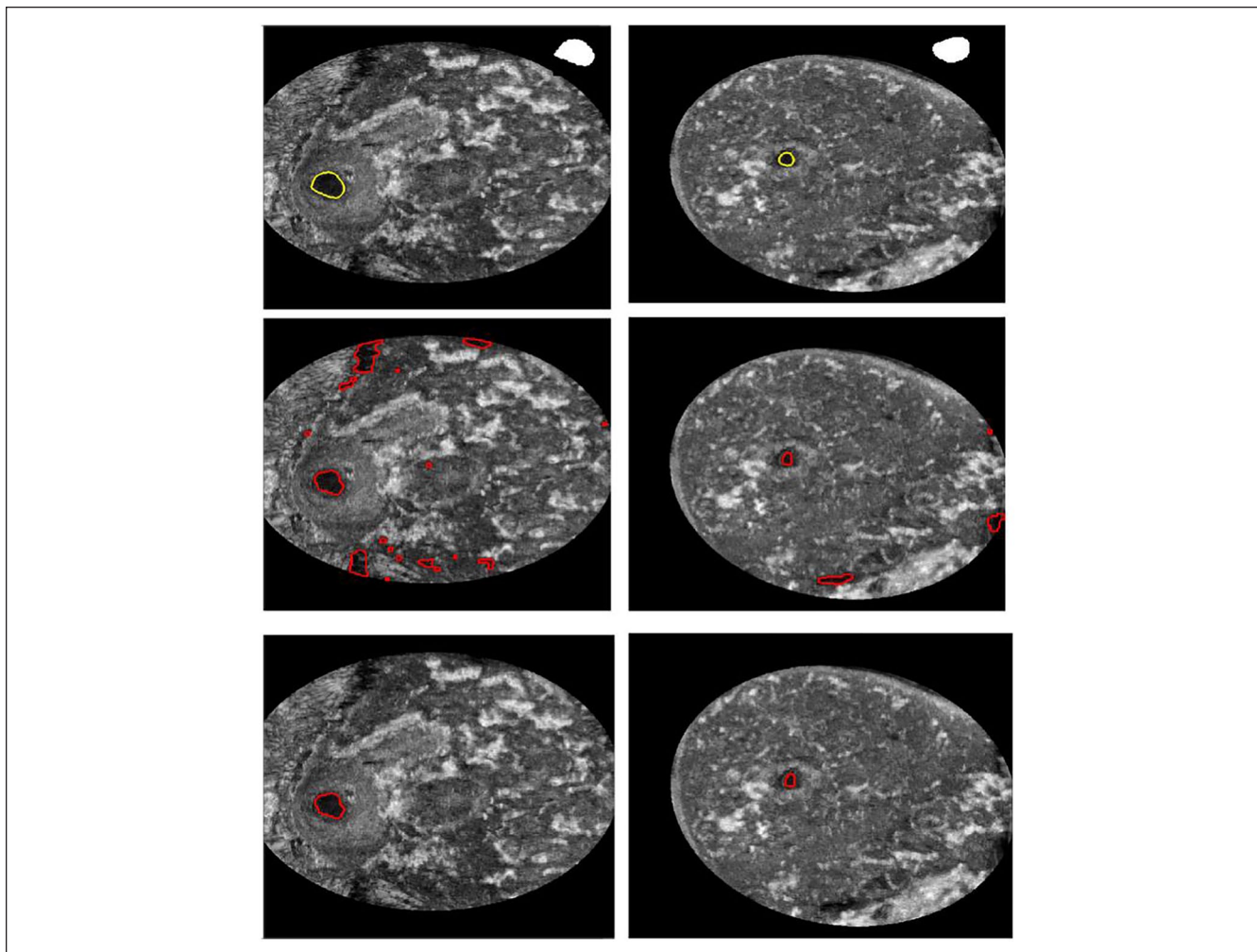
Figure 10. Scans of two different patients at different phases of the algorithm. Row1—Original image along with the ground truth, Row2—probable NSAs predicted by ANN and Row3—True NSAs predicted by SVM.

of 56 regions after data augmentation. Thus, 56 NSA regions[1] and 154 non-NSA[2] regions were available as the training set, from which we manually separated the predicted contours of the nipple (NSAs) and non-nipple (non-NSAs). Accordingly, these contours that were selected formed a balanced dataset for training the SVM classifier.

During training, the regularisation (C) and the kernel (γ) parameters were varied between 0.1 and 10 under 100 iterations to tune the classifier model avoiding the consequences of over and underfitting. Accordingly, the optimized model was obtained for C = 1 and γ = 0.5 and provided an accuracy of 98% on the training set. Later, the trained SVM model was then validated on the 185 probable NSA regions of the testing data set and their corresponding performance metrics are presented in the next section. The results of classification of true NSA regions from a set of probable NSA regions predicted using ANN for 2 different patients is shown in Figure 10.

Results

Further, the performance of the model was evaluated on the pixels in the probable NSA regions using confusion matrix and Receiver Operating Characteristics (ROC). The pixels in the regions, whose contours were classified as true NSA contours by the SVM were considered as true NSA pixels. The confusion matrix and performance metrics for these pixels on the test data are given in Tables 5 and 6 respectively.

The ROC obtained for the predicted true NSAs on the test data is displayed in Figure 11. From the curve, the performance metric Area under the curve (AUC) is computed which was found to be 0.9001. The value of the AUC indicates the better accuracy of the model.

Discussion

A comparison of our method to (a) Active Contours (b) Local Graphcut and (c) K means clustering techniques shows that he

Table 5. Confusion Matrix of Predicted True NSAs ON TEST DATA.

Predicted class	Non-NSA	115538 (TN) 61.62%	12106 (FN) 6.45%
	NSA	2894 (FP) 1.54%	56951 (TP) 30.37%
		Non-NSA	NSA
		Actual class	

Table 6. Performance Metrics After Contour Classification on TEST data.

F-score	0.88
Specificity	0.9755
Sensitivity	0.8246
IOU	0.7915
Accuracy	0.9199
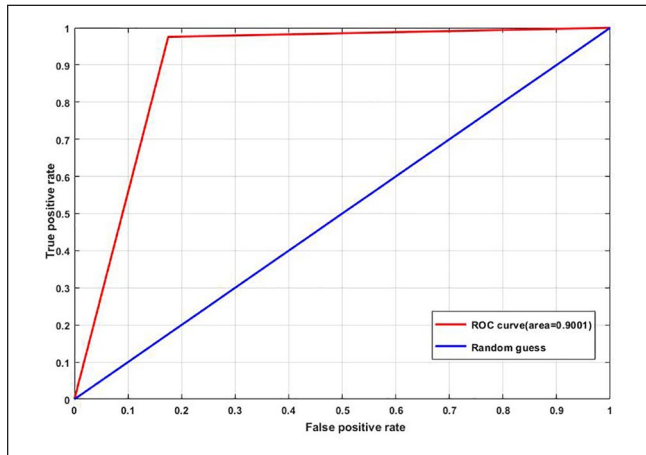
Figure 11. ROC of the predicted true NSA regions on the test data.

first two methods are semi-automatic. That is, for these techniques, a region of interest had to be marked manually and the NSA regions could be segmented. In the case of Active contours, the seed point was marked within the NSA and the region-based active contour method proposed by Chan and Vese[39] was used for segmentation. In the case of Local Graphcut method,[40] a six-sided polygon was chosen around the NSA region as the initialization ROI for the algorithm. The resulting segmentation output along with the selected ROI is shown in Figure 12(c1–c3). Finally, the segmented results of the k means[41] are presented in Figure 12(d1–d3). Here the number of clusters k was chosen as 2 (NSA and non—NSA) and resulting clusters were post-processed using morphological operators to obtain the NSA regions.

When qualitatively compared to the ground truths using the Disc Similarity Coefficient (DSC) (equation (24)) index,

the segmentation results presented scores of 0.8045, 0.8286, and 0.8539 for the Active Contours, Local Graphcut and the proposed technique respectively. Though the results of Active Contours and Local graph cut methods are comparable to ground truths, the final segmentation results depend on initialization and the number of iterations. K means, though it automatically predicts the NSA regions, fails due to a large number of false positives and hence presents poor DSC results. Therefore, the proposed two-step procedure of first finding probable NSA regions through shape and texture features and then classifying the extracted regions using shape complexity, edge distance and area presents better qualitative and quantitative results as shown in Figure 12(e1–e3).

DSC is calculated as,

$$DSC = \frac{2\,|\,G \cap P\,|}{|\,G\,| + |\,P\,|} \qquad (26)$$

where; G is manually annotated ground truth by the expert w and P is the predicted segmentation. Higher DSC values confirm that segmentation results are identical to the ground truths. Finally Table 7 presents a concise comparison between our proposed method and various techniques mentioned in.[39-41]

Conclusion

The paper presents a CADe system to locate the Nipple Shadow Area (NSA) from Automated Whole Breast ultrasound (AWBUS) scans, which are used as a reference point to detect the position of any suspected breast lesions. For accurate NSA detection, it is required to reduce false detection by employing efficient segmentation methods. Thus the proposed technique utilizes the shape and texture features of NSA extracted using the Hu moments and GLCM through an iterative sliding window to train ANN for detecting the probable NSAs. Later, the true NSAs are identified from the predicted probable NSAs using SVM adopting the contour features: fractal dimension, edge detection, and contour area.

The proposed methodology was tested on the AWBUS dataset provided by the First Affiliated Hospital of Shantou University. Also, since the algorithm was tested on individual AWBUS coronal slices which are 2D, therefore could be employed to segment NSA region from BUS images obtained from other ultrasound instruments. Furthermore, the extraction of coronal planes (2D) by stacking sagittal slices helps in viewing the nipple region as a circular disk, thereby aiding segmentation through shape and intensity features.

From the experimental results, it is evident that SVM based classification of NSA from the contour predicted by ANN using the Hu moment and GLCM values gives better segmentation accuracy, and it also outperforms other similar techniques in various metrics, which can contribute to the design of better health care systems.
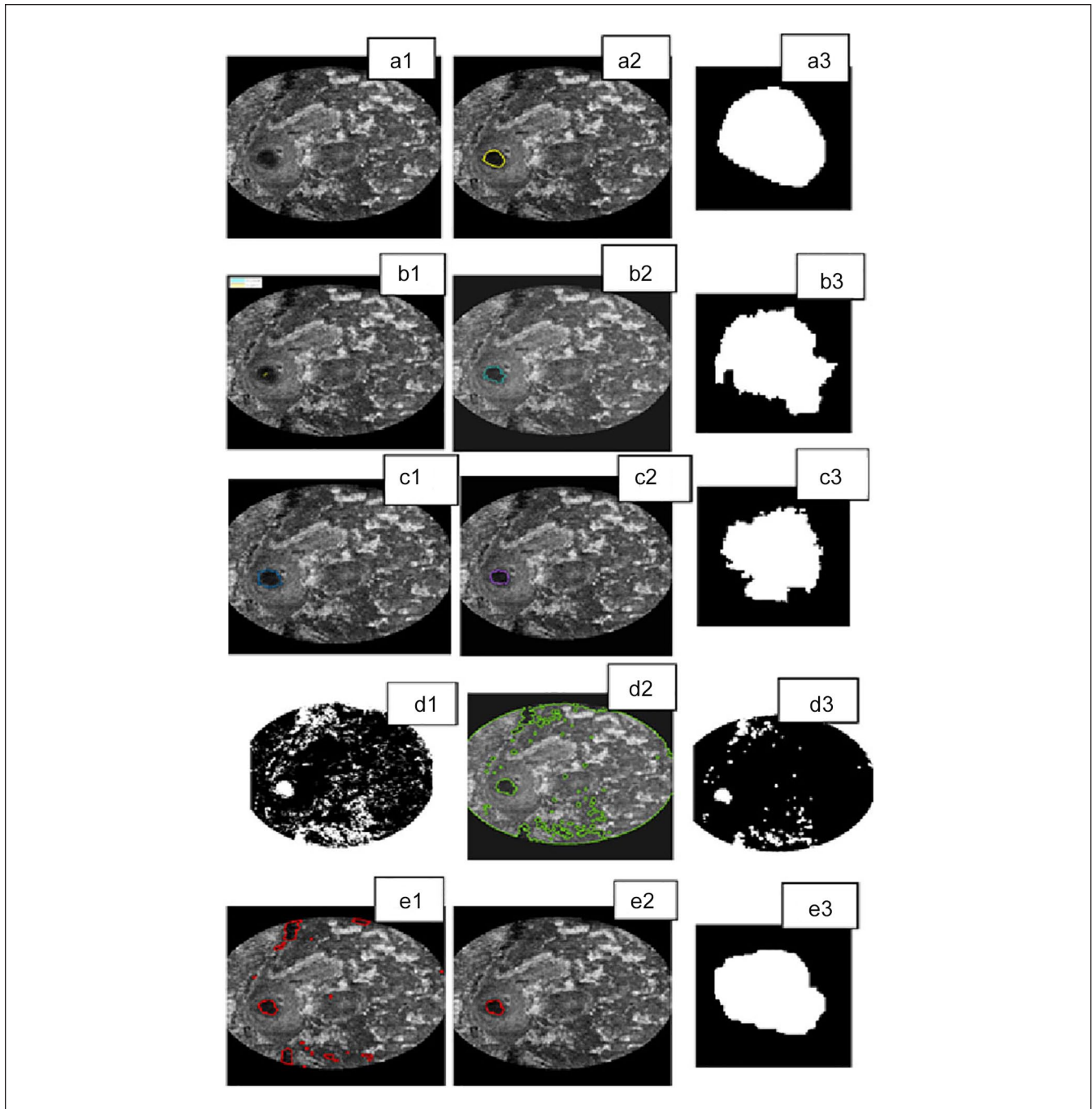
Figure 12. Comparison with different segmentation techniques; Row 1: Original image along with annotations and ground truths shown in Figure 12(a1–a3) respectively. Row 2: Segmentation output obtained using the Active Contours method: Figure 12(b1) illustrates the initialization. Figure 12(b2) denotes the overlaid contour and Figure 12(b3) illustrates the segmented output. Row 3: segmentation output from the Local Graph cut method. Figure 12(c1) denotes the initialization ROI, followed by overlaid segmentation contour and the final segmentation output in Figure 12(c2) and (c3) respectively. Row 4: Presents segmentation results from K mean clustering technique along with the segmented contours and results in Figure 12(d2) and (d3) respectively. Row 5: illustrates the segmentation outputs from the proposed technique where Figure 12(e1) illustrates the probable NSA regions from the ANN followed by the final output obtained from the trained SVM classifier in Figures 12(e3) and (e2) respectively.

Table 7. Comparative Results of DSC with Other Methods.

Method	Performance metric(DSC)
Active contours[39]	0.8045
Local graphcut[40]	0.8286
K means clustering, K = 2[41]	0.4832
Proposed method	0.8539

Declaration of Conflicting Interests

Funding

ORCID iD

Alex Noel Joseph Raj https://orcid.org/0000-0003-1505-3159

Supplemental Material

Supplemental material for this article is available online.

Notes

1. True NSA region (1 per image) was manually annotated and 4 transformations were applied to the training set comprising of 14 images.
2. For the Non-NSA regions, no data augmentation was used but was manually annotated and used for training the SVM.

References

1. Bray F, Ferlay J, Soerjomataram I, Siegel RL, Torre LA, Jemal A. Global cancer statistics 2018: GLOBOCAN estimates of incidence and mortality worldwide for 36 cancers in 185 countries. CA Cancer J Clin. 2018;68(6):394-424.
2. Agarwal G, Ramakant P. Breast cancer care in India: the current scenario and the challenges for the future. Breast Care. 2008;3(1):21-7.
3. Francescatti DS, and Silverstein MJ. (eds.). Breast Cancer: A New Era in Management. New York, NY: Springer Science & Business Media, 2013, pp. 43-6.
4. Kolb TM, Lichy J, Newhouse JH. Comparison of the performance of screening mammography, physical examination, and breast US and evaluation of factors that influence them: an analysis of 27,825 patient evaluations. Radiology. 2002;225(1):165-75.
5. Kotsianos-Hermle D, Hiltawsky KM, Wirth S, Fischer T, Friese K, Reiser M. Analysis of 107 breast lesions with automated 3D ultrasound and comparison with mammography and manual ultrasound. Eur J Radiol. 2009;71(1):109-15.
6. Xu Y, Wang Y, Yuan J, Cheng Q, Wang X, Carson PL. Medical breast ultrasound image segmentation by machine learning. Ultrasonics. 2019;91:1-9.
7. Chae EY, Cha JH, Kim HH, Shin HJ. Comparison of lesion detection in the transverse and coronal views on automated breast sonography. J Ultrasound Med. 2015;34(1):125-35.
8. Panigrahi L, Verma K, Singh BK. Ultrasound image segmentation using a novel multi-scale Gaussian kernel fuzzy clustering and multi-scale vector field convolution. Expert Syst Appl. 2019;115:486-98.
9. Xian M, Zhang Y, Cheng HD, Xu F, Zhang B, Ding J. (2018). Automatic breast ultrasound image segmentation: a survey. Pattern Recognit. 2018;79:340-55.
10. Shin HJ, Kim HH, Cha JH. Current status of automated breast ultrasonography. Ultrasonography. 2015;34(3):165.
11. Wang L, Zoehrer F, Friman O, Hahn H. A fully automatic method for nipple detection in 3D breast ultrasound images. Int J Comput Assist Radiol Surg. 2011;6:S191-2.
12. Cheng HD, Shan J, Ju W, Guo Y, Zhang L. Automated breast cancer detection and classification using ultrasound images: a survey. Pattern Recognit. 2010;43(1):299-317.
13. Tan T, Platel B, Mus R, Tabar L, Mann RM, Karssemeijer N. Computer-aided detection of cancer in automated 3-D breast ultrasound. IEEE Trans Med Imaging. 2013;32(9):1698-706.
14. Méndez AJ, Tahoces PG, Lado MJ, Souto M, Correa J, Vidal JJ. Automatic detection of breast border and nipple in digital mammograms. Comput Methods Programs Biomed. 1996;49(3):253-62.
15. Jas M, Mukhopadhyay S, Chakraborty J, Sadhu A, Khandelwal N. A heuristic approach to automated nipple detection in digital mammograms. J Digit Imaging. 2013;26(5):932-40.
16. Petroudi S, Brady M. Automatic nipple detection on mammograms. In: International Conference on Medical Image Computing and Computer-Assisted Intervention, Springer, Berlin, Heidelberg, November 2003, pp. 971-2.
17. Mustra M, Bozek J, Grgic M. Nipple detection in craniocaudal digital mammograms. In: 2009 International Symposium ELMAR, Zadar, Croatia, 28-30 September 2009, pp. 15-18. New York: IEEE.
18. Popli MB, Teotia R, Narang M, Krishna H. Breast positioning during mammography: mistakes to be avoided. Breast Cancer. 2014;8:119-24.
19. Sweeney RJI, Lewis SJ, Hogg P, McEntee MF. A review of mammographic positioning image quality criteria for the craniocaudal projection. Br J Radiol. 2018;91(1082):20170611.
20. Chae SH, Jeong JW, Choi JH, Chae EY, Kim HH, Choi YW, et al. Fully automated nipple detection in digital breast tomosynthesis. Comput Methods Programs Biomed. 2017;143:113-20.
21. Zhuang Z, Raj ANJ, Jain A, Ruban N, Chaurasia S, Li N, et al. Nipple segmentation and localization using modified U-net on breast ultrasound images. J Med Imaging Health Inform. 2019;9(9):1827-37.
22. Konyar MZ, Ertürk S. Enhancement of ultrasound images with bilateral filter and Rayleigh CLAHE. In: 2015 23nd Signal Processing and Communications Applications Conference (SIU), Malatya, Turkey, 16-19 May 2015, pp.1861-4. New York: IEEE.
23. Mousania Y, Karimi S. Contrast improvement of ultrasound images of focal liver lesions using a new histogram equalization. In: Fundamental Research in Electrical Engineering. Singapore: Springer; 2019, pp. 43-53.
24. Campos GFC, Mastelini SM, Aguiar GJ, Mantovani RG, de Melo LF, Barbon S. Machine learning hyperparameter selection

for contrast limited adaptive histogram Equalization. EURASIP J Image Video Process. 2019;2019(1):59.

25. Tripathi AK, Mukhopadhyay S, Dhara AK. Performance metrics for image contrast. In: 2011 International Conference on Image Information Processing, Shimla, Himachal Pradesh, India, 3-5 November 2011, pp. 1-4. New York: IEEE.

26. Horsch K, Giger ML, Venta LA, Vyborny CJ. Automatic segmentation of breast lesions on ultrasound. Med Phys. 2001;28(8):1652-9.

27. Noble JA, Boukerroui D. Ultrasound image segmentation: a survey. IEEE Trans Med Imaging. 2006;25(8):987-1010.

28. Hu MK. Visual pattern recognition by moment invariants. IEEE Trans Inf Theory. 1962;8(2):179-87.

29. Zhang X, Yang J, Nguyen E. Breast cancer detection via Hu moment invariant and feedforward neural network. In: AIP Conference Proceedings, Ho Chi Minh City, Vietnam, 24-25 May 2018, vol. 1954, no. 1, p. 030014. AIP Publishing LLC.

30. Haralick RM, Shanmugam K, Dinstein IH. Textural features for image classification. IEEE Trans Syst Man Cybern. 1973;6:610-21.

31. Wu K, Shu H, Dillenseger JL. Region and boundary feature estimation on ultrasound images using moment invariants. Comput Methods Programs Biomed. 2014;113(2):446-55.

32. Milosevic M, Jankovic D, Peulic A. Thermography based breast cancer detection using texture features and minimum variance quantization. EXCLI J. 2014;13:1204.

33. Hossain MRI, Ahmed I, Kabir MH. Automatic lung tumor detection based on GLCM features. In: Asian Conference on Computer Vision, Singapore, 1-5 November 2014, pp. 109-21. Springer, Cham.

34. Anuradha K. Statistical feature extraction to classify oral cancers. J Glob Res Comput Sci. 2013;4(2):8-12.

35. Mandelbrot BB. Fractals: Form, Chance, and Dimension, 365 pp., W.H. Freeman, San Francisco (1977) (Revised edition, The Fractal Geometry of Nature, 468 pp., W. H. Freeman, San Francisco, Calif. (1983)).

36. Zhuang Z, Lei N, Raj ANJ, Qiu S. Application of fractal theory and fuzzy enhancement in ultrasound image segmentation. Med Biol Eng Comput. 2019;57(3):623-32.

37. Okubo PG, Aki K. Fractal geometry in the San Andreas fault system. J Geophys Res Solid Earth. 1987;92(B1):345-55.

38. Green G. An Essay on The Application of Mathematical Analysis to The Theories of Electricity and Magnetism. Nottingham: Wezäta-MelinsAktiebolag, 1828.

39. Chan TF, Vese LA. Active contours without edges. IEEE Trans Image Process. 2001;10(2):266-77.

40. Boykov Y, Funka-Lea G. Graph cuts and efficient ND image segmentation. Int J Comput Vis. 2006;70(2):109-31.

41. Wagstaff K, Cardie C, Rogers S, Schrödl S. Constrained k-means clustering with background knowledge. In: ICML, Williamstown, MA, 28 June-1 July, vol. 1, pp. 577-84.

ADID-UNET—a segmentation model for COVID-19 infection from lung CT scans

Alex Noel Joseph Raj[1,*], Haipeng Zhu[1,*], Asiya Khan[2], Zhemin Zhuang[1], Zengbiao Yang[1], Vijayalakshmi G. V. Mahesh[3] and Ganesan Karthik[4]

[1] Department of Electronic Engineering, College of Engineering, Shantou University, Shantou, China
[2] School of Engineering, Computing and Mathematics, University of Plymouth, Plymouth, UK
[3] Department of Electronics and Communication, BMS Institute of Technology and Management, Bangalore, India
[4] COVID CARE - Institute of Orthopedics and Traumatology, Madras Medical College, Chennai, India
* These authors contributed equally to this work.

ABSTRACT

Currently, the new coronavirus disease (COVID-19) is one of the biggest health crises threatening the world. Automatic detection from computed tomography (CT) scans is a classic method to detect lung infection, but it faces problems such as high variations in intensity, indistinct edges near lung infected region and noise due to data acquisition process. Therefore, this article proposes a new COVID-19 pulmonary infection segmentation depth network referred as the Attention Gate-Dense Network- Improved Dilation Convolution-UNET (ADID-UNET). The dense network replaces convolution and maximum pooling function to enhance feature propagation and solves gradient disappearance problem. An improved dilation convolution is used to increase the receptive field of the encoder output to further obtain more edge features from the small infected regions. The integration of attention gate into the model suppresses the background and improves prediction accuracy. The experimental results show that the ADID-UNET model can accurately segment COVID-19 lung infected areas, with performance measures greater than 80% for metrics like Accuracy, Specificity and Dice Coefficient (DC). Further when compared to other state-of-the-art architectures, the proposed model showed excellent segmentation effects with a high DC and F1 score of 0.8031 and 0.82 respectively.

INTRODUCTION

COVID-19 has caused a worldwide health crisis. The World Health Organization (WHO) announced COVID-19 as a pandemic on March 11, 2020. The clinical manifestations of COVID-19 range from influenza-like symptoms to respiratory failure (i.e., diffuse alveolar injury) and its treatment requires advanced respiratory assistance and artificial ventilation. According to the global case statistics from the Center for Systems Science and Engineering (CSSE) of Johns Hopkins University (JHU) (*Wang et al., 2020a*)
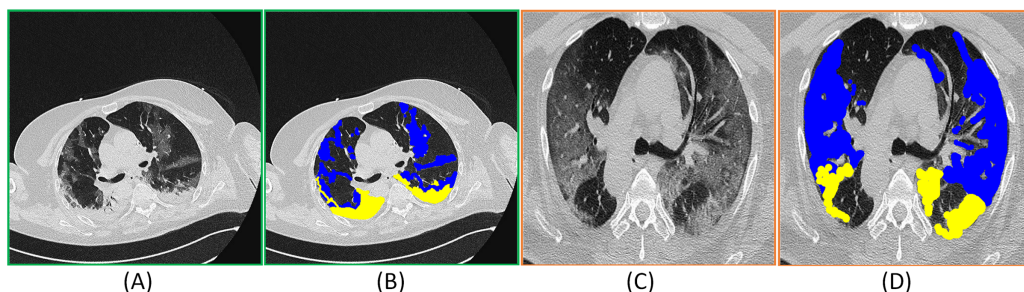
Figure 1 **(A) and (C) represent the CT images and (B) and (D) correspond to the COVID-19 infected areas in CT axial section.** Here the blue and yellow masks represent the Ground Glass Opacity(GGO) and the late lung consolidation segments respectively. The images were obtained from (*MedSeg, 2020*).
Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-1

(updated August 30, 2020), 24,824,247 confirmed COVID-19 cases, including 836,615 deaths, have been reported so far with pronounced effect in more than 180 countries. COVID-19 can be detected and screened by Reverse Transcription Polymerase Chain Reaction (RT-PCR). However, the shortage of equipment and the strict requirements on the detection environment limit the rapid and accurate screening of suspected cases. Moreover, the sensitivity of RT-PCR is not high enough, resulting in a large number of false-negatives (*Ai et al., 2020*), which presents early detection and treatment of patients with presumed COVID-19 (*Fang et al., 2020*). As an important supplement to RT-PCR, CT scans clearly describe the characteristic lung manifestations related to COVID-19 (*Chung et al., 2020*), the early Ground Glass Opacity (GGO), and late lung consolidation are shown in Fig. 1. Nevertheless, CT scans also show imaging features that are similar to other types of pneumonia, making it difficult to differentiate them. Moreover, the manual depiction of lung infection is a tedious and time-consuming job, which is often influenced by personal bias and clinical experience.

In recent years, deep learning has been gaining popularity in the field of medical imaging due to it's intelligent and efficient feature extraction ability (*Kong et al., 2019*; *Ye, Gao & Yin, 2019*), and has achieved great success. An earliest classic example is the application of deep learning to children's chest X-rays to detect and distinguish bacterial and viral pneumonia (*Kermany et al., 2018*; *Rajaraman et al., 2018*). Also using deep learning methods have been applied to detect various imaging features of chest CT images (*Depeursinge et al., 2015*; *Anthimopoulos et al., 2016*). Recently, researchers proposed to detect COVID-19 infections in patients by radiation imaging combined with deep learning technology. *Li et al. (2020)* proposed a simple Cov-Net deep learning network in combination with a deep learning algorithm, which was used to distinguish COVID-19 and Community-Acquired Pneumonia (CAP) from chest CT scans. *Wang & Wong (2020)* proposed Covid-Net to detect COVID-19 cases from chest X-ray images, with an accuracy rate of 93.3%. The infection probability of COVID-19 *Xu et al. (2020)* was calculated from CT scans by adopting a position-oriented attention model that presented accuracy close to 87%. However, the above models rarely involved the segmentation of COVID-19 infection (*Chaganti et al., 2020*; *Shan et al., 2020*). The challenges involved in segmentation

include: (a) variations in texture, size, and position of the infected areas in CT scans. For example, some infection areas are small, which easily lead to a high probability of false negatives in CT scans. (b) The boundary of GGO is usually of low contrast and fuzzy in appearance, which makes it difficult to distinguish from the healthy regions during the segmentation process. (c) The noise around the infected area is high, which greatly affects the segmentation accuracy and (d) finally the cost and time consumed in obtaining high-quality pixel-level annotation of lung infection in CT scans is high. Therefore, most of the COVID-19 CT scan datasets are focused on diagnosis, and only a few of them provide segmentation labels. However, with the passage of time, the annotated datasets for the segmentation of COVID-19 pulmonary infection were released but due to a lesser amount of data, the phenomenon of overfitting could cause problems while training thus necessitating the need for more segmentation datasets and better algorithms for accurate results.

Therefore, to address the challenges stated above, we propose a new deep learning network called Attention Gate-Dense Network- Improved Dilation Convolution-UNET (ADID-UNET) for the segmentation of COVID-19 from lung infection CT scans. Experimental results on a publicly available dataset illustrate that the proposed model presents reliable segmentation results that are comparable to the ground truths annotated by experts. Also, in terms of performance, the proposed model surpasses other state-of-the-art segmentation models, both qualitatively and quantitatively.

Our contributions in this paper are as follows:

1. To address the problem that the gradient disappearance in the deep learning network pose, we employ a dense network (*Huang et al., 2017*) instead of a traditional convolution and max-pooling operations. The dense network extracts dense features and enhances feature propagation through the model. Moreover, the training parameters of the dense network are less, which reduces the size and the computational cost.

2. To increase the size of the respective field and to compensate for the problems due to blurry edges, an improved dilation convolution (IDC) module is used to connect the encoder and decoder pipelines. The IDC model increases the receptive field of the predicted region providing more edge information, which enhances the edge recognition ability of the model.

3. Since the edge contrast of GGO is very low, we use the attention gate (AG) instead of simple cropping and copying. This further improves the accuracy of the model to detect the infection areas by learning the characteristics of the infected regions.

4. Due to the limited number of COVID-19 segmented datasets with segmentation labels, which is less than the minimum number of samples required for training a complex model, we employ data augmentation techniques and expand the dataset on the basis of the collected public datasets.

The rest of the paper is organized as follows: "Related Work" describes the work related to the proposed model. "Methods" introduces the basic structure of ADID-UNET. Details of the dataset, experimental results and discussion are dealt with in "Experience Results". Finally, "Conclusion" presents the conclusion.

RELATED WORK

ADID-UNET model proposed in this paper is based on UNET (*Ronneberger, Fischer & Brox, 2015*) architecture and therefore, we will discuss the literature related to our work which includes: deep learning and medical image segmentation, improvement of medical image segmentation algorithms, CT scan segmentation, and application of deep learning in segmentation of COVID-19 lesions from lung CT scans.

Deep learning and medical image segmentation

In recent years, deep learning algorithms have become more mature leading to various artificial intelligence (AI) systems based on deep learning algorithms being developed. Also, semantic segmentation using deep learning algorithms (*Oktay et al., 2018*) has developed rapidly with applications in both natural and medical images. *Long, Shelhamer & Darrell (2015)* pioneered the use of a fully connected CNN (FCN) to present rough segmentation outputs that were of the input resolution through fractionally strided convolution process also referred as the upsampling or deconvolution. The model was tested on PASCAL VOC, NYUDv2, and SIFT datasets and, presented a Mean Intersection of Union (M-IOU) of 62.7%, 34%, 39.5%, respectively. They also reported that upsampling, part of the in-network, was fast, accurate, and provided dense segmentation predictions. Later through a series of improvements and extensions to FCN (*Ronneberger, Fischer & Brox, 2015*; *Badrinarayanan, Kendall & Cipolla, 2017*; *Xu et al., 2018*), a symmetrical structure composed of encoder and decoder pipelines, called UNET (*Ronneberger, Fischer & Brox, 2015*), was proposed for biomedical or medical image segmentation. The encoder structure predicted the segmentation area, and then the decoder recovered the resolution and achieved accurate spatial positioning. Also, the UNET used crop and copy operations for the precise segmentation of the lesions. Further, the model achieved good segmentation performance at the International Symposium on Biomedical Imaging (ISBI) challenge (*Cardona et al., 2010*) with the M-IOU of 0.9203. Moreover, an improved network referred as the SegNet was proposed by *Badrinarayanan, Kendall & Cipolla (2017)*. The model used the first 13 convolution layers of the VGG16 network (*Karen & Andrew, 2014*) to form an encoder to extract features and predict segmentation regions. Later by using a combination of convolution layers, unpooling and softmax activation function in the decoder, segmentation outputs of input resolution were obtained. When tested with the CamVid dataset (*Brostow, Fauqueur & Cipolla, 2009*), the M-IOU index of SegNet was nearly 10% higher than that of FCN (*Long, Shelhamer & Darrell, 2015*). *Xu et al. (2018)* regarded segmentation as a classification problem in which each pixel was associated with a class label and designed a CNN network composed of three layers of convolution and pooling, a fully connected layer (FC) and softmax function. The model of successfully segmented three-dimensional breast ultrasound (BUS) image datasets was presented into four parts: skin, fibroglandular tissue, mass, and fatty tissue and achieved a recall rate of 88.9%, an accuracy of 90.1%, precision of 80.3% and F1 score of 0.844. According to the aforementioned literature, FCN (*Long, Shelhamer & Darrell, 2015*) and their improved variants presented accurate segmentation results for both natural or medical images. Therefore, the UNET and variants (*Almajalid et al., 2019*; *Negi et al., 2020*),

due to its advantages of fast training and high segmentation accuracy are widely used in the field of medical image segmentation.

Improvement of medical image segmentation algorithms

Medical images such as the ultrasound images are generally prone to speckle noise, uneven intensity distribution, and low contrast between the lesions and the backgrounds which affect the segmentation ability of the traditional UNET (*Ronneberger, Fischer & Brox, 2015*) structure. Therefore, considerable efforts were invested in improving the architecture. *Xia & Kulis (2017)* proposed a fully unsupervised deep learning network called W-Net model that connects two UNETs to predict and reconstruct the segmentation results. *Schlemper et al. (2019)* proposed an attention UNET network, which integrated attention modules into the UNET (*Ronneberger, Fischer & Brox, 2015*) model to achieve spatial positioning and subsequent segmentation. The model presented a segmentation accuracy of 15% higher than the traditional UNET architecture. *Zhuang et al. (2019a)* combined the goodness of the attention gate system and the dilation convolution module and proposed a hybrid architecture referred as the RDA-UNET. By introducing residual network (*He et al., 2016*) instead of traditional convolution layers they reported a segmentation accuracy of 97.91% towards the extraction of lesions in breast ultrasound images. Also, the GRA-UNET (*Zhuang et al., 2019b*) model included a group convolution module in-between the encoder and decoder pipelines to improve the segmentation of the nipple region in breast ultrasound images. Therefore, from the literature, it can be inferred that introducing additional modules like attention gate instead of traditional cropping and copying, inclusion of dilation convolution to increase the receptive fields and use of residual networks can favorably improve the accuracy of the segmentation model. However, these successful segmentation models (*Schlemper et al., 2019*; *Zhuang et al., 2019a*; *Xia & Kulis, 2017*) were rarely tested with CT scans, hence the next section concentrates on the segmentation of CT scans.

CT scan segmentation

CT imaging is a commonly used technology in the diagnosis of lung diseases since lesions can be segmented more intuitively from the chest CT scans. The segmented lesion aid the specialist in the diagnosis and quantification of the lung diseases (*Gordaliza et al., 2018*). In recent years, most of the classifier models and algorithms based on feature extraction have achieved good segmentation results in chest CT scans. *Ye et al. (2009)* proposed a shape-based Computer-Aided Detection (CAD) method where a 3D adaptive fuzzy threshold segmentation method combined with chain code was used to estimate infected regions in lung CT scans. In feature-based techniques, due to the low contrast between nodules and backgrounds, the boundary discrimination is unclear leading to inaccurate segmentation results. Therefore, many segmentation techniques based on deep learning algorithms have been proposed. *Wang et al. (2017)* developed a central focusing convolutional neural network for segmenting pulmonary nodules from heterogeneous CT scans. *Jue et al. (2018)* designed two deep networks (an incremental and dense multiple resolution residually connected network) to segment lung tumors from

CT scans by adding multiple residual flows with different resolutions. *Guofeng et al. (2018)* proposed a UNET model to segment pulmonary nodules in CT scans which improved the overall segmentation output through the avoidance of overfitting. Compared with other segmentation algorithms such as graph-cut (*Ye, Beddoe & Slabaugh, 2009*), their model had better segmentation results with a Dice coefficient of 0.73. Recently, *Peng et al. (2020)* proposed an automatic CT lung boundary segmentation method, called Pixel-based Two-Scan Connected Component Labeling-Convex Hull-Closed Principal Curve method (PSCCL-CH-CPC). The model included the following: (a) the image preprocessing step to extract the coarse lung contour and (b) coarse to finer segmentation algorithm based on the improved principal curve and machine learning model. The model presented good segmentation results with Dice coefficient as high as 96.9%. *Agarwal et al. (2020)* proposed a weakly supervised lesion segmentation method for CT scans based on an attention-based co-segmentation model (*Mukherjee, Lall & Lattupally, 2018*). The encoder structure composed of a variety of CNN architectures that includes VGG-16 (*Karen & Andrew, 2014*), Res-Net101 (*He et al., 2016*), and an attention gate module between the encoder-decoder pipeline, while decoder composed of upsampling operation. The proposed method first generated the initial lesion areas from the Response Evaluation Criteria in Solid Tumors (RECIST) measurements and then used co-segmentation to learn more discriminative features and refine the initial areas. The paper reported a Dice coefficient of 89.8%. The above literatures suggest that deep learning techniques are effective in segmenting lesions in lung CT scans and many researchers have proposed different deep learning architectures to deal with COVID-19 CT scans. Therefore, in the next section we will further study their related works.

Application of deep learning in segmentation of COVID-19 lesions from lung CT scans

In recent months, COVID-19 has become a hot topic of concern all over the world and CT imaging is considered to be a convincing method to detect COVID-19. However, due to the limited datasets and the time and labor involved in annotations, segmentation datasets related to COVID-19 CT scans are less readily available. But, many researchers have still proposed advanced methods to deal with COVID-19 diagnosis, which also includes segmentation techniques (*Fan et al., 2020*; *Wang et al., 2020b*; *Yan et al., 2020*; *Zhou, Canu & Ruan, 2020*; *Elharrouss et al., 2020*; *Chen, Yao & Zhang, 2020*). On the premise of insufficient datasets with segmentation labels, the Inf-Net network proposed by *Fan et al. (2020)*, combined a semi-supervised learning model and FCN8s network (*Long, Shelhamer & Darrell, 2015*) with implicit reverse attention and explicit edge attention mechanism to improve the recognition rate of infected areas. The model successfully segmented COVID-19 infected areas from CT scans and reported a sensitivity and accuracy of 72.5% and 96.0%, respectively. *Elharrouss et al. (2020)* proposed an encoder-decoder-based CNN method for COVID-19 lung infection segmentation based on a multi-task deep-learning based method, which overcame the shortage of labeled datasets, and segmented lung infected regions with a high sensitivity of 71.1%. *Wang et al. (2020b)* proposed a noise-robust COVID-19 pneumonia lesions segmentation network

which included a noise-robust dice loss function along with convolution function, residual network, and Atrous Spatial Pyramid Pooling (ASPP) module. The model was referred as Cople-Net presented automatic segmentation of COVID-19 pneumonia lesions from CT scans. The method proved that the proposed new loss function was better than the existing noise-robust loss functions such as Mean absolute error (MAE) loss (*Ghosh, Kumar & Sastry, 2017*) and Generalized Cross-Entropy (GCE) loss (*Zhang & Sabuncu, 2018*) and achieved a Dice coefficient and Relative Volume Error (RVE) of 80.72% and 15.96%, respectively. *Yan et al. (2020)* employed an encoder-decoder deep CNN structure composed of convolution function, Feature Variation (FV) module (mainly contains convolution, pooling, and sigmoid function), Progressive Atrous Spatial Pyramid Pool (PASPP) module (including convolution, dilation convolution, and addition operation) and softmax function. The convolution function obtained features, FV block enhanced the feature representation ability and the PASPP was used between encoder and decoder pipelines compensated for the various morphologies of the infected regions. The model achieved a good segmentation performance with a Dice coefficient of 0.726 and a sensitivity of 0.751 when tested on the COVID-19 lung CT scan datasets. *Zhou, Canu & Ruan (2020)* proposed an encoder-decoder structure based UNET model for the segmentation of the COVID-19 lung CT scan. The encoder structure was used to extract features and predict rough lesion areas which composed convolution function and Res-dil block (combines residual block (*He et al., 2016*) and dilation convolution module). The decoder pipeline was used to restore the resolution of the segmented regions through the upsampling and the attention mechanism between the encoder-decoder framework to capture rich contextual relationships for better feature learning. The proposed method can achieve an accurate and rapid segmentation on COVID-19 lung CT scans with a Dice coefficient, sensitivity, and specificity of 69.1%, 81.1%, and 97.2%, respectively. Further, *Chen, Yao & Zhang (2020)* proposed a residual attention UNET for automated multi-class segmentation of COVID-19 lung CT scans, which used residual blocks to replace traditional convolutions and upsampling functions to learn robust features. Again, a soft attention mechanism was applied to improve the feature learning capability of the model to segment infected regions of COVID-19. The proposed model demonstrates a good performance with a segmentation accuracy of 0.89 for lesions in COVID-19 lung CT scans. Therefore, the deep learning algorithms are helpful in segmenting the infected regions from COVID-19 lung CT scans which aid the clinicians to evaluate the severity of infection (*Tang et al., 2020*), large-scale screening of COVID-19 cases (*Shi et al., 2020*) and quantification of the lung infection (*Ye et al., 2020*). Table 1 summarizes the deep learning-based segmentation techniques available for COVID-19 lung infections.

METHODS

In this section, we first introduce the proposed ADID-UNET network with detailed discussion on the core network components including dense network, improved dilation convolution, and attention gate system. To present realistic comparisons, experimental results are presented at each subsection to illustrate the performance and superiority of the

Table 1 The summary of various deep learning algorithms for COVID-19 lung CT scans and the segmentation results. RVE, ACC, DC, Sen, Sp and F1 represent relative volume error, accuracy, Dice coefficient, sensitivity, specificity and F1 score, respectively.

Literature	Data Type	Dataset	Technique	Segmentation results
Fan et al. (2020)	CT Scan	100 CT images	Semi supervised CNN	73.9% (DC)
			FCN8s network	96.0% (S_p)
Wang et al. (2020b)	CT Scan	558 CT images	Residual connection	80.7% (DC)
			CNN	16.0% (RVE)
Yan et al. (2020)	CT Scan	21,658 CT images	Deep CNN	72.6% (DC)
				75.1% (S_{en})
Zhou, Canu & Ruan (2020)	CT Scan	100 CT images	Attention mechanism	69.1% (DC)
			Res-Net, dilation convolution	81.1% (S_{en})
Elharrouss et al. (2020)	CT Scan	100 CT images	Encoder-decoder-based CNN	78.6% (Dice)
				71.1% (S_{en})
Chen, Yao & Zhang (2020)	CT Scan	110 CT images	Encoder-decoder-based CNN	83.0% (DC)
				89.0% (ACC)
Xu et al. (2020)	CT Scan	110 CT images	CNN	86.7% (ACC)
				83.9% (F1)
Shuai et al. (2020)	CT Scan	670 CT images	CNN	73.1% (ACC)
				67.0% (S_p)

model after adding core components. Further in "Experience Results" we have presented a summary of the % improvements achieved when compared to the traditional UNET architecture.

ADID-UNET architecture

ADID-UNET is based on UNET (*Ronneberger, Fischer & Brox, 2015*) architecture with the following improvements: (a) The dense network proposed by *Huang et al. (2017)* is used in addition to the convolution modules of encoder and decoder structures, (b) an improved dilation convolution (IDC) is introduced between the frameworks, and (c) the attention gate (AG) system is used instead of the simple cropping and copying operations. The structure of ADID-UNET is shown in Fig. 2. Here f_{en}, f_{upn}, f_{idc} describe the features at the n-th layer of the encoder, decoder, and IDC modules, respectively.

When COVID-19 CT scans are presented to the encoder, the first four layers (each layer has convolutions, rectification, and max pooling functions) extract features (f_1–f_4) that are passed to dense networks. Here dense networks are used instead of convolution and max-pooling layers to further enhance the features (f_5–f_6) and in "Dense Network", we elaborate the need for the dense network and present experimental results to prove its significance. Next, an improved dilation convolution module referred as the IDC model, is used between the encoder–decoder structure to increase the receptive field and gather detailed edge information that assists in extracting the characteristic. The module accepts the feature f_6 from the dense networks and after improvement, present f_{idc} them as inputs to the decoder structure. To ensure consistency in the architecture and to avoid losing information, the decoder mirrors the encoder with two dense networks that replace the
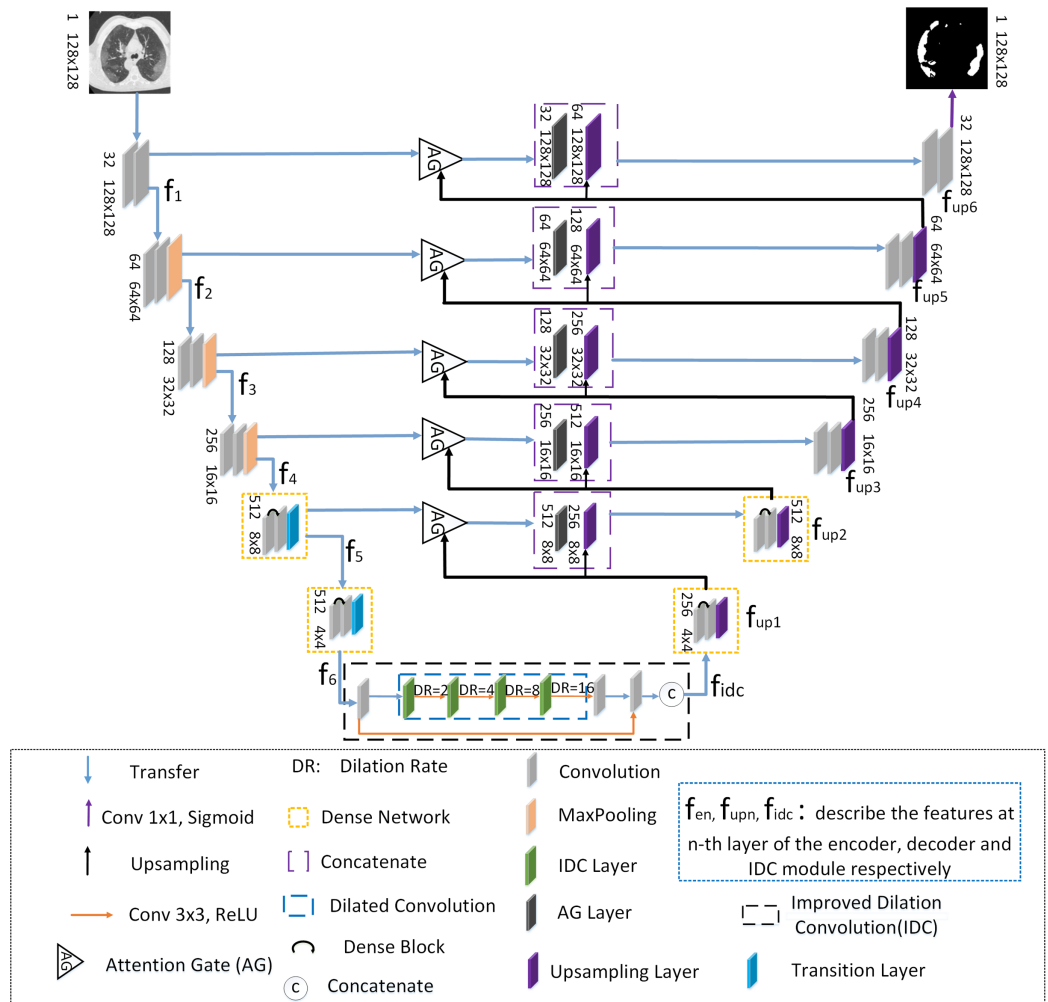
Figure 2 The structure of the ADID-UNET network. The blue, purple, black and orange arrows represent transfer function, convolution function with 1×1 convolution kernel and sigmoid function, upsampling function, convolution function with 3×3 convolution kernel and RELU function, respectively. The triangle represents the attention gate system. The orange, purple, blue and black dotted boxes represent the dense network, concatenate function, dilation convolution and improved dilation convolution, respectively. The curved black arrows within the orange rectangle indicate the dense block. The C within the circle represents concatenate function. The gray, orange, green, black, purple and blue squares represent convolution function, maximum pooling function, improved void convolution layer, attention gate layer, upsampling layer and transition layer, respectively, and describe the features at n-th layer of the encoder, decoder and IDC module, respectively.

Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-2

first two upsampling operations. Further for the better use of the context information between the encoder-decoder pipeline, the AG model is used instead of cropping and copying operations, which aggregates the corresponding layer-wise encoder features with the decoder and presents it to the subsequent upsampling layers. Likewise, the decoder framework presents upsampled features f_{up1} to f_{up6} and final feature map (f_{up6}) is presented to the sigmoid activation function to predict and segment the COVID-19 lung infected regions. The following section explains the components of ADID-UNET in detail.
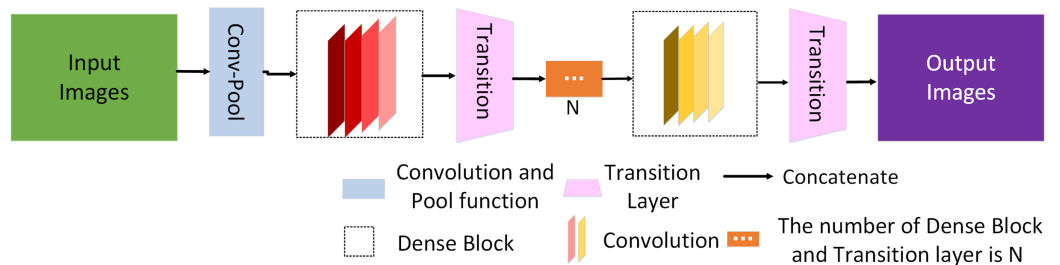
Figure 3 Schematic diagram of the dense network. Green and purple squares represent the input image and the output image, respectively. The blue rectangles represent convolution function and pooling function. The dashed black box represents the dense block and the red and yellow rectangle module within represent convolution function. The pink trapezoidal clock represents the transition function and N represents the number of Dense Block and Transition layer is N.

Full-size ⬛ DOI: 10.7717/peerj-cs.349/fig-3

Dense network

It was presumed that with the increase of network layers, the learning ability of the network will gradually improve, but during the training, for deep networks, the gradient information that is helpful for the generalization may disappear or expand excessively. In literature, the problem is referred as vanishing or explosion of the gradient. As the network begins to converge, due to the disappearance of the gradient the network saturates, resulting in a sharp decline in network performance. Therefore, *Zhuang et al. (2019a)* introduced residual units proposed by *He et al. (2016)* into UNET structure to avoid performance degradation during training. The residual learning correction scheme to avoid performance degradation is described in (1):

$$y = G(x, \{F_i\}) + x \tag{1}$$

Here x and y are the input and output vectors of the residual block, F_i is the weight of the corresponding layer. The function $G(x, \{F_i\})$ is a residue when added to x, avoids vanishing gradient problems, and enables efficient learning.

From (1) the summation of $G(x, \{F_i\})$ and x in Res-Net (*He et al., 2016*) avoids the vanishing gradient problems but forwarding the gradient information alone to the proceeding layers may hinder the information flow in the network and the recent work by *Huang et al. (2016)* illustrated that of Res-Nets discard features randomly during training. Moreover, Res-Nets include large number of parameters, which increases the training time. To solve this problem, *Huang et al. (2017)* proposed a dense network (as shown in Fig. 3), which directly connects all layers, and thus skillfully obtains all features of the previous layer without convolution.

The dense network is mainly composed of convolution layers, pooling function, multiple dense blocks, and transition layers. Let us consider a network with L layers, and each layer implements a nonlinear transformation H_i. Let x_0 represent the input image, i represents layer i, x_{i-1} is the output of layer $i - 1$. H_i can be a composite operation, such as

X₀ : The Input Images

Xᵢ : Output Images of Layer i

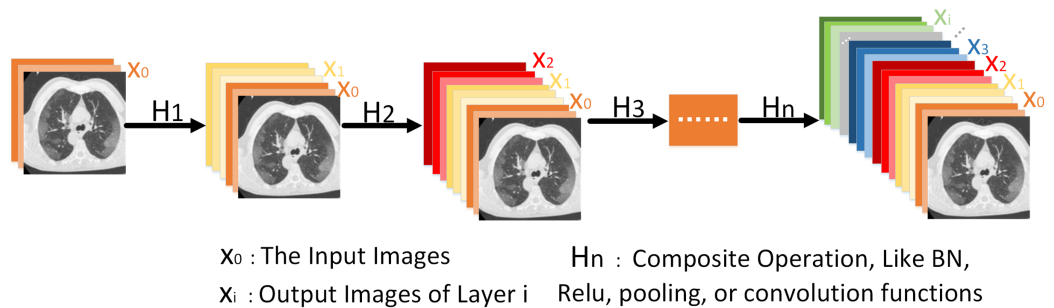Hₙ : Composite Operation, Like BN, Relu, pooling, or convolution functions

Figure 4 It is more intuitive to understand the forward connection mode of a dense network. The forward connection model of the Dense Network, illustrating that the output x_i includes inputs from $x_0, x_1,\ldots,x_{i-1}$.   Full-size ⊡ DOI: 10.7717/peerj-cs.349/fig-4

batch normalization (BN), rectified linear function (RELU), pooling, or convolution functions. Generally, the output of traditional network in layer i is as follows:

$$x_i = H_i \times (x_{i-1}) \tag{2}$$

For the residual network, only the identity function from the upper layer is added:

$$x_i = H_i \times (x_{i-1}) + x_{i-1} \tag{3}$$

For a dense network, the feature mapping $x_0, x_1,\ldots, x_{i-1}$ of all layers before layer i is directly connected, which is represented by Eq. (4):

$$x_i = H_i \times ([x_0, x_1, \ldots, x_{i-1}]) \tag{4}$$

where $[x_0, x_1, \ldots, x_{i-1}]$ denotes the cascade of characteristic graphs and $\times$ represents the multiplication operation. Figure 4 shows the forward connection mechanism of the dense network where the output of layers is connected directly to all previous layers.

Generally, a dense network is composed of several dense blocks and transition layers. Here we only use two dense blocks and transition layers to form simple dense networks. Using Eq. (5) to express the dense block:

$$\gamma = \alpha([x_0, x_1, \ldots, x_{i-1}], \beta_i) \tag{5}$$

where $[x_0, x_1, \ldots, x_{i-1}]$ denotes the cascade of characteristic graphs, β_i is the weight of the corresponding layer. In the ADID-UNET model proposed in this paper, the feature f_4 (refer to Fig. 2) is fed to the transition layer, which is mainly composed of BN, RELU, and average pooling operation. Later the feature is batch standardized and rectified before convolving with a 1×1 kernel function. Again, the filtered outputs go through the same operation and are convoluted with 3×3 kernel, before concatenating with the input feature f_4. The detailed structure of the two dense blocks and transition layers used in the encoder structure is shown in Fig. 5A. Here w, h correspond to the width and height of the input, respectively, and b represents the number of channels. Besides, s represents the step size of the pooling operation, n represents the number of filtering operations performed by each layer. In our model, n takes values 32, 64, 128, 256, and 512.

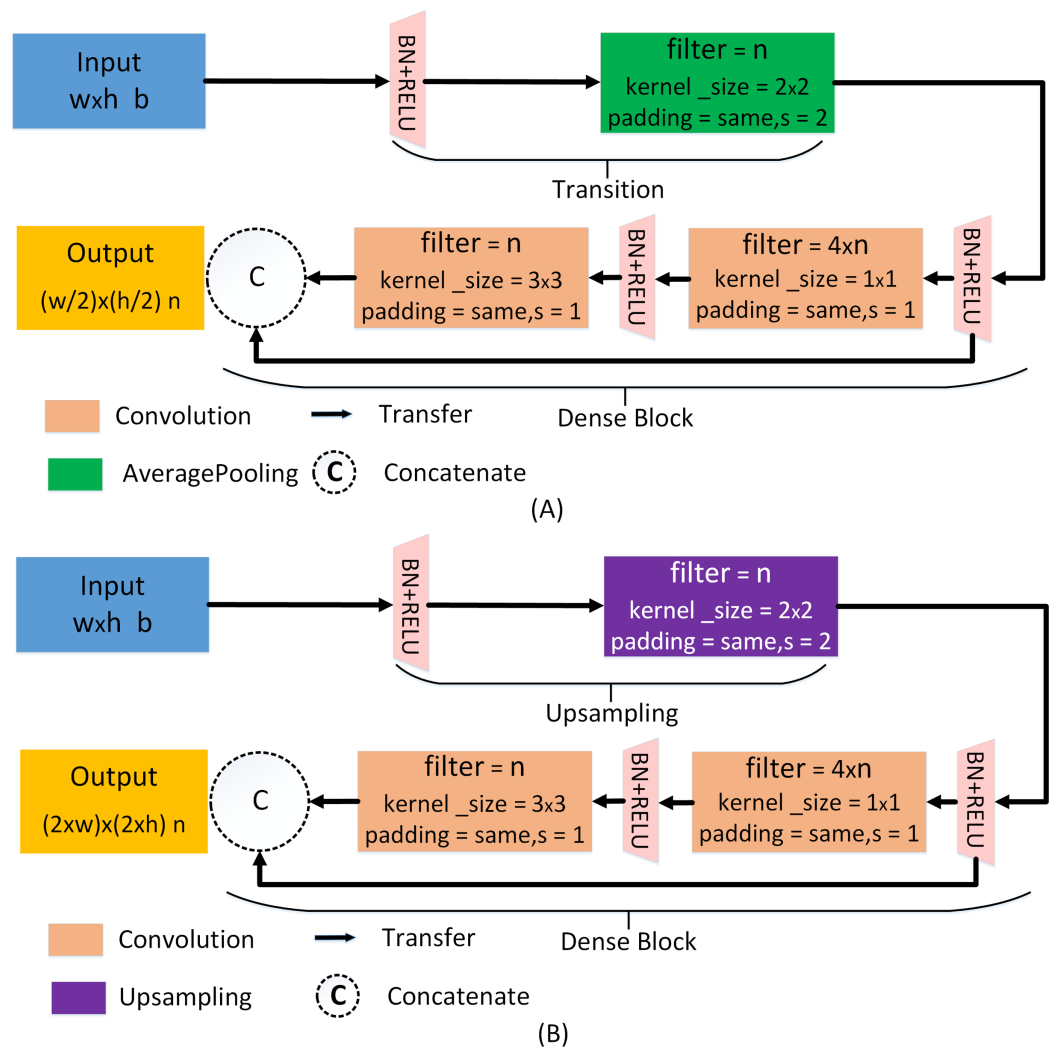It should be noted that the output of the first dense layer is the aggregated result of 4

Figure 5 Dense network for encoder and decoder pipelines. (A) The dense network of the encoder pipeline. (B) The dense network of the decoder pipeline. Here, w, h, b, s, and n correspond to the width and height of the input, the number of channels, the step size of the pooling operation, and the number of filtering operations performed by each layer, respectively. For layers 6, 5, 4, 3, 2, 1, the values of n are 512, 512, 256, 128, 64, and 32, respectively. Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-5

convolution operations ($4 \times n$), which is employed to emphasize the features learning by reducing the loss of features. In the decoding structure, to restore the resolution of the predicted segmentation, a traditional upsampling layer of the UNET (*Ronneberger, Fischer & Brox, 2015*) is used instead of the transition layer. The detailed structure is shown in Fig. 5B.

For the proposed network, we use only two dense networks mainly (a) to reduce the computation costs and (b) experiments with different layers of dense networks suggest that the use of two dense networks was sufficient since the segmentation results were accurate and comparable to the ground truth. Figure 6 and Table 2 illustrate the qualitative and quantitative comparisons with different numbers of dense network in the encoder-decoder framework.
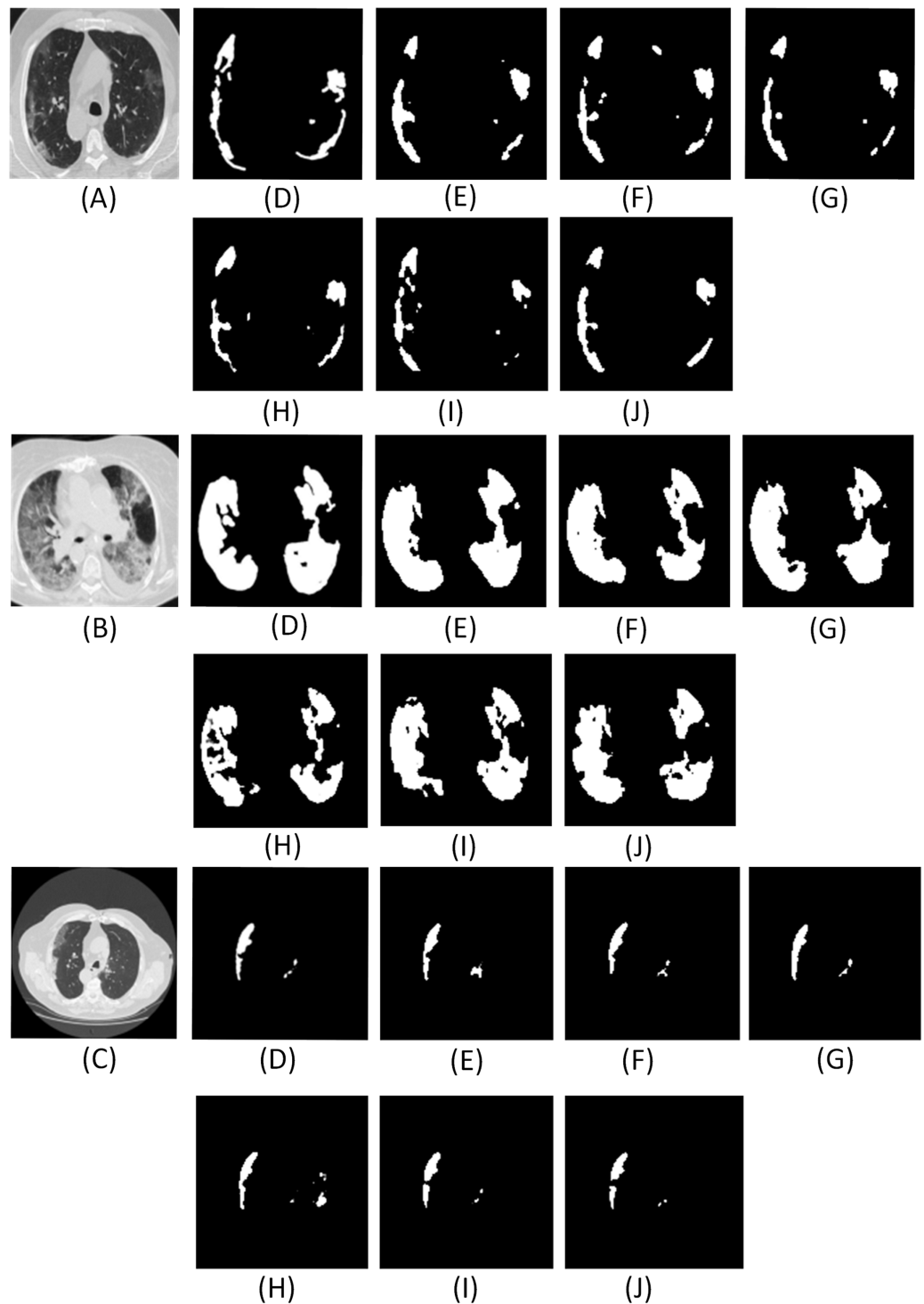
Figure 6 Results of adding different numbers of dense networks. Here (A–C) are the CT scans. For each row, (D) denotes the ground truth, (E–J) illustrate the segment results from UNET, Num1, Num2, Num3, Num4 and Res-Net, respectively. Num1~Num4 denotes the number of dense networks used in the encoder and decoder pipelines. Res-Net refers to the network where convolution operations are replaced by the residual network (*He et al., 2016*). UNET (*Ronneberger, Fischer & Brox, 2015*) denotes the traditional architecture without dense network. Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-6

Table 2 Quantitative comparisons with respect to ground truth for different dense layers included the UNET (*Ronneberger, Fischer & Brox, 2015*) and Res-Net (*He et al., 2016*) architecture. ACC, DC, Sen, Sp, Pc, AUC, F1, Sm, Ea and MAE represent accuracy, Dice coefficient, sensitivity, specificity, precision, the area under the curve, F1 score, structural metric, enhancement alignment meter and mean absolute error, respectively.

Number of dense network	ACC	DC	S_{en}	S_p	P_c	AUC	F1	S_m	E_α	MAE
Num_1	0.9696	0.7971	0.8011	0.9958	0.8290	0.9513	0.8129	0.8411	0.9315	0.0088
Num_2	0.9700	0.8011	0.8096	0.9966	0.8596	0.9492	0.8184	0.8528	0.9394	0.0083
Num_3	0.9686	0.7569	0.7546	0.9957	0.8200	0.9334	0.7806	0.8349	0.9379	0.0104
Num_4	0.9699	0.7869	0.7579	0.9961	0.8485	0.9495	0.8241	0.8341	0.9348	0.0090
UNET	0.9696	0.7998	0.8052	0.9957	0.8247	0.9347	0.8154	0.8400	0.9390	0.0088
Res-Net	0.9698	0.8002	0.7978	0.9962	0.8344	0.9504	0.8180	0.8415	0.9352	0.0094

From the analysis of results in Fig. 6 and Table 2, it is found that the effect of using two dense networks in the model is obvious and can present accurate segments of the infected areas that can be inferred directly from the qualitative and quantitative metrics.

Moreover, with high accuracy and a good Dice coefficient, the choice of two dense networks is the best choice in the encoder decoder pipeline. Also, using two dense networks in place of traditional convolutions or residual networks enable global feature propagation, encourage feature reuse, and also solve the gradient disappearance problems associated with deep networks thereby significantly improving the segmentation outcomes.

Improved dilation convolution

Since the encoder pipeline of the UNET structure is analogous to the traditional CNN architecture, the pooling operations involved at each layer propagate either the maximum or the average characteristics of the extracted features, hence connecting the encoder outputs directly to decoder, thus limiting the segmentation accuracy of the network. The RDA-UNET proposed by *Zhuang et al. (2019a)* utilized a dilation convolution (DC) module between the encoder-decoder pipeline to increase the receptive field and further learn the boundary information accurately. Also, the DC module is often used in many variant UNETs (*Chen et al., 2019*; *Yu & Koltun, 2015*) to improve the receptive field, hence, we use the DC module and introduce additional novelty in the DC module.

Equation (6) describes the DC operation between the input image $f(x, y)$ and the kernel $g(i, j)$.

$$p(x, y) = \alpha \left\{ \sum_{i,j} f(x + i \times r, y + j \times r) \times g(i, j) + k \right\} \qquad (6)$$

where α is the RELU function, k is a bias unit (i, j) and (x, y) denote the coordinates of the kernel and those of the input images respectively, and r is the dilation rate that controls the size of receptive fields. The size of the receptive field obtained can be expressed as follows:

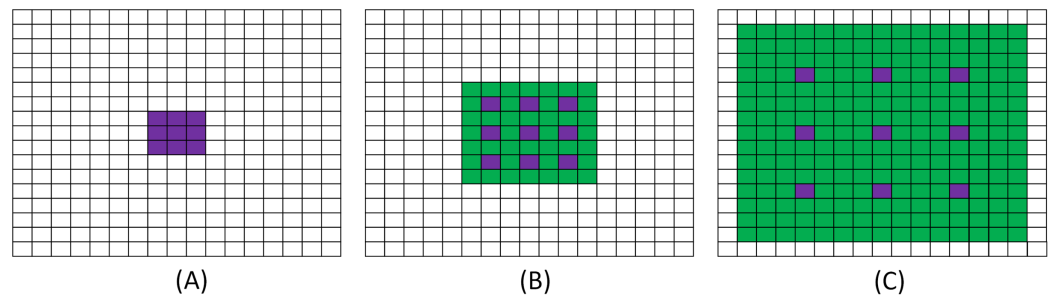$$N = ((k\_f_{\text{size}} + 1) \times (r - 1) + k\_f_{\text{size}}) \qquad (7)$$

Figure 7 Schematic diagram of dilation convolution. (A) It shows the visual field effect of the classical 3 × 3 convolution kernel, covering 3 × 3 field of view each time (purple part of figure (A)); (B) corresponds to 3 × 3 with $r = 2$. Though the size of the convolution kernel is still 3 × 3, but the receptive field of convolution kernel is increased to 7 × 7 (purple and green parts in figure (B)); (C) corresponds to 3 × 3 with $r = 3$ and a receptive field of 15 × 15. Full-size ⬛ DOI: 10.7717/peerj-cs.349/fig-7
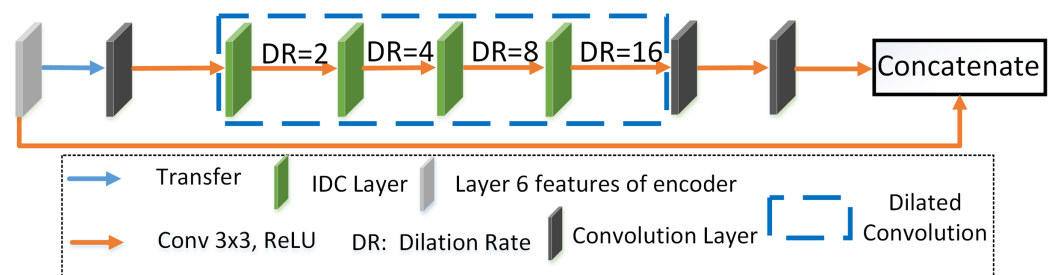
Figure 8 Improved dilation convolution (IDC). The gray, black and green rectangles represent layer 6 features of the encoder, convolution layer and improved dilation convolution layer, respectively. The blue dotted box indicates the dilation convolution. Blue and orange arrows represent the transfer function and the convolution function with the convolution kernel of 3×3 and the RELU function, respectively. Full-size ⬛ DOI: 10.7717/peerj-cs.349/fig-8

where k_f_{size} is the convolution kernel size, r is the convolution rate of the dilation and N is the size of the receptive field. As shown in Fig. 7.

Based on our experimental analysis we understand that DC module has a pronounced effect in extracting information for larger objects or lesions and considering that most of the early ground-glass opacity (GGO) or late lung consolidation lesions have smaller areas, we present an improved dilation convolution (IDC) module between the encoder–decoder framework to accurately segment smaller regions.

Figure 8 illustrates the IDC module that consists of several convolution functions with different dilation rates and rectified linear functions (RELU). Our improvements are as follows: (a) combining single strided convolution operations and dilated convolutions with dilation rate such as 2, 4, 8, and 16, respectively. The above combination helps in the extraction of features from both smaller and larger receptive fields thus assisting in the isolation of the small infected COVID-19 regions seen in lung CT scans and (b) referring to the idea of the dense network (*Huang et al., 2017*), we concatenate the input of the IDC module to its output and use the information of input features to further enhance feature learning. The input of IDC module is the rough segmentation regions obtained by encoder structure. The combination of the original segmentation region features

and the accurate features extracted by IDC module not only avoids the loss of useful information, but also provides accurate input for the decoding pipeline, which is conducive to improve the segmentation accuracy of the model. As the inputs advance (left to right in Fig. 8), they get convolved with a 3×3 kernel of convolution layers and the dilation rate of IDC is 2, 4, 8, and 16, respectively. From the comparative experiments with the traditional DC model (the dilation rate is the same for both the models), we find that the computational cost and computation time required for the IDC module is less than that of the DC module, as shown in Table 3.

From Fig. 9 and Table 4, it is found that the use of layers with convolution and smaller dilation rates at the end along with others ensures the cumulative extraction of features from both smaller and larger receptive fields thus assisting in the isolation of the small infected COVID-19 regions seen in lung CT scans. Also, the performance scores specifically the Dice coefficient is higher (about 3%) for DID-UNET compared to DD-UNET. In summary, the IDC model connected between the encoder–decoder structure, reduces loss of the original features but additionally expands the field of the segmented areas thereby improving the overall segmentation effect.

Attention gate

Although the improved dilation convolution improves the feature learning ability of the network, due to the loss of spatial information in the feature mapping at the end of the encoder structure, the network has difficulties in reducing false prediction for (a) small COVID-19 infected regions and (b) areas with blurry edges with poor contrast between the lesion and background. To solve this problem, we introduce the attention gate (AG) model shown in Fig. 10 mechanism into our model instead of simple cropping and copying. AG model computes the attention coefficient $\sigma \in [0, 1]$, based on Eq. (8):

$$\sigma = \varepsilon_2 \left\{ p_k \left[p_i \left(\varepsilon_1 \left(p_n \times n + p_m \times m + b_{m,n} \right) \right) + b_{int} \right] + b_k \right\} \tag{8}$$

$$\varepsilon_2(x) = \frac{1}{1 + \exp(-x)} \tag{9}$$

where n and m represent the feature mapping of the AG module input from the decoder and encoder pipelines, respectively. And p_m, p_n, p_i, p_k are the convolution kernels of size 1×1. $b_{m,n}$, b_{int}, b_k represent the offset unit. ε_1 and ε_2 denote the RELU and sigmoid activation function respectively. Here ε_2 limits the range between 0 and 1.

Finally, the attention coefficient σ is multiplied by the input feature map f_i to present the output g_o as shown in Eq. (10):

$$g_o = \sigma \times f_i \tag{10}$$

From Fig. 11 and Table 5, results showed that the inclusion AG module improved the performance of the network (ADID-UNET), with segmentation accuracy of almost 97%. Therefore, by introducing the AG model, the network makes full use of the output feature information of encoder and decoder, which greatly reduces the probability of

Table 3 Comparison results of the number of parameters of the UNET model with dense network incorporating either improved dilation convolution module (IDC) or dilation convolution module. Here DID-UNET refers to the inclusion of dense networks and IDC module and DD-UNET denotes dense networks and traditional dilation convolution added to the UNET structure.

Method	Total parameters	Trainable parameters	Non-trainable parameters	Train time epoch/(s)	Test time (s)
DD-UNET	56,223,034	56,190,272	32,762	145	8
DID-UNET	52,162,362	52,132,416	29,946	135	3

(A) (B) (C) (D) (E)
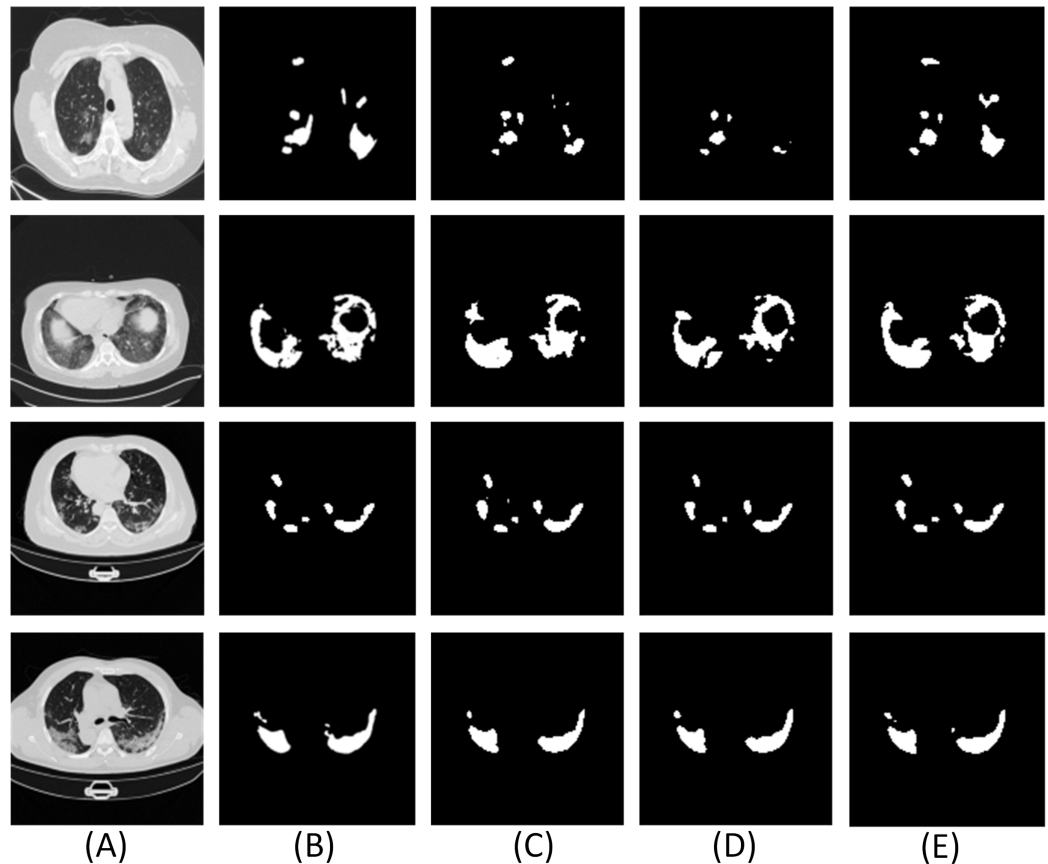
Figure 9 Experimental results of the improved dilation convolution and the traditional dilation convolution. Here (A) is the CT scan, (B) is the ground truth, (C–E) are the segment results from UNET, DD UNET and DID UNET, respectively. Here DID-UNET refers to the inclusion of dense networks and IDC module. DD-UNET denotes dense networks with traditional dilation convolution added to the UNET. Full-size ◨ DOI: 10.7717/peerj-cs.349/fig-9

Table 4 The results of comparison indees of the improved dilation convolution experiment and the traditional dilation convolution experiment.

Method	ACC	DC	S_{en}	S_p	P_c	AUC	F1	S_m	E_α	MAE
UNET	0.9696	0.7998	0.8052	0.9957	0.8247	0.9347	0.8154	0.8400	0.9390	0.0088
DD-UNET	0.9697	0.7757	0.7402	0.9971	0.8622	0.9214	0.7923	0.8401	0.9312	0.0094
DID-UNET	0.9700	0.8023	0.7987	0.9964	0.8425	0.9549	0.8241	0.8447	0.9374	0.0084
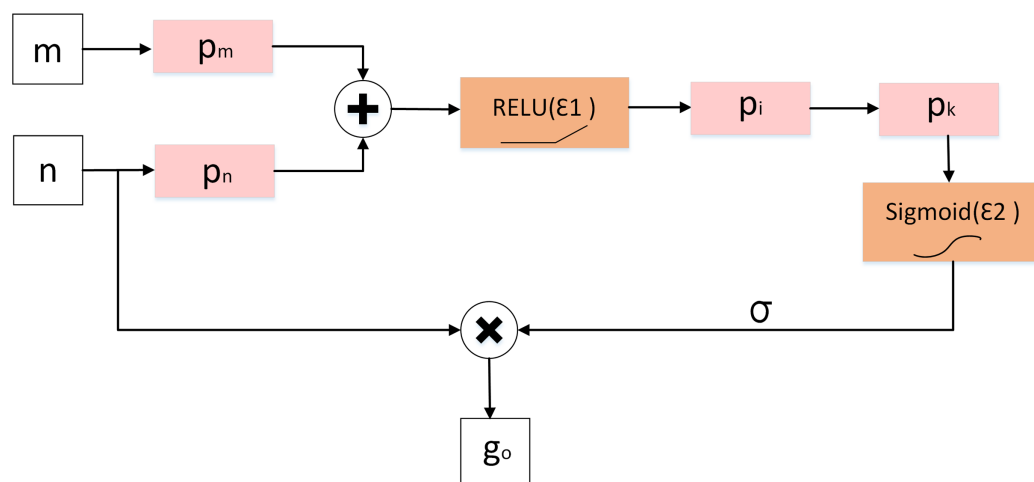
false prediction of small targets, and effectively improves the sensitivity and accuracy of the model.

EXPERIMENT RESULTS

COVID-19 segmentation dataset collection and processing

Organizing a COVID-19 segmentation dataset is time-consuming and hence there are not many CT scan segmentation datasets. At present, there was only one standard dataset namely the COVID-19 segmentation dataset (*MedSeg, 2020*), which was composed of 100 axial CT scans from different COVID-19 patients. All CT scans were segmented by radiologists associated with the Italian Association of medicine and interventional radiology. Since the database was updated regularly, on April 13, 2020, another segmented CT scans dataset with segment labels from Radiopaedia was added. The whole datasets that contained both positive and negative slices (373 out of the total of 829 slices have been evaluated by a radiologist as positive and segmented), were selected for training and testing the proposed model.

The dataset consists of 1,838 images with annotated ground truth was randomly divided into 1,318 training samples, 320 validation samples, and 200 test samples. Since the number of training images is less, we expand the training dataset where we first merge the COVID-19 lung CT scans with the ground scene and then perform six affine transformations as mentioned in *Krizhevsky, Sutskever & Hinton (2012)*. Later the transformed image is separated from the new background truth value and added to the training dataset as additional training images. Therefore, the 1,318 images of the training dataset are expanded, and 9,226 images are obtained for training. Figure 12 illustrates the data expansion process.

Segmentation evaluation index

The commonly used evaluation indicators for segmentation such as accuracy (ACC), precision (P_c), Dice coefficient (DC), the area under the curve (AUC), sensitivity (S_{en}),
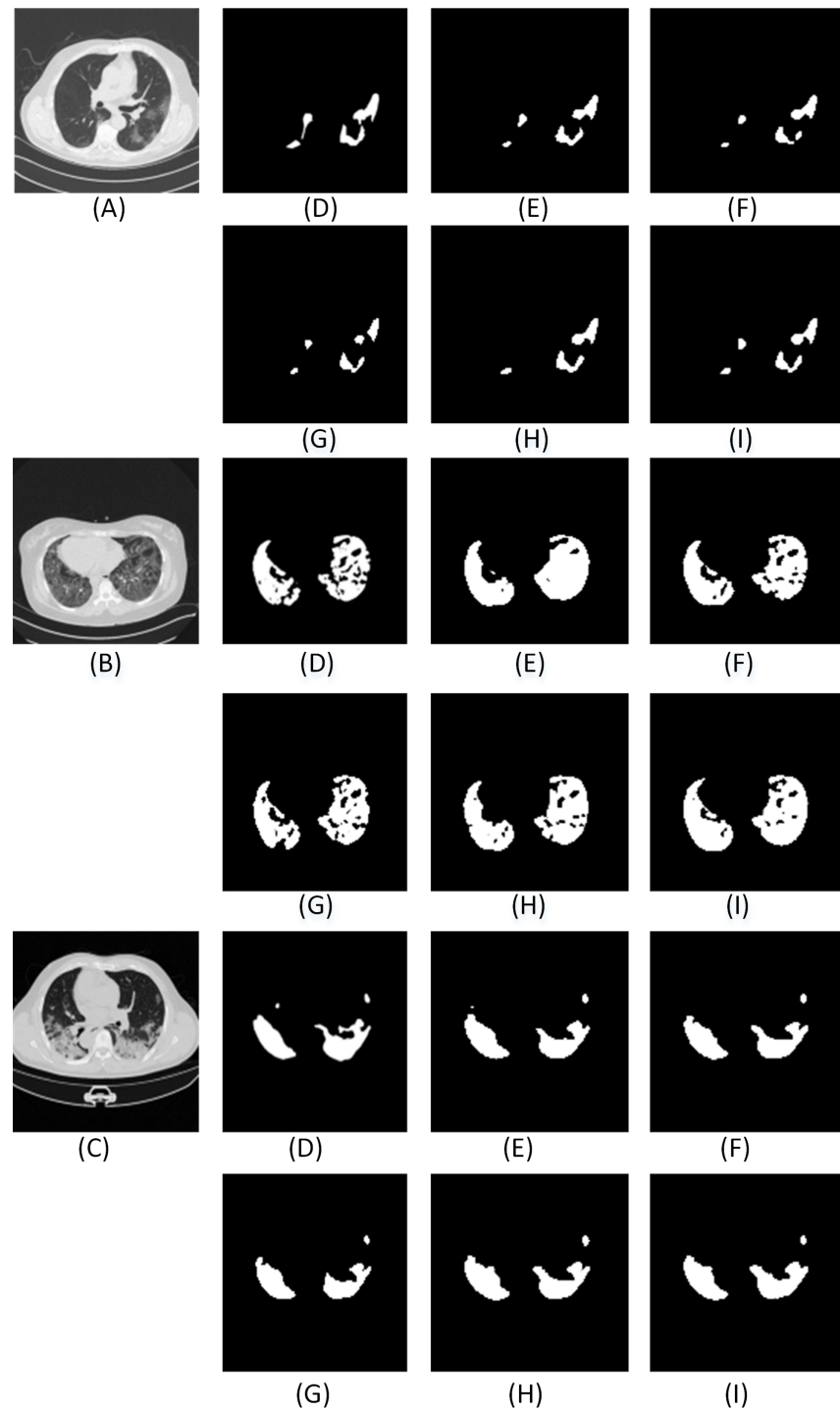
Figure 11 The qualitative results of comparative experiments with or without Attention Gate in the network. Here (A–C) are three test images. For each row, (D) denotes the ground truth, (E–I) illustrate the segmentation results from UNET, AG-UNET, DA-UNET, IDA-UNET and ADID-UNET, respectively. AG-UNET– the addition of AG module to UNET, DA-UNET–adding two dense networks and AG module to the network without including the IDC module. IDA-UNET refers to adding IDC and AG modules to the UNET without adding dense networks, and ADID-UNET indicates that dense networks, IDC and AG module are added to the network. Full-size ⬜ DOI: 10.7717/peerj-cs.349/fig-11

Table 5 The quantitative results of the comparison with or without the AG model experiment. AG-UNET– the addition of AG module to UNET, DA-UNET–adding two dense networks and AG module to the network without including the IDC module. IDA-UNET refers to adding IDC and AG modules to the UNET without adding dense networks, and ADID-UNET indicates that dense network, IDC and AG module are added to the network.

Method	ACC	DC	S_{en}	S_p	P_c	AUC	F1	S_m	E_α	MAE
UNET	0.9696	0.7998	0.8052	0.9957	0.8247	0.9347	0.8154	0.8400	0.9390	0.0088
AG-UNET	0.9697	0.8020	0.8106	0.9962	0.8347	0.9571	0.8116	0.8511	0.9345	0.0087
DA-UNET	0.9698	0.7754	0.7400	0.9959	0.8470	0.9274	0.7930	0.8334	0.9104	0.0091
IDA-UNET	0.9698	0.7961	0.7834	0.9964	0.8469	0.9450	0.8126	0.8513	0.9437	0.0085
ADID-UNET	0.9701	0.8031	0.7973	0.9966	0.8476	0.9551	0.8200	0.8509	0.9449	0.0082

specificity (S_p) and F1 score (F1) were used to evaluate the performance of the model. These performance indicators are calculated as follows:

(1) For computing accuracy, precision, sensitivity, specificity, and F1 score we generate the confusion matrix where the definitions of true positive (TP), true negative (TN), false positive (FP), and false negative (FN) are shown in Table 6.

(1) Accuracy (ACC): A ratio of the number of correctly predicted pixels to the total number of pixels in the image.

$$\text{Accuracy (ACC)} = \frac{\text{TP} + \text{TN}}{\text{TP} + \text{TN} + \text{FP} + \text{FN}} \tag{11}$$

(2) Precision (P_c): A ratio of the number of correctly predicted lesion pixels to the total number of predicted lesion pixels.

$$\text{Precision }(P_c) = \frac{\text{TP}}{\text{TP} + \text{FP}} \tag{12}$$

(3) Sensitivity (Sen): A ratio of the number of correctly predicted lesion pixels to the total number of actual lesion pixels.

$$\text{Sensitivity (Sen)} = \frac{\text{TP}}{\text{TP} + \text{FN}} \tag{13}$$

(4) F1 score (F1): A measure of balanced accuracy obtained from a combination of precision and sensitivity results.

$$\text{F1 score (F1)} = 2 \times \frac{P_c \times S_{en}}{P_c + S_{en}} \tag{14}$$

(5) Specificity (S_p): A ratio of the number of correctly predicted non-lesion pixels to the total number of actual non-lesion pixels.

$$\text{Specificity }(S_p) = \frac{\text{TN}}{\text{TN} + \text{FP}} \tag{15}$$

(6) Dice coefficient (DC): Represents the similarity between the model segment output (Y) and the ground truth (X). The higher the similarity between the lesion and the

ground truth, the larger the Dice coefficient and the better the segmentation effect. Dice coefficient is calculated as follows:

$$\text{Dice Coefficient (DC)} = \frac{2 \times (X \cap Y)}{X + Y} \qquad (16)$$

Also, we use a Dice coefficient (*Dice, 1945*) loss (dice_loss) as the training loss of the model, the calculation is as follows:

$$\text{Train Loss} = \text{Dice Coefficient Loss} = 1.0 - \frac{2 \times (X \cap Y)}{X + Y} \qquad (17)$$

(7) The area under the curve (AUC): AUC is the area under the receiver operating characteristic (ROC) curve. It represents the degree or the measure of separability and indicates the capability of the model in distinguishing the classes. Higher the AUC better is the segmentation output and hence the model.

In addition to the above widely used indicators, we also introduce the Structural metric (S_m) (*Fan et al., 2017*), Enhanced alignment metric (E_α) (*Fan et al., 2018*) and Mean Absolute Error (MAE) (*Fan et al., 2020*; *Elharrouss et al., 2020*) to measure the segmentation similarity with respect to the ground truth.

(8) Structural metric (S_m): Measures the structural similarity between the prediction map and ground truth segmented mask, it is more in line with the human visual system than Dice coefficient.

$$S_m = (1 - \beta) \times S_{os}\big(S_{op}, S_{gt}\big) + \beta \times S_{or}\big(S_{op}, S_{gt}\big) \qquad (18)$$

where S_{os} stands for target perception similarity, S_{or} stands for regional perceptual similarity, $\beta = 0.5$ is a balance factor between S_{os} and S_{or}. And S_{op} stands for the final prediction result and S_{gt} represents the ground truth.

(9) Enhance alignment metric (E_α): Evaluates the local and global similarity between two binary maps computed based on Eq. (19):

$$E_\alpha = \frac{1}{w \times h} \sum_{i}^{w} \sum_{j}^{h} \alpha \times \big(S_{op}(i,j), S_{gt}(i,j)\big) \qquad (19)$$

where w and h are the width and height of ground truth S_{gt}, (i,j) denotes the coordinates of each pixel in S_{gt}. α represents the enhanced alignment matrix:

$$\alpha = \frac{2 \times \big(S_{gt} - \sqrt{S_{gt}}\big) \times \big(S_{op} - \sqrt{S_{op}}\big)}{\big(S_{gt} - \sqrt{S_{gt}}\big)^2 + \big(S_{op} - \sqrt{S_{op}}\big)^2} \qquad (20)$$

(10) Mean Absolute Error (MAE): Measures the pixel-wise difference between S_{op} and S_{gt}, defined as:

$$\text{MAE} = \frac{1}{w \times h} \sum_{i}^{w} \sum_{j}^{h} \big|S_{op}(i,j) - S_{gt}(i,j)\big| \qquad (21)$$

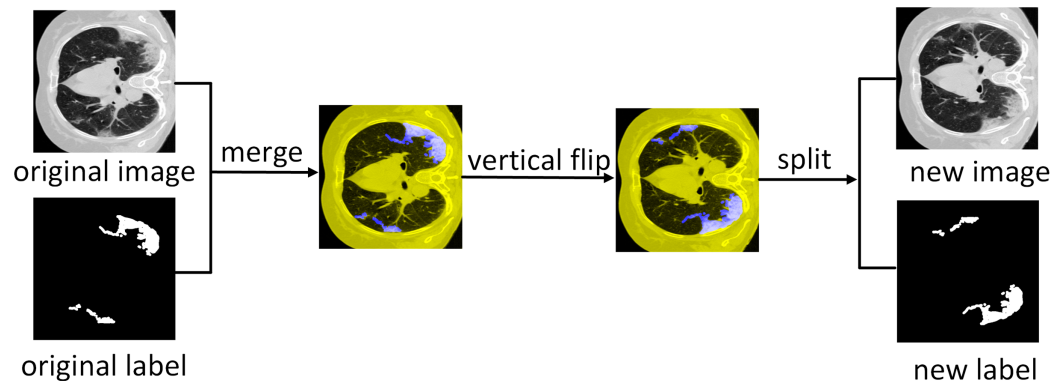**Joseph Raj et al. (2021),** *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.349

21/34

Figure 12 Data augmentation. Illustration of vertical flipping process showing the expansion of the training dataset. Full-size ☑ DOI: 10.7717/peerj-cs.349/fig-12

Table 6 Definition of TP, FP, FN, TN.

Category	Actual lesion	Actual non-lesion
Predicted Lesion	True Position (TP)	False Position (FP)
Predicted Non-Lesion	False Negative (FN)	True Negative (TN)

Experimental Details

The ADID-UNET proposed in this paper is implemented in Keras framework and is trained and tested by using the workstation with NVIDIA GPU P5000. During the training process, we set the learning rate as $l_r = 1 \times 10^{-3}$, and Adam optimizer was selected as the optimization technique. The 9,226 training samples, 320 verification samples, and 200 test samples were resized to 128×128 and trained with a batch size of 32 for 300 epochs. Figures 13 and 14 shows the performance curves obtained for the proposed ADID-UNET during training, validation, and testing.

Segmentation results and discussion

Qualitative results

To show the performance of the ADID-UNET model, we used 200 pairs of COVID-19 lung infection CT scans as test data, and the segmentation results are shown in Fig. 15. From the analysis of Fig. 15, it was found that the ADID-UNET model can accurately segment the COVID-19 lung infection areas from the CT scans, especially the smaller infected areas, and the segmentation result is very close to the ground truth. This illustrates the effectiveness of the proposed method for the segmentation of COVID-19 lung infection regions from CT scans. Moreover, we can also see that ADID-UNET can accurately segment the complicated infection areas (single COVID-19 lung infection areas and more complex uneven distribution infection areas) in CT scans, which further proves the power of the model proposed in this paper. In a word, the ADID-UNET model proposed in this paper can effectively and accurately segment COVID-19 lung infection
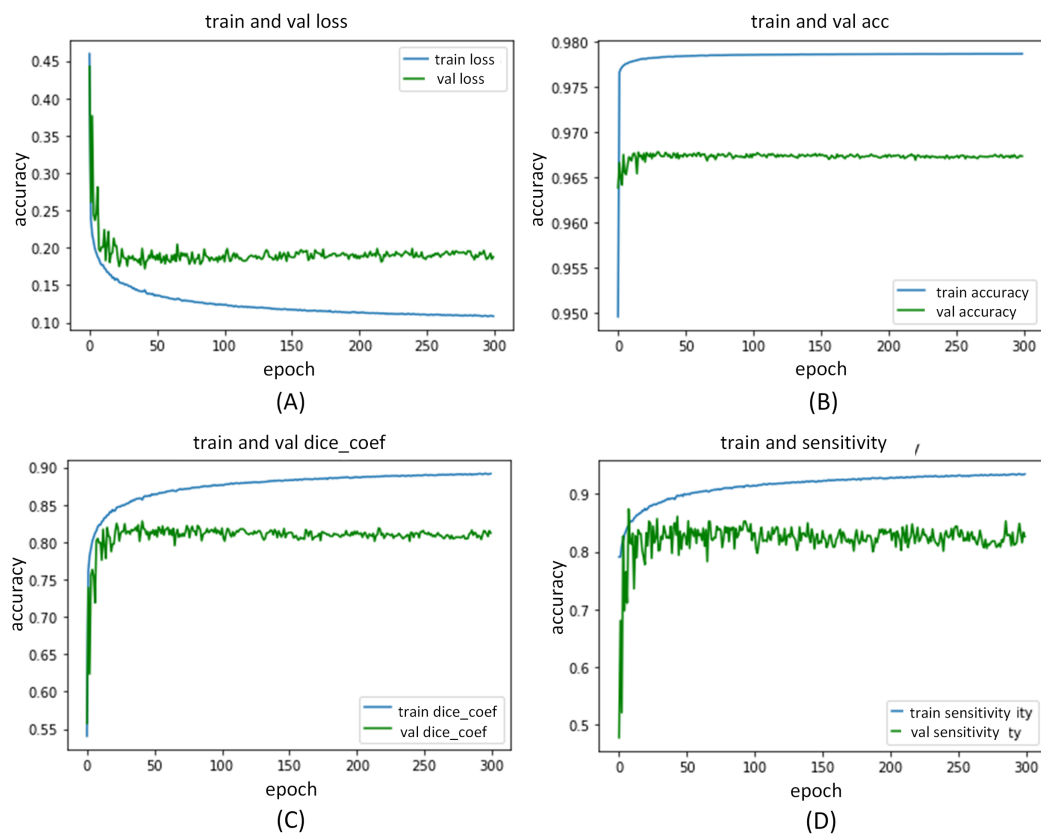
Figure 13 ADID-UNET training and validation performance index curve. (A–D) indicate the loss of training and validation, accuracy, Dice coefficient, and sensitivity performance curves, respectively.

Full-size ⊡ DOI: 10.7717/peerj-cs.349/fig-13

areas with different sizes and uneven distribution, and the visual effect of segmentation is very close to the gold standard.

Further, we also compare the proposed model with other state-of-art segmentation models. From the results (Figs. A1 and A2 and Table 7), we can infer that the ADID-UNET model presents segmentation outputs closer to the ground truth. In contrast, the FCN8s network (*Long, Shelhamer & Darrell, 2015*) presents more under and over segmented regions. Further RAD-UNET (*Zhuang et al., 2019a*) presents comparable segmentation results but its effect is less pronounced for smaller segments. Analyzing the segmentation visual results from Figs. A1 and A2, we can clearly find that the ADID-UNET model proposed in this paper can accurately segment the COVID-19 lung infection regions than other state-of-the-art model with results close to the ground truth, which proves the efficacy of the proposed ADID-UNET model.

Quantitative results

Table 7, presents the performance scores for various indicators mentioned in "Experiment Results". Here, for ADID-UNET the scores such as the Dice coefficient, precision, F1 score, specificity and AUC are 80.31%, 84.76%, 82.00%, 99.66% and 95.51%, respectively. Further, most of the performance indexes are above 0.8 with the highest
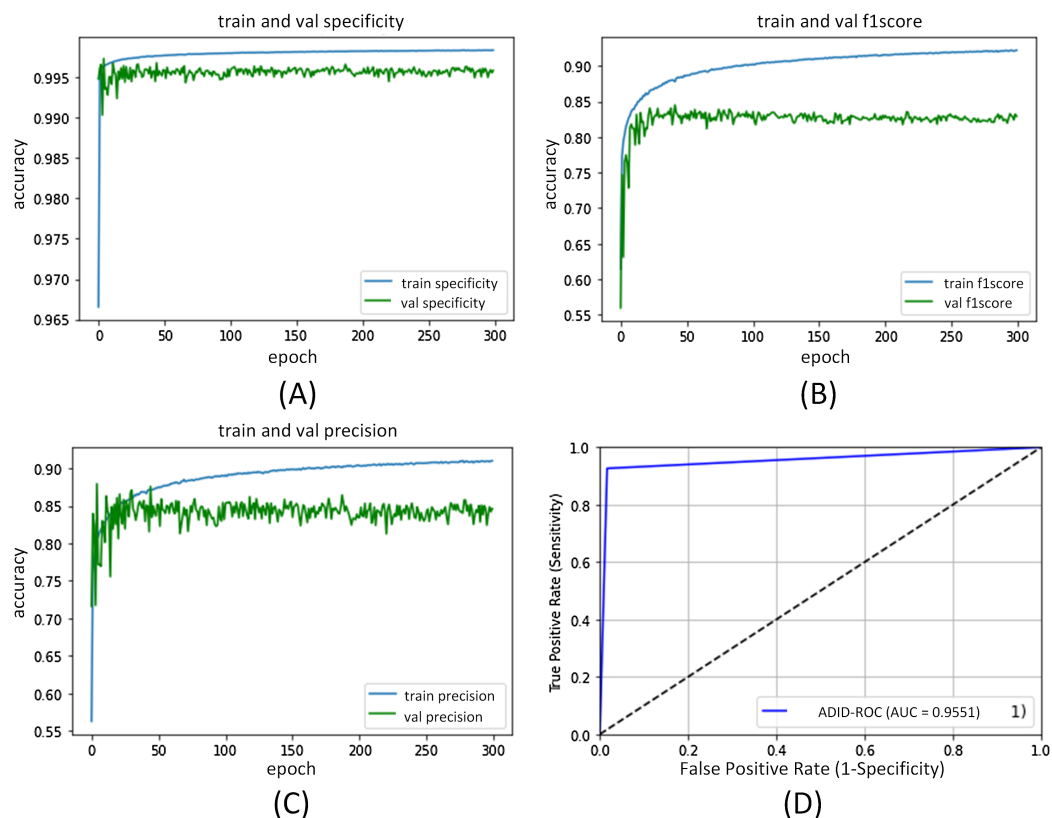
Figure 14 ADID-UNET training and validation performance index curve. (A–D) indicate the training and validation of the specificity, F1 score, precision, and AUC performance curves, respectively.
Full-size DOI: 10.7717/peerj-cs.349/fig-14

segmentation accuracy of 97.01%. The above results clearly indicates that the proposed model presents segmentation outputs closer to ground truth annotations.

Discussion

The proposed model presents an improved version of the UNET model obtained by the inclusion of modules such as the dense network, IDC and the attention gates to the existing UNET (*Ronneberger, Fischer & Brox, 2015*) structure. The effectiveness of these additions were experimentally verified in "Methods". Further, to summarize the effectiveness of the addition of each module to the UNET architecture, Table 8 tabulates the improvement at each stage of the addition. From Table 8, it is found that adding additional components to the UNET (*Ronneberger, Fischer & Brox, 2015*) structure can obviously improve the overall segmentation accuracy of the network. For example, with the inclusion of the dense networks (D-UNET), the metrics such as Dice coefficien (DC) and AUC reached 79.98% and 93.47%, respectively.

Further, the inclusion of the IDC improved the scores further (DID-UNET). Finally, the proposed model with dense network, IDC and the AG modules (namely ADID-UNET) presented the best performance scores and provided an improvement of 0.05%, 0.33%,
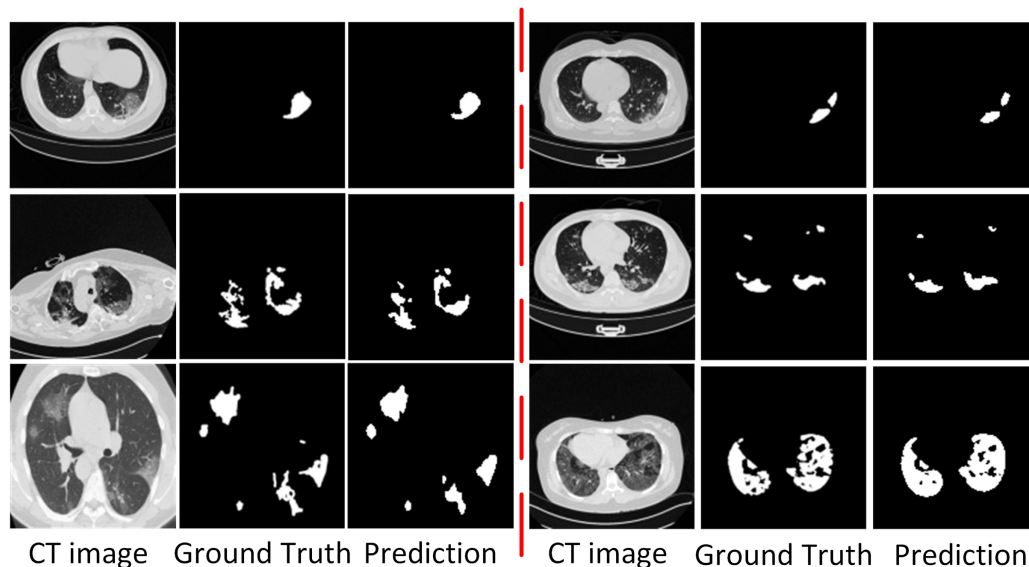
Figure 15 Visual comparison of the segmentation results of COVID-19 lung infection obtained from the proposed ADID-UNET. Column 1 represents the original CT Image, column 2 the ground truth and column 3 the predicted segmentation results from the proposed network called the Attention Gate-Dense Network-Improved Dilation Convolution-UNET (ADID-UNET).

Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-15

Table 7 Quantitative results of infected areas in the COVID-19 dataset. - - - means no relevant data in the original literature.										
Method	ACC	DC	S_{en}	S_p	P_c	AUC	F1	S_m	E_α	MAE
FCN8s	0.9666	0.6697	0.6692	0.9923	0.6860	0.9485	0.6724	0.7539	0.9134	0.0157
UNET	0.9696	0.7998	0.8052	0.9957	0.8247	0.9347	0.8154	0.8400	0.9390	0.0088
Segnet	0.9684	0.7408	0.7608	0.9937	0.7549	0.9492	0.7558	0.8080	0.9374	0.0125
Squeeze UNET	0.9689	0.7681	0.7827	0.9946	0.7776	0.9446	0.7785	0.8227	0.9326	0.0107
Residual UNET	0.9697	0.7924	0.7905	0.9961	0.8248	0.9444	0.8055	0.8397	0.9324	0.0094
RAD UNET	0.9699	0.7895	0.7625	0.9970	0.8601	0.9419	0.8062	0.8475	0.9328	0.0096
Fan et al. (2020)	- - -	0.7390	0.7250	0.9600	- - -	- - -	- - -	0.8000	0.8940	0.0640
Elharrouss et al. (2020)	- - -	0.7860	0.7110	0.9930	0.8560	- - -	0.7940	- - -	- - -	0.0760
Yan et al. (2020)	- - -	0.7260	0.7510	- - -	0.7260	- - -	- - -	- - -	- - -	- - -
Zhou, Canu & Ruan (2020)	- - -	0.6910	0.8110	0.9720	- - -	- - -	- - -	- - -	- - -	- - -
Chen, Yao & Zhang (2020)	0.8900	- - -	- - -	0.9930	0.9500	- - -	- - -	- - -	- - -	- - -
ADID-UNET	0.9701	0.8031	0.7973	0.9966	0.8476	0.9551	0.8200	0.8509	0.9449	0.0082

2.29%, 2.04% and 1.09% for metrics such as accuracy, DC, precision, AUC and structural metric respectively when compared to traditional UNET architecture.

Furthermore, from Figs. A1 and A2, it is obvious that ADID-UNET performs better than other well-known segmentation models in terms of visualization. Specifically, ADID-UNET can segment relatively smaller infected regions which is of great significance for clinical accurate diagnosis of COVID-19 infection location. The use of (a) dense

Table 8 The quantitative results showing percentages improvements of the model after adding additional components to UNET (*Ronneberger, Fischer & Brox, 2015*) structure. D-UNET denotes dense networks with UNET structure, DID-UNET represents dense networks and improved dilation convolution to the structure of UNET, and ADID-UNET refers to proposed model with dense networks improved dilation convolution and attention gate modules to the UNET structure. ↑ indicates that the performance index is higher than that of UNET structure, ↓ indicates that the performance index is lower than that of UNET structure.

Method	ACC	DC	S_{en}	S_p	P_c	AUC	F1	S_m	E_α	MAE
UNET	0.9696	0.7998	0.8052	0.9957	0.8247	0.9347	0.8154	0.8400	0.9390	0.0088
D-UNET	0.9700	0.8011	0.8096	0.9966	0.8596	0.9492	0.8184	0.8528	0.9394	0.0083
DID-UNET	0.9700	0.8023	0.7987	0.9964	0.8425	0.9549	0.8241	0.8447	0.9374	0.0084
ADID-UNET	0.9701	0.8031	0.7973	0.9966	0.8476	0.9551	0.8200	0.8509	0.9449	0.0082
Improvement of D-UNET	↑0.04%	↑0.13%	↑0.44%	↑0.09%	↑3.49%	↑1.45%	↑0.30%	↑1.28%	↑0.04%	↓0.05%
Improvement of DID-UNET	↑0.04%	↑0.25%	↓0.65%	↑0.07%	↑1.78%	↑2.02%	↑0.87%	↑0.47%	↓0.16%	↓0.04%
Improvement of ADID-UNET	↑0.05%	↑0.33%	↓0.79%	↑0.09%	↑2.29%	↑2.04%	↑0.46%	↑1.09%	↑0.59%	↓0.06%

networks instead of traditional convolution and max-pooling function, (b) inclusion of improved dilation convolution module between the encoder-decoder pipeline and (c) the presence of attention gate network in the skip connections have presented accurate segmentation outputs for various types of COVID-19 infections (GGO and pulmonary consolidation). However, ADID-UNET still has room for improvement in terms of Dice coefficient and sensitivity and also computational costs which can be researched in future.

CONCLUSION

The paper proposes a new variant of UNET (*Ronneberger, Fischer & Brox, 2015*) architecture to accurately segment the COVID-19 lung infections in CT scans. The model, ADID-UNET includes dense networks, improved dilation convolution, and attention gate, which has strong feature extraction and segment capabilities. The experimental results show that ADID-UNET is effective in segmenting small infection regions, with performance metrics such as accuracy, precision and F1 score of 97.01%, 84.76%, and 82.00%, respectively. The segmentation results of the ADID-UNET network can aid the clinicians in faster screening, quantification of the lesion areas and provide an overall improvement in the diagnosis of COVID-19 lung infection.

APPENDIX

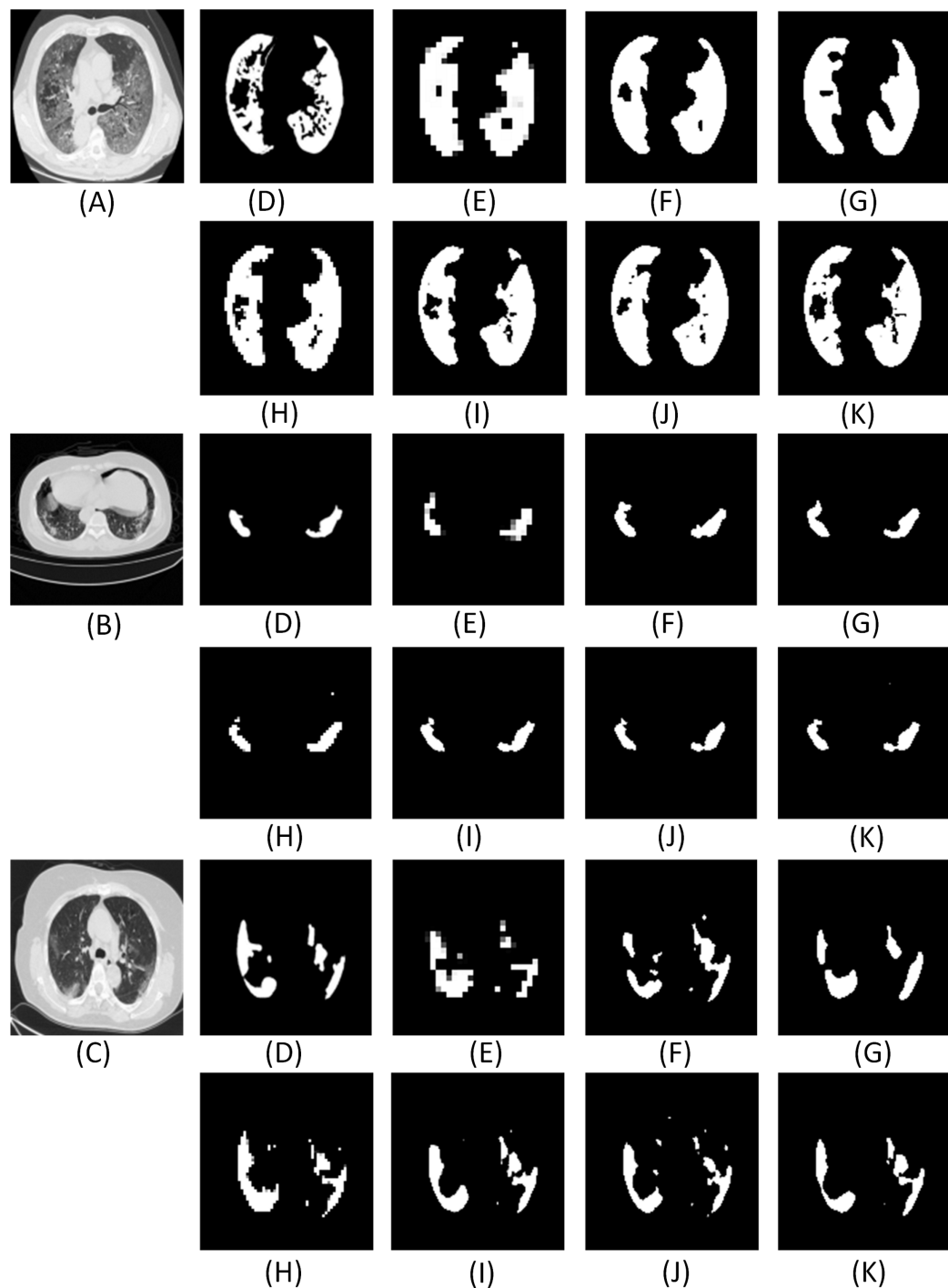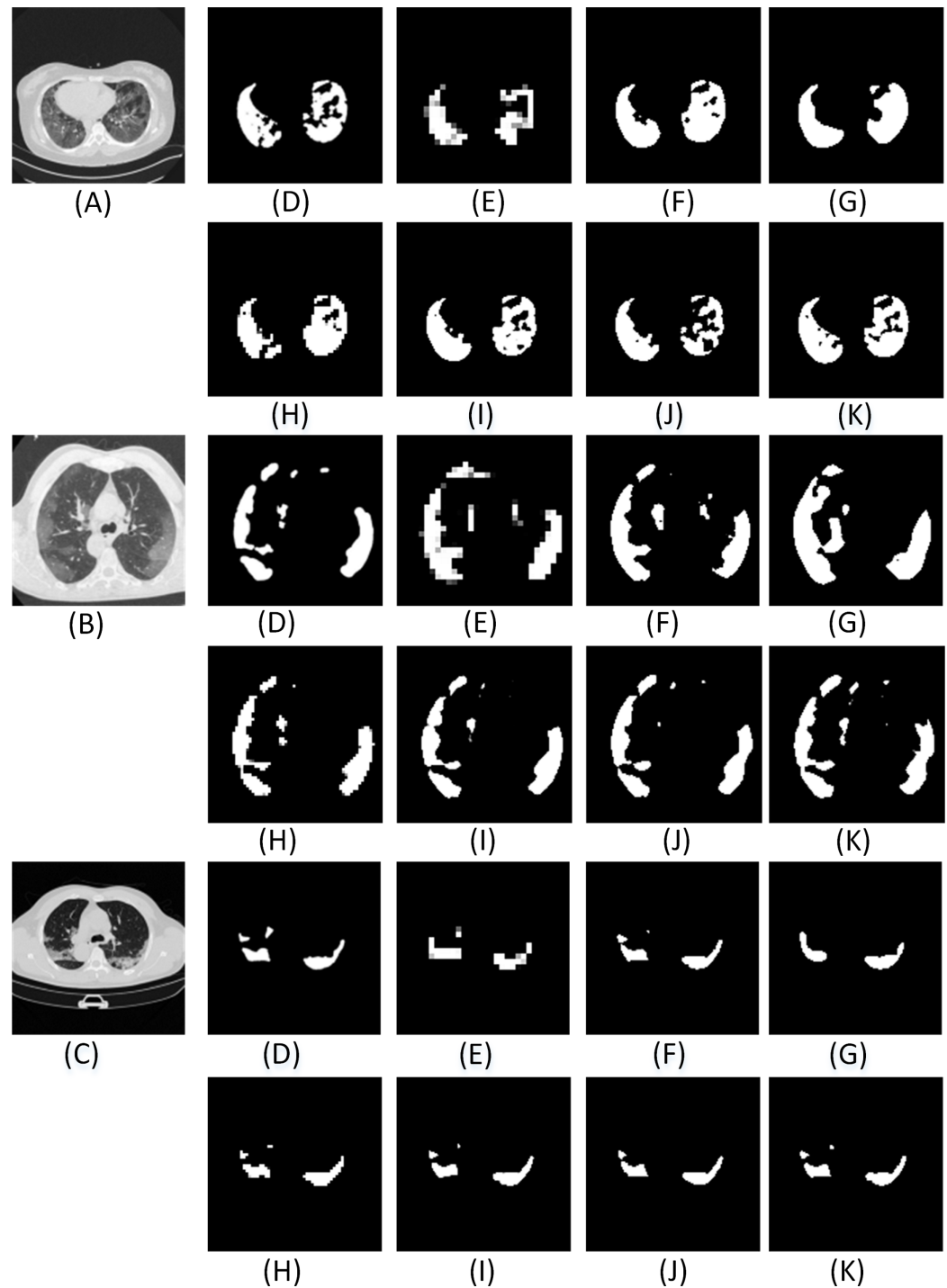We describe the abbreviations of this paper in detail, as shown in Table A1.

Figure A1 The visual comparison of the segmentation results of COVID-19 lung infection compared with other advanced models. Here (A–C) illustrate three test images obtained from (*MedSeg, 2020*). For each row, (D) denotes the ground truth, and (E–K) illustrate the segmentation results from FCN8s (*Long, Shelhamer & Darrell, 2015*), UNET (*Ronneberger, Fischer & Brox, 2015*), Segnet (*Badrinarayanan, Kendall & Cipolla, 2017*), Squeeze UNET (*Iandola et al., 2016*), Residual UNET (*Alom et al., 2018*), RAD UNET (*Zhuang et al., 2019a*), ADID-UNET, respectively.

Full-size 🖼 DOI: 10.7717/peerj-cs.349/fig-A1

Joseph Raj et al. (2021), *PeerJ Comput. Sci.*, DOI 10.7717/peerj-cs.349

27/34

Figure A2 **The visual comparison of the segmentation results of COVID-19 lung infection compared with other advanced models.** Here (A–C) illustrate three test images obtained from (*MedSeg, 2020*). For each row, (D) denotes the ground truth, and (E–K) illustrate the segmentation results from FCN8s (*Long, Shelhamer & Darrell, 2015*), UNET (*Ronneberger, Fischer & Brox, 2015*), Segnet (*Badrinarayanan, Kendall & Cipolla, 2017*), Squeeze UNET (*Iandola et al., 2016*), Residual UNET (*Alom et al., 2018*), RAD UNET (*Zhuang et al., 2019a*), ADID-UNET, respectively.

Full-size ⬚ DOI: 10.7717/peerj-cs.349/fig-A2

Table A1 An explanation of the acronyms that appears in this article.

Abbreviation	Explanation
D-UNET	Inclusion of Dense networks to the UNET structure
AG-UNET	Inclusion of Attention gate module to the UNET structure
DA-UNET	Inclusion of both dense networks and attention gate module to the UNET structure
IDA-UNET	Inclusion of Improved dilation convolution and Attention Gate module to the UNET structure
DID-UNET	Inclusion of dense networks and improved dilation convolution to the UNET structure
ADID-UNET	Inclusion of dense networks, Improved dilation convolution and Attention Gate modules to the UNET structure

Competing Interests

The authors declare that they have no competing interests.

Author Contributions

- Alex Noel Joseph Raj conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Haipeng Zhu conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

- Asiya Khan analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Zhemin Zhuang conceived and designed the experiments, analyzed the data, prepared figures and/or tables, and approved the final draft.
- Zengbiao Yang conceived and designed the experiments, performed the experiments, analyzed the data, performed the computation work, prepared figures and/or tables, and approved the final draft.
- Vijayalakshmi G. V. Mahesh performed the experiments, analyzed the data, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.
- Ganesan Karthik analyzed the data, performed the computation work, prepared figures and/or tables, authored or reviewed drafts of the paper, and approved the final draft.

Data Availability

The following information was supplied regarding data availability:

Code and data is available in the Supplemental Files.

The trained model and usage procedure are available at GitHub:

https://github.com/jalexnoel/ADID-UNET.git.

The dataset is available at MedSeg through Figshare:

MedSeg; Jenssen, Håvard Bjørke; Sakinis, Tomas (2021): MedSeg Covid Dataset 1. figshare. Dataset. DOI 10.6084/m9.figshare.13521488.v2.

MedSeg; Jenssen, Håvard Bjørke; Sakinis, Tomas (2021): MedSeg Covid Dataset 2. figshare. Dataset. DOI 10.6084/m9.figshare.13521509.v2.

Supplemental Information

Supplemental information for this article can be found online at http://dx.doi.org/10.7717/peerj-cs.349#supplemental-information.

REFERENCES

Agarwal V, Tang Y, Xiao J, Summers RM. 2020. Weakly-supervised lesion segmentation on CT scans using co-segmentation. In: *Medical Imaging 2020: Computer-Aided Diagnosis.* 356–361.

Ai T, Yang Z, Hou H, Zhan C, Xia L. 2020. Correlation of chest CT and RT-PCR testing in coronavirus disease 2019 (COVID-19) in China: a report of 1014 cases. *Radiology* **296(2)**:200642 DOI 10.1148/radiol.2020200642.

Almajalid R, Shan J, Du Y, Zhang M. 2019. Development of a deep-learning-based method for breast ultrasound image segmentation. In: *2018 17th IEEE International Conference on Machine Learning and Applications (ICMLA).* Piscataway: IEEE.

Alom MZ, Hasan M, Yakopcic C, Taha TM, Asari VK. 2018. Recurrent residual convolutional neural network based on UNET (R2U-Net) for medical image segmentation. *Available at http://arxiv.org/abs/1802.06955.*

Anthimopoulos M, Christodoulidis S, Ebner L, Christe A, Mougiakakou S. 2016. Lung pattern classification for interstitial lung diseases using a deep convolutional neural network. *IEEE Transactions on Medical Imaging* **35(5)**:1207–1216 DOI 10.1109/TMI.2016.2535865.

Badrinarayanan V, Kendall A, Cipolla R. 2017. egNet: a deep convolutional encoder-decoder architecture for image segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39(12)**:2481–2495.

Brostow GJ, Fauqueur J, Cipolla R. 2009. Semantic object classes in video: a high-definition ground truth database. *Pattern Recognition Letters* **30(2)**:88–97 DOI 10.1016/j.patrec.2008.04.005.

Cardona A, Saalfeld S, Preibisch S, Schmid B, Cheng A, Pulokas J, Tomancak P, Hartenstein V, Harris KM. 2010. An integrated micro- and macroarchitectural analysis of the drosophila brain by computer-assisted serial section electron microscopy. *PLOS Biology* **8(10)**:e1000502 DOI 10.1371/journal.pbio.1000502.

Chaganti S, Balachandran A, Chabin G, Cohen S, Flohr T, Georgescu B, Grenier P, Grbic S, Liu S, Mellot F, Murray N, Nicolaou S, Parker W, Re T, Sanelli P, Sauter AW, Xu Z, Yoo Y, Ziebandt V, Comaniciu D. 2020. Quantification of tomographic patterns associated with COVID-19 from chest CT. ArXiv preprint arXiv:2004.01289v5.

Chen X, Yao L, Zhang Y. 2020. Residual attention UNET for automated multi-class segmentation of COVID-19 chest CT images. ArXiv preprint arXiv:2004.05645.

Chen Y, Guo Q, Liang X, Wang J, Qian Y. 2019. Environmental sound classification with dilated convolutions. *Applied Acoustics* **148**:123–132 DOI 10.1016/j.apacoust.2018.12.019.

Chung M, Bernheim A, Mei X, Zhang N, Huang M, Zeng X, Cui J, Xu W, Yang Y, Fayad ZA, Jacobi A, Li K, Li S, Shan H. 2020. CT imaging features of 2019 novel coronavirus (2019-nCoV). *Radiology* **295(1)**:202–207 DOI 10.1148/radiol.2020200230.

Depeursinge A, Chin AS, Leung AN, Terrone D, Bristow M, Rosen G, Rubin DL. 2015. Automated classification of usual interstitial pneumonia using regional volumetric texture analysis in high-resolution computed tomography. *Investigative Radiology* **50(4)**:261–267 DOI 10.1097/RLI.0000000000000127.

Dice LR. 1945. Measures of the amount of ecologic association between species. *Ecology* **26(3)**:297–302 DOI 10.2307/1932409.

Elharrouss O, Almaadeed N, Subramanian N, Al-Maadeed S. 2020. An encoder-decoder-based method for COVID-19 lung infection segmentation. ArXiv preprint arXiv:2007.00861.

Fan D-P, Cheng M-M, Liu Y, Li T, Borji A. 2017. Structure-measure: a new way to evaluate foreground maps. In: *Proceedings of the IEEE International Conference on Computer Vision.* Piscataway: IEEE, 4548–4557.

Fan D-P, Gong C, Cao Y, Ren B, Cheng M-M, Borji A. 2018. Enhanced-alignment measure for binary foreground map evaluation. *Available at http://arxiv.org/abs/1805.10421.*

Fan DP, Zhou T, Ji GP, Zhou Y, Shao L. 2020. Inf-Net: Automatic COVID-19 lung infection segmentation from CT scans. *IEEE Transactions on Medical Imaging* **39(8)**:2626–2637.

Fang Y, Zhang H, Xie J, Lin M, Ying L, Pang P, Ji W. 2020. Sensitivity of chest CT for COVID-19: comparison to RT-PCR. *Radiology* **296(2)**:200432 DOI 10.1148/radiol.2020200432.

Ghosh A, Kumar H, Sastry PS. 2017. Robust loss functions under label noise for deep neural networks. ArXiv preprint arXiv:1712.09482.

Gordaliza PM, Muñoz-Barrutia A, Abella M, Desco M, Sharpe S, Vaquero JJ. 2018. Unsupervised CT lung image segmentation of a mycobacterium tuberculosis infection model. *Scientific Reports* **8(1)**:9802 DOI 10.1038/s41598-018-28100-x.

Guofeng T, Yong L, Huairong C, Qingchun Z, Huiying J. 2018. Improved UNET network for pulmonary nodules segmentation. *Optik - International Journal for Light and Electron Optics* **174**:460–469 DOI 10.1016/j.ijleo.2018.08.086.

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*, *Las Vegas, NV, USA*. Piscataway: IEEE.

Huang G, Liu Z, Maaten LVD, Weinberger KQ. 2017. Densely connected convolutional networks. In: *2017 IEEE Conference on Computer Vision and Pattern Recognition (CVPR)*, *Honolulu, HI, USA*. Piscataway: IEEE.

Huang G, Sun Y, Liu Z, Sedra D, Weinberger K. 2016. Deep networks with stochastic depth. In: Leibe B, Matas J, Sebe N, Welling M, eds. *Computer Vision – ECCV 2016. ECCV 2016: Lecture Notes in Computer Science*. Vol. 9908. Cham: Springer.

Iandola FN, Han S, Moskewicz MW, Ashraf K, Dally WJ, Keutzer K. 2016. Squeezenet: AlexNet-level accuracy with 50x fewer parameters and <0.5 mb model size. ArXiv preprint arXiv:1602.07360.

Jue J, Yu-Chi H, Chia-Ju L, Darragh H, Hellmann MD, Deasy JO, Gig M, Harini V. 2018. Multiple resolution residually connected feature streams for automatic lung tumor segmentation from CT images. *IEEE Transactions on Medical Imaging* **38(1)**:134–144.

Karen S, Andrew Z. 2014. Very deep convolutional networks for large-scale image recognition. ArXiv preprint arXiv:1409.1556.

Kermany DS, Goldbaum M, Cai W, Valentim CCS, Liang H, Baxter SL, McKeown A, Yang G, Wu X, Yan F, Dong J, Prasadha MK, Pei J, Ting MYL, Zhu J, Li C, Hewett S, Dong J, Ziyar I, Shi A, Zhang R, Zheng L, Hou R, Shi W, Fu X, Duan Y, Huu VAN, Wen C, Zhang ED, Zhang CL, Li O, Wang X, Singer MA, Sun X, Xu J, Tafreshi A, Lewis MA, Xia H, Zhang K. 2018. Identifying medical diagnoses and treatable diseases by image-based deep learning. *Cell* **172(5)**:1122–1131 DOI 10.1016/j.cell.2018.02.010.

Kong B, Wang X, Bai J, Lu Y, Yin Y. 2019. Learning tree-structured representation for 3D coronary artery segmentation. *Computerized Medical Imaging and Graphics* **80**:101688 DOI 10.1016/j.compmedimag.2019.101688.

Krizhevsky A, Sutskever I, Hinton G. 2012. Imagenet classification with deep convolutional neural networks. *Neural Information Processing Systems* **25(6)**:84–90 DOI 10.1145/3065386.

Li L, Qin L, Xu Z, Yin Y, Wang X, Kong B, Bai J, Lu Y, Fang Z, Song Q, Cao K, Liu D, Wang G, Xu Q, Fang X, Zhang S, Xia J, Xia J. 2020. Using artificial intelligence to detect COVID-19 and community-acquired pneumonia based on pulmonary CT: evaluation of the diagnostic accuracy. *Radiology* **296(2)**:200905 DOI 10.1148/radiol.2020200905.

Long J, Shelhamer E, Darrell T. 2015. Fully convolutional networks for semantic segmentation. *IEEE Transactions on Pattern Analysis and Machine Intelligence* **39(4)**:640–651.

Ma J, Ge C, Wang Y, An X, Gao J, Yu Z, Zhang M, Liu X, Deng X, Cao S, Wei H, Mei S, Yang X, Nie Z, Li C, Tian L, Zhu Y, Zhu Q, Dong G, He J. 2020. COVID-19 CT segmentation dataset. *Available at http://radiopaedia.org*.

MedSeg. 2020. Segmentation dataset, COVID-19 CT. *Available at https://medicalsegmentation.com/covid19/*.

Mukherjee P, Lall B, Lattupally S. 2018. Object co-segmentation using deep siamese network. ArXiv preprint arXiv:1803.02555.

Negi A, Raj ANJ, Nersisson R, Zhuang Z, Murugappan M. 2020. RDA-UNET-WGAN: an accurate breast ultrasound lesion segmentation using wasserstein generative adversarial networks. *Arabian Journal for Science and Engineering* **45**:6399–6410.

Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Mori K, Mcdonagh S, Hammerla NY, Kainz B. 2018. Attention UNET: Learning where to look for the pancreas. ArXiv preprint arXiv:1804.03999.

Peng T, Xu TC, Wang Y, Zhou H, Wang J. 2020. Hybrid automatic lung segmentation on chest ct scans. *IEEE Access* **8**:73293–73306 DOI 10.1109/ACCESS.2020.2987925.

Rajaraman S, Candemir S, Kim I, Thoma G, Antani S. 2018. Visualization and interpretation of convolutional neural network predictions in detecting pneumonia in pediatric chest radiographs. *Applied Sciences* **8(10)**:1715 DOI 10.3390/app8101715.

Ronneberger O, Fischer P, Brox T. 2015. UNET: convolutional networks for biomedical image segmentation. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention*.

Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, Rueckert D. 2019. Attention gated networks: learning to leverage salient regions in medical images. *Medical Image Analysis* **53**:197–207.

Shan F, Gao Y, Wang J, Shi W, Shi N, Han M, Xue Z, Shen D. 2020. Lung infection quantification of COVID-19 in CT images with deep learning. ArXiv preprint arXiv:200304655.

Shi F, Xia L, Shan F, Wu D, Wei Y, Yuan H, Jiang H, Gao Y, Sui H, Shen D. 2020. Large-scale screening of COVID-19 from community acquired pneumonia using infection size-aware classification. ArXiv preprint arXiv:2003.09860.

Shuai W, Bo K, Jinlu M, Xianjun Z, Mingming X, Jia G, Mengjiao C, Jingyi Y, Yaodong L, Xiangfei M, Bo X. 2020. A deep learning algorithm using CT images to screen for corona virus disease (COVID-19). *MedRxiv* DOI 10.1101/2020.02.14.20023028.

Tang Z, Zhao W, Xie X, Zhong Z, Shi F, Liu J, Shen D. 2020. Severity assessment of coronavirus disease 2019 (COVID-19) using quantitative features from chest ct images. ArXiv preprint arXiv:2003.11988.

Wang C, Horby PW, Hayden FG, Gao GF. 2020a. A novel coronavirus outbreak of global health concern. *The Lancet* **395(10223)**:470–473 DOI 10.1016/S0140-6736(20)30185-9.

Wang G, Liu X, Li C, Xu Z, Ruan J, Zhu H, Meng T, Li K, Huang N, Zhang S. 2020b. A noise-robust framework for automatic segmentation of COVID-19 pneumonia lesions from CT images. *IEEE Transactions on Medical Imaging* **39(8)**:2653–2663 DOI 10.1109/TMI.2020.3000314.

Wang L, Wong A. 2020. COVID-Net: a tailored deep convolutional neural network design for detection of COVID-19 cases from chest X-Ray images. ArXiv preprint arXiv:2003.09871.

Wang S, Zhou M, Liu Z, Liu Z, Gu D, Zang Y, Dong D, Gevaert O, Tian J. 2017. Central focused convolutional neural networks: developing a data-driven model for lung nodule segmentation. *Medical Image Analysis* **40**:172–183 DOI 10.1016/j.media.2017.06.014.

Xia X, Kulis B. 2017. W-Net: A deep model for fully unsupervised image segmentation. ArXiv preprint arXiv:1711.08506.

Xu X, Jiang X, Ma C, Du P, Li X, Lv S, Yú L, Chen Y, wei Su J, jing Lang G, Li Y-T, Zhao H, Xu K, Ruan L, Wu W. 2020. Deep learning system to screen coronavirus disease 2019 pneumonia. ArXiv preprint arXiv:2002.09334.

Xu Y, Wang Y, Yuan J, Cheng Q, Wang X, Carson PL. 2018. Medical breast ultrasound image segmentation by machine learning. *Ultrasonics* **91**:1–9 DOI 10.1016/j.ultras.2018.07.006.

Yan Q, Wang B, Gong D, Luo C, Zhao W, Shen J, Shi Q, Jin S, Zhang L, You Z. 2020. COVID-19 chest CT image segmentation—a deep convolutional neural network solution. *Available at http://arxiv.org/abs/2004.10987*.

Ye H, Gao F, Yin Y. 2019. Precise diagnosis of intracranial hemorrhage and subtypes using a three-dimensional joint convolutional and recurrent neural network. *European Radiology* **29(11)**:6191–6201 DOI 10.1007/s00330-019-06163-2.

Ye X, Beddoe G, Slabaugh G. 2009. Graph cut-based automatic segmentation of lung nodules using shape, intensity, and spatial features. In: *Workshop on Pulmonary Image Analysis, MICCAI.*

Ye X, Lin X, Dehmeshki J, Slabaugh G, Beddoe G. 2009. Shape-based computer-aided detection of lung nodules in thoracic CT images. *IEEE Transactions on Biomedical Engineering* **56(7)**:1810–1820 DOI 10.1109/TBME.2009.2017027.

Ye Z, Zhang Y, Wang Y, Huang Z, Song B. 2020. Chest CT manifestations of new coronavirus disease 2019 (COVID-19): a pictorial review. *European Radiology* **30(8)**:4381–4389 DOI 10.1007/s00330-020-06801-0.

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

Zhang Z, Sabuncu MR. 2018. Generalized cross entropy loss for training deep neural networks with noisy labels. In: *Proceedings of the 32nd International Conference on Neural Information Processing Systems.* 8792–8802.

Zhou T, Canu S, Ruan S. 2020. An automatic COVID-19 CT segmentation based on UNET with attention mechanism. ArXiv preprint arXiv:2004.06673.

Zhuang Z, Li N, Raj ANJ, Mahesh VGV, Qiu S. 2019a. An RDAU-NET model for lesion segmentation in breast ultrasound images. *PLOS ONE* **14(8)**:e0221535.

Zhuang Z, Noel A, Raj J, Jain A, Murugappan M. 2019b. Nipple segmentation and localization using modified UNET on breast ultrasound images. *Journal of Medical Imaging and Health Informatics* **9(9)**:1827–1837 DOI 10.1166/jmihi.2019.2828.

Shape and Texture Aware Facial Expression Recognition Using Spatial Pyramid Zernike Moments and Law's Textures Feature Set

VIJAYALAKSHMI G. V. MAHESH[ID]**1, CHENGJI CHEN**2,
VIJAYARAJAN RAJANGAM[ID]**3, (Member, IEEE), ALEX NOEL JOSEPH RAJ**[ID]**2, (Member, IEEE),
AND PALANI THANARAJ KRISHNAN**[ID]**4, (Member, IEEE)**

[1]Department of Electronics and Communication, BMS Institute of Technology and Management, Bengaluru 560064, India
[2]Department of Electronic Engineering, College of Engineering, Shantou University, Shantou 515063, China
[3]Centre for Healthcare Advancement, Innovation and Research, Vellore Institute of Technology, Chennai 600127, India
[4]Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai 600119, India

Corresponding author: Alex Noel Joseph Raj (jalexnoel@stu.edu.cn)

ABSTRACT Facial expression recognition (FER) requires better descriptors to represent the face patterns as the facial region changes due to the movement of the face muscles during an expression. In this paper, a method of concatenating spatial pyramid Zernike moments based shape features and Law's texture features is proposed to uniquely capture the macro and micro details of each facial expression. The proposed method employs multilayer perceptron and radial basis function feed forward artificial neural networks for recognizing the facial expressions. The suitability of the features in recognizing the expressions is explored across the datasets independent of the subjects or persons. The experiments conducted on JAFFE and KDEF datasets demonstrate that the concatenated feature vectors are capable of representing the facial expressions with better accuracy and least errors. The radial basis function based classifier delivers a performance with an average recognition accuracy of 95.86% and 88.87% on the JAFFE and KDEF datasets respectively for subject dependent FER.

INDEX TERMS Facial expressions, emotions, Zernike moments, Law's texture features, neural network classifier.

I. INTRODUCTION

A face is a unique trait for identifying or recognizing people around us. Facial expressions are the reflection of emotional states and play an important role in non-verbal communication. A facial expression is a result of the motions of the muscles underneath the skin of the face. Analysis of these expressions helps in understanding the behavior of a person and certain anatomical changes. For example, the heart rate is higher in anger as compared to happiness and the skin resistance decreases during sadness revealing high stress. Thus, automated FER and subsequent analysis have found its applications in various domains such as surveillance, crowd emotion monitoring, psychological disorder detection, human-computer interaction, driver safety assistance and so on. Ekman and Friesen [1] has formalized

six universal facial expressions such as surprise, fear, disgust, contempt, anger, sadness and happiness. These expressions have evolved through social learning and very much essential for survival.

FER is a pattern recognition system and the basis for FER is observation/identification, understanding and interpretation of the visual cues on the face. The main components of FER are (i) face detection (ii) feature extraction and (iii) expression classification. The spatial arrangement of the facial features like shape, fine lines, wrinkles on the facial skin and the structural elements of the face such as forehead, eyebrow, eyes, mouth create different patterns on the face during an expression. These patterns form the key characteristics comprising both micro and macro details. The micro details include variation in the wrinkles and fine lines changing the facial appearance and texture of the skin. Whereas, the movements in the structural elements of the face indicates the macro details. A raised and arched eyebrow

The associate editor coordinating the review of this manuscript and approving it for publication was Rosalia Maglietta[ID].

with eyes wide open represents surprise. A lowered eyebrow with intensely staring eyes indicates anger. Disgust expression is expressed with a wrinkled nose and lowered eyebrows. Similarly, an open mouth represents fear whereas corners raised mouth shows happiness. The extraction of these patterns forms the feature descriptors for recognizing facial expressions.

Feature extraction and description play a major role in deciding the accuracy of the FER system. The features should be redundant, reliable, robust and unique with the best discriminating ability. Extracting the right features is critical and also a challenging task. Most of the FER methods have been proposed for capturing the facial expressions through shape and texture features as descriptors that include Zernike moments (ZM), histogram of oriented gradients (HOG), active shape model, local binary pattern (LBP), local directional pattern (LDP), statistical measures and gray level co-occurrence matrix (GLCM) [2]. These methods extract either the macro or micro details alone to describe the facial expressions. But, it is important to provide a maximum possible description to enhance the performance of the system. Thus, both texture and shape information provided by the Laws texture energy measures and Zernike Moments are utilized for FER. Texture information is obtained by five types of kernels such as level, edge, spot, ripple and wave. Distinct information is obtained with each kernel revealing the changes in the wrinkles and fine lines of the facial skin. Zernike moments provide the shape information of the changes in the facial appearance due to the movements in the structural elements. In this perspective, ZM is extended to spatial pyramid representation with three-level decomposition to capture the shape information in each facial sub-region thus forming Spatial pyramid Zernike moments (SPZM). The Law's texture energy features (LTexM) and SPZM are combined to form the integrated feature vector that has improved discriminating ability to recognize and classify the facial expressions. The contributions of this paper are (i) an integrated feature set for capturing the macro and micro details from the facial expressions using SPZM and Law's texture features (ii) quantitative assessment of improved recognition accuracy by considering the images with different orientations under subject dependent and independent FER (iii) integration of features for the effective representation of the facial expressions (iv) robust features for recognizing the facial expressions from the full left and right profile of facial images.

The literature review is presented in section II. Section III describes the Law's texture energy measures and SPZM. The proposed methodology and the experimental results are presented in sections III and IV respectively. Conclusion is presented in section V

II. RELATED WORK
The different approaches for extracting the facial feature to recognize the expressions include (i) Facial action coding system (FACS) and Action unit (AU) (ii) geometric, appearance

and hybrid methods. In AU method, the movement of muscles responsible for producing a facial expression is encoded into 46 facial AUs [3]. FER system detects the face AUs to classify the expressions using observations and comparisons. FACS describes the relationship between the muscle movements of the face and expressions and was introduced by Ekman and Friesen [1] based on the characteristics of AUs. The second method depends on the extraction of content-based facial features. It depends upon the appearance, geometric and hybrid characteristics. Appearance-based approaches are holistic capturing global information from the facial images to generate the feature vector for distinguishing facial expressions. Non-holistic approaches make use of geometric features such as eyes, nose, mouth, chin, head outline of a face and their relationships.

Holistic approaches apply transformations and use statistical methods to extract the features representing the texture characteristics of the image. Gabor filter [4]–[8] provides texture descriptors with good discrimination ability. It provides both magnitude and phase components, but magnitudes are selected as features as they are invariant to translation. Gabor filter provides multidimensional feature vectors with high computation cost and the dimension of features can be reduced by principal component analysis (PCA). LBP [9], [10] generates binary patterns to represent the texture by comparing the center pixel with the neighborhood pixels of a region from the facial image. Variants of LBP were further introduced for improvement in the performance of the FER system. A weighted multi-scale method for LBP was proposed by [11], in which multiple weighted LBP features are extracted with different scales. The extended LBP is combined with the Karhunen-Loeve (KL) transform in [12]. The role of subspace analysis methods such as PCA and Independent Component Analysis (ICA) is investigated for the extraction of the facial expression features [13]. Other methods like local directional numbers (LDN) [14], local ternary pattern [15], discrete wavelet transform (DWT) [16] and sparse local Fisher discriminant analysis (SLFDA) [17] were also explored for FER in recent years.

Non-holistic approaches extract the geometrical features providing the position of facial landmarks and shape information. Geometric features were extracted using the curvelet transform. The coefficients of the transform with varying scales and angles form the feature vectors to represent the facial expressions [18]. Histogram of oriented gradients, originally proposed for object recognition, was found significant for FER. HOG provides the magnitude and phase of the gradients from which the dominant gradients are selected relating to the edge information [19], [20]. In [21], facial landmarks are tracked based on elastic bunch graph matching (EBGM) displacement. The facial landmarks or combination of the landmarks form the feature vector representing a facial expression. In [22], optical flow based facial points were tracked from consecutive frames to detect the movement of facial points to provide dynamic features. Moments based shape descriptors are critical in representing the facial

expressions. Zernike moments with magnitude features, exhibiting orthogonal and rotation invariant properties, are used in [2], [23], [24]. Pseudo Zernike moments (PZM) [25] also provide a good description for FER. Appearance-based methods are affected by lighting and orientation conditions. Geometry based methods provide better feature descriptors working well irrespective of the variations in the facial image.

Further, the recent years have also witnessed the success of deep learning methods with multilayered architectures in facial expression recognition. The deep learning methods automatically compute the features for data representation while reducing the requirement for extracting the hand crafted features. The work presented in [27] utilized the ZM for deriving the coefficients of the convolution kernels in convolution layer of CNN architecture. This was found to be significant in extracting the shape features and improved the classification accuracy. Deep sparse auto encoders were implemented by [28] to learn discriminative and robust features. The work presented in [29] introduced a part based hierarchical bi directional recurrent neural network to analyze the dynamic evolution and morphological variations in facial expressions which proved to be effective in reducing the error rates. Modifications in CNN architectures are introduced by [30]–[33] to enhance the performance of FER system. Generative adversial networks(GAN) [34], [35] with generator and discriminator networks have also emerged to be better models in discriminating the facial expressions. Though deep features are efficient and have outperformed the existing hand crafted feature methods, the deep learning methods require larger datasets for training and are computationally expensive.

Thus this paper proposes a method considering the advantages of appearance and geometric based methods. Hence, the features from holistic and non-holistic approaches are combined to form a robust feature vector for improved classification. The efficacy of the selected feature descriptors in the proposed method is proved by focusing on both subject dependent and subject independent FER.

III. SPZM AND LTexM INTEGRATED FEATURE SET MODEL FOR FER

The schematic of the proposed feature concatenation strategy for recognizing facial expressions is shown in Fig. 1. The FER is a pattern recognition problem where a facial pattern (FP) is assigned one of the m expression labels $\{c_i\}_{i=1,2,........m}$. The key characteristics or patterns are extracted from the facial images with expressions in the form of feature vectors. The selection of the right features improves the degree of accuracy in classification. LTex and SPZM are extracted from the facial images and integrated to form a feature vector for training a neural network classifier. Upon training, the classifier model is observed to classify the facial expression of an image. Finally, the performance of the model is quantitatively assessed using various performance measures based on the elements of the confusion matrix.

The process of the proposed approach is presented in Algorithm 1.

Algorithm 1 The process flow of SPZM and LTexM integrated feature set model

Begin

Input: Facial expression images database with seven emotions

Dataset division: Training and Testing dataset

Training phase:

1) Assign M=Number of training images
2) **for** i=1 to M
   a) Extract the Law's texture features
   b) Extract Zernike moment based shape features from spatial pyramid of the image SPZM = $\{|A_{nm}|_k^l\}$
   c) Integrate the texture and shape features to form integrated feature vector FV_{int} = {LTexM, SPZM}

   end
3) Assign the class labels to the integrated feature set with seven expressions C={angry, disgust, fear, happy, sad, surprise and neutral} and create label feature set matrix $\{FV_{int}, C\}$
4) Provide $\{FV_{int}, C\}$ to the neural network classifiers (MLPNN and RBFNN)
   a) Define the learning parameters of the neural network
   b) Tune the parameters to improve the performance and derive the classifier model

Testing phase (Output):

1) Assign N=Number of testing images
2) **for** i=1 to N
   a) Extract the Law's texture features
   b) Extract Zernike moment based shape features from spatial pyramid of the image SPZM = $\{|A_{nm}|_k^l\}$
   c) Integrate the texture and shape features to form integrated feature vector FV_{intq} = {LTexM, SPZM}$_Q$

   end
3) Provide FV_{intq} to the classifer model to predict the class label which belongs to set C

End

The subsections present the comprehension of the proposed work.

A. FACIAL EXPRESSION DATASET

The two databases (i) Japanese female facial expression (JAFFE) and (ii) Karolina Directed Emotional Faces (KDEF) are experimented to analyze the performance of the proposed FER

The JAFFE database [36] is widely used for evaluating the performance of FER systems. JAFFE is a dataset collected from 10 Japanese females. Each individual has posed with six basic facial emotions like angry, disgust, fear, happy, sad and
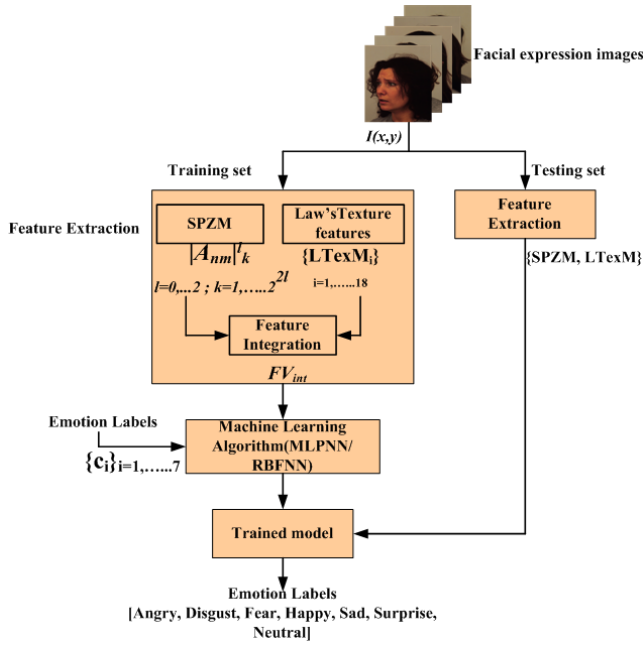
FIGURE 1. Proposed methodology for facial expression recognition.

FIGURE 2. Sample images from the JAFFE database.

FIGURE 3. Sample images (a) the emotions of an individual (KDEF database) and (b) five different profiles/orientations of an expression.

surprise with a neutral pose. This dataset has 213 grayscale images with labels of facial expression and each image has a size of 256×256 pixels. Fig. 2 shows the sample images of a Japanese female with facial expressions from the database.

The KDEF database [37] developed by the department of clinical neuroscience, a section of psychology from Sweden for psychological and medical research purposes. The database has 4900 color images of facial expressions acquired from 70 individuals (35 females and 35 males) of age ranging from 20 to 30 years. Here, each individual has displayed seven emotions such as angry, disgust, fear, happy, sad, surprise and neutral. The facial expressions of each person is captured under 2 sessions from 5 different angles: -90, -45, 0, 45 and 90 degrees, providing full left, half left, straight, half right and full right profiles under uniform lighting. The spatial resolution of the facial image is 562×762. Fig. 3 shows the sample images of an individual with facial expressions from the KDEF database.

For the proposed work, 700 facial expression images are acquired from 20 female subjects with all five orientations. All the images were converted to gray-scale for computational reasons. The problem of facial expression classification is more challenging with this dataset as some images are captured with different orientations and only partial face is visible in full left and full right profiles. Thus, the FER system with a robust feature extraction method is required to overcome the complexity of the problem.
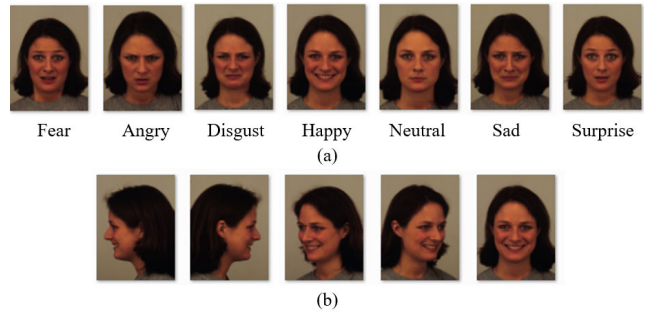
The performance of the proposed framework was validated considering subject dependent and subject independent facial expression recognition.

1) SUBJECT DEPENDENT FER
For subject dependent expression recognition, the JAFFE and KDEF databases were used individually. A single dataset, containing the facial expressions of all the subjects, is divided into training and testing sets with no overlapping using hold out(HO) method. 5% of the images from the JAFFE database and 60% from KDEF database are used for training to build the classifier model. The remaining samples are used to evaluate the model after testing. Though the testing and training sets are disjoint, they are framed considering the images of all the subjects from the dataset as shown in Fig. 4(a) for the JAFFE database.

2) SUBJECT INDEPENDENT FER
For Subject independent FER, the subjects or persons considered with their facial expressions for training and testing should be disjoint. Thus the images of the 20 female subjects from the KDEF database were considered for training while the facial expression images of the 10 subjects from the JAFFE database were used for testing the model.

B. FEATURE EXTRACTION
Features were extracted from the facial expression images to carry out classification and recognition.

1) FEATURE EXTRACTION USING ZERNIKE MOMENTS
During a facial expression, a lot of movements in the structural elements of the face: eyebrows, eyes, nose, mouth and chin can be observed, thus resulting in changes in the appearance of the face. The most prominent change is in the shape characteristics of the face. These changes need to be captured and represented by a suitable shape descriptor to recognize and classify the facial expression to interpret the emotion. The preferred shape descriptor should have two characteristics to provide the best discrimination between the facial expressions. The characteristics are (i) invariance to rotation, translation and scaling (ii) lower redundancy between the features. Zernike Moments [38], from the class
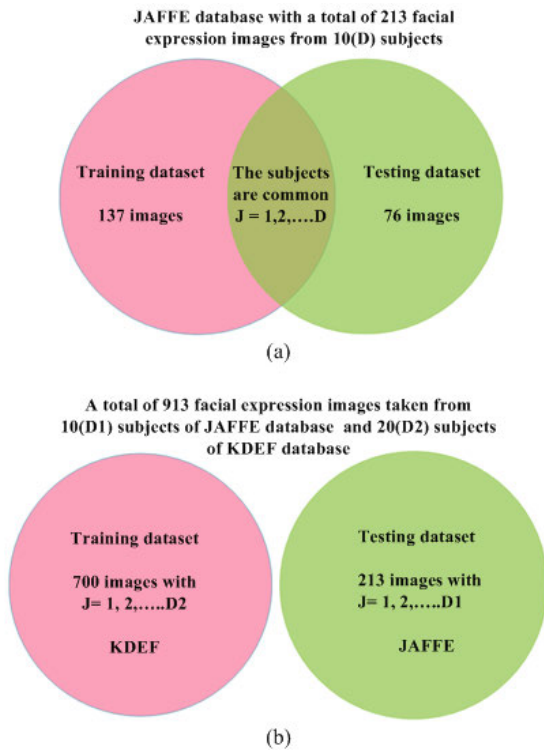
FIGURE 4. Dataset division (a) Subject dependent-JAFFE and (b) Subject independent FER.

FIGURE 5. SPZM feature of a sample image.

of orthogonal moments, proved acceptable to be the shape descriptors with a higher degree of information satisfying the above requirements [39]–[41]. The orthogonality of the moments provides lower possible redundancy and the magnitude of the ZM is invariant to rotation. Further, the lower and higher-order moments provide global shape and detailed information respectively. Thus, to capture the complete shape information from an image, a set of ZM is computed by varying the moment orders from a lower value to the higher. This process involves a significantly larger number of computations.

To reduce the complexity of the computation, the size of the set of ZM is predefined and combined with the spatial pyramid of the image to obtain the spatial pyramid ZM feature. A key characteristic of ZM that best complies with the spatial pyramid of the image is its hierarchical nature. The lower order moments provide the global shape information and the higher-order moments reveal the local shape information.

2) SPZM FEATURES

The SPZM feature is motivated by the concept of spatial pyramid representation of an image [42], [43], the shape descriptors and, Zernike moments. The idea is to extract the shape information at different levels of the pyramid to create a multilevel shape representation to capture the complete pattern created from the facial expression. The ZMs are computed at each pyramid level and are concatenated to form SPZM feature vector. Thus, this descriptor can be viewed as
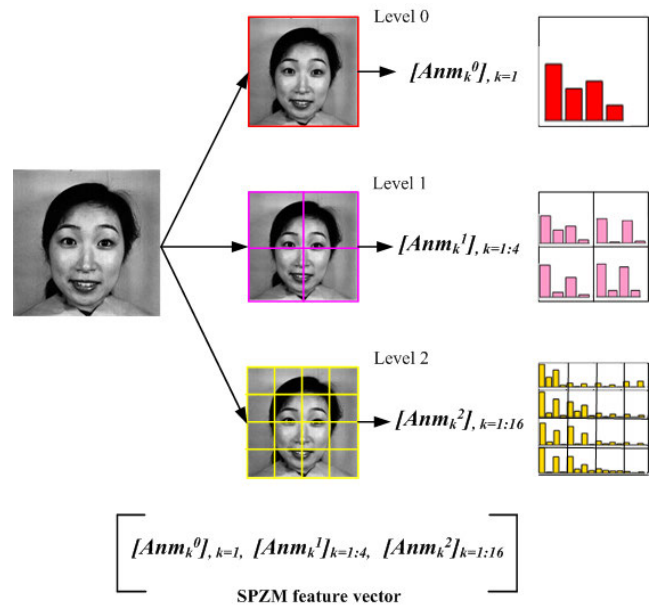
a combination of global shape characteristics and local shape characteristics of an image that can efficiently discriminate facial expressions. The extraction of SPZM involves (i) construction of a set of grids at different levels. (ii) computation of ZM for each sub-region or cell of the pyramid levels and concatenation of the moments.

The facial image $I(x, y)$ is divided into a set of grids at levels 0,..,L where L=2, such that the grid at each level l has 2^{2l} sub-regions or cells $C(x, y) = I(x, y)_k^l$ along each axis with $l = 0, \ldots\ldots L$ and $k = 1, \ldots. 2^{2l}$. Now for each cell $C(x, y)$, ZM is computed to extract the features. The ZM is computed for all the cells at different levels of the image pyramid and integrate them to form SPZM feature as shown in Fig. 5.

The ZM is computed by projecting each cell $C(x, y)$ on to the set of complex Zernike polynomials $V_{nm}(x, y)$

$$A_{nm} = \frac{m+1}{\pi} \sum_x \sum_y C(x, y) V_{nm}^*(\rho, \theta) \quad with \ x^2 + y^2 \leqslant 1$$

(1)

such that

$$V_{nm}(x, y) = V_{nm}(\rho, \theta) = R_{nm}(\rho) e^{jm\theta}$$

(2)

where,

n is the order of the polynomial

m is the repetition factor such that $|m| \leqslant n$ and $n - |m|$ is even

ρ is the length of the vector from the origin to the pixel located at spatial location (x, y) and is presented by $\rho = \sqrt{x^2 + y^2}$

θ angle of the vector from origin to the pixel located at spatial location, (x, y) from the x-axis in counter clockwise direction and

$R_{nm}(\rho)$ is the radial polynomial defined as

$$R_{nm}(\rho) = \sum_{s=0}^{n-|m|} (-1)^s \frac{(2n+1-s)!}{s!(n-|m|-s)!(n+|m|+1-s)!} \rho^{n-s}$$

(3)

The ZM thus computed is a complex quantity, from which the magnitude $|A_{nm}|$ and phase Arg (A_{nm}) are obtained. The magnitude $|A_{nm}|$ is selected since it is preserved as the shape descriptor by varying the order of the moment and the repetition factor. For a cell $C(x, y)$ a set of ZM magnitudes $[|A_{11}|, |A_{20}|, |A_{22}|, |A_{31}|]$ is computed by varying the moment order from 1 to 3 with the associated repetition factors. The moment $|A_{00}|$ is ignored as the zeroth-order Zernike polynomial. V_{00} is flat, in turn, an image projected on to this polynomial does not provide any edge maps or shape information. In ZM, the order of the polynomial indicates the radial component and the repetition factor indicates the sinusoidal component. With higher-order, the number of turning points in Zernike polynomial rises and thus provides better shape information of an image. The next higher-order polynomials next to zero are considered for image description in this work. The Zernike polynomial basis functions $V_{00}, V_{11}, V_{20}, V_{22}, V_{31}$ over a unit circle and their magnitudes $|V_{nm}|$ are provided in Fig. 6(a) and Fig. 6(b). Fig. 6(b) illustrates the rotational invariance characteristics of the ZM. The one dimension profile of the 2D Zernike polynomials is displayed in Fig. 6(c) to emphasize the number of turning points obtained with a variation in moment order. The complete shape description for an image (Fig. 5) is obtained by concatenating $[|A_{11}|, |A_{20}|, |A_{22}|, |A_{31}|]$ obtained for each $C(x, y)$ at levels $l = 0, 1, 2$ as indicated in equation (4) to form SPZM features.

$$\text{SPZM} = \{[|A_{nm}|_{k\,k=1}^0], [|A_{nm}|_{k\,k=1:4}^1], [|A_{nm}|_{k\,k=1:16}^2]\}_{n=1,2,3}$$

(4)

3) LAW's TEXTURE FEATURES

Texture features provide the statistical measures of an image based on the spatial arrangement of the pixel intensities. In the proposed work, the texture features are utilized to capture the micro details of the facial expression reflecting the texture of the skin through the formation of wrinkles and fine lines. These details are significant and contribute to identifying and recognizing the emotional state associated with expression. Kenneth I Laws [44] proposed texture energy measures for classifying the textures. These features are invariant to changes in rotation, contrast and luminance. These measures find applications in various domains [45]–[47].

The Law's texture feature extraction method is presented as,

1) Law's texture features used a set of one-dimensional convolution masks that are center-weighted and symmetrical or anti-symmetrical with varying dimensions of 1×3, 1×5 and 1×7. It is proved that masks with dimension 1×5 provide better texture descriptors. Each
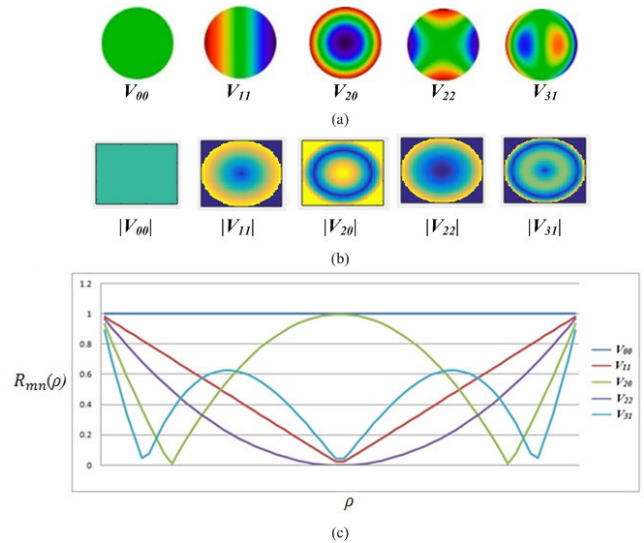

FIGURE 6. Zernike polynomial basis functions (a) $V_{00}, V_{11}, V_{20}, V_{22}, V_{31}$ over a unit circle (b) Magnitudes of the basis functions (c) Zernike polynomials in 1D.

of the masks is used to analyze the level, edge, spot and ripple characteristics of the image and named appropriately. The details of the 1×5 masks are as shown

L5(level)	[1 4 6 4 1]
E5(Edge)	[−1 2 0 2 1]
S5(Spot)	[−1 0 2 0 −1]
R5(Ripple)	[1 −4 6 −4 1]

2) One dimensional masks were combined to generate two-dimensional masks by computing the outer product of each vector as indicated,

	L5	E5	S5	R5
L5	L5L5	L5E5	L5S5	L5R5
E5	E5L5	E5E5	E5S5	E5R5
S5	S5L5	S5E5	S5S5	S5R5
R5	R5L5	R5E5	R5S5	R5R5

This resulted in a set of 16 convolution masks $\{CM_i\}_{i=1,2,\ldots\ldots16}$, each with a dimension of 5×5. The appropriate CM were selected from the convolution masks based on the listed criteria: (i) the mask L5L5 was dropped as it is sensitive to changes in the intensities and the sum of the elements of the mask adds up to a non zero value. (ii) The masks that provide a similar type of information were combined. For example, L5E5 measures vertical edge content while E5L5 presents horizontal edge content, thus the average will provide the total edge information. Accordingly, a set of 9 convolution masks $\{CM_i\}_{i=1,2,\ldots\ldots9}$ were finally selected as presented in equation 5.

$$\{CM_i\}$$
$$= \{L5E5/E5L5, L5S5/S5L5, L5R5/R5L5, E5S5/S5$$
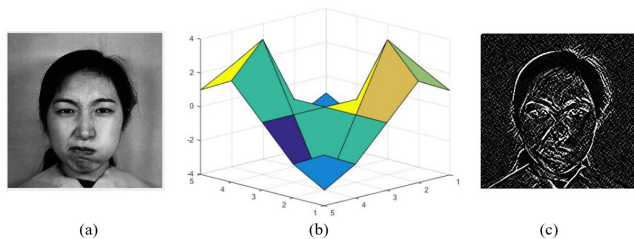$$E5, E5R5/R5E5, R5S5/S5L5, E5E5, S5S5, R5R5\}$$

(5)

FIGURE 7. Sample of an individual (a) with an angry expression, (b) Convolution mask E5E5 and (c) Law's texture image resulted from convolution with E5E5 mask from the JAFFE database.

FIGURE 8. Computation of Law's texture feature for a sample image.

3) Later, each of the facial images $I(x, y)$ is convolved with a set of masks $\{CM_i\}_{i=1,2........9}$ to produce a set of Law's texture images $\{LT_i\}_{i=1,2........9}$ representing the nine texture descriptors of a pixel for the input image. Fig. 7 presents a sample of the convolution mask with the corresponding texture image.

4) Finally, Law's texture energy measures, mean and standard deviation were computed for every image of the set $\{LT_i\}_{i=1,2........9}$, which form the texture descriptor with a length of 18 describing the variations in the image during an expression as indicated in Fig. 8.

C. FEATURE INTEGRATION

The Law's texture features and shape features from ZM are concatenated to form an integrated feature vector FV_{int}. As the contributions from the texture and shape features are different, the feature set consolidates the information from both the features using a series fusion rule.

$$FV_{int} = \{\text{Law's texture features, Spatial pyramid ZM features}\}$$
$$= \{\text{LTexM, SPZM}\} \qquad (6)$$

During integration, no weights were considered to give equal preference to both the feature types. The integrated feature set is normalized using Z-score normalization to form a new vector FV_{int_norm} with mean 0 and standard deviation 1. Feature normalization is crucial as it is one of the requirements of the machine learning algorithm and the magnitudes of the features in the feature vector influence the weight update during training process [49]. The normalized feature set is represented by,

$$FV_{int\_norm} = \frac{FV_{int} - \mu}{\sigma} \qquad (7)$$

where μ is the mean and σ is the standard deviation of the features. The features are normalized by retaining their original properties.

D. CLASSIFICATION USING ARTIFICIAL NEURAL NETWORKS (ANN)

The extracted texture and shape features from the facial expression image are fused to form an integrated feature set FV_{int_norm}. The concatenated feature vector is provided to the feed forward artificial neural network (ANN) for classification. ANNs are used extensively for facial expression
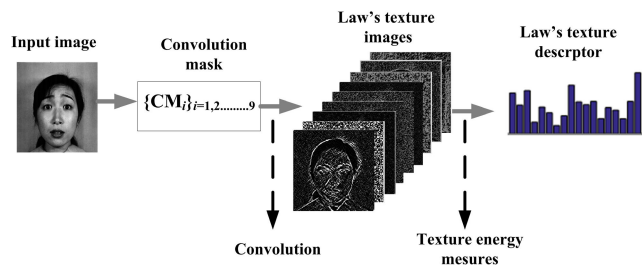
recognition [40] as they have the ability (i) to learn from the examples and model complex non-linear relationships and (ii) to generalize well by predicting the class labels of unseen data. This work proposes to use multilayer perceptron feed forward artificial neural network (MLPNN) and radial basis function neural network for classification

To train both MLPNN and RBFNN, the training images from the JAFFE and KDEF databases were considered as mentioned in section III.A and remaining images were used for testing. The FV_{int_norm} extracted from the training images were labeled for emotions from $C = [c_1, c_2, \ldots, c_7]$, where c_i represents the facial expressions like angry, disgust, fear, happy, sad, surprise and neutral.

1) MULTILAYER PERCEPTRON NEURAL NETWORK

The neural network architecture is designed with an input layer, one hidden layer associated with a hyperbolic tangent activation function $f(x) = \left(\frac{2}{1+e^{-2x}-1}\right)$ and an output layer with a linear activation function. Each layer is associated with the trainable parameters known as weights and biases. The input layer accepts an input vector, and passes to the hidden layer for processing. The hidden layer output is fed to the output layer to present the network's final output.

The MLPNN network is trained with the labeled feature vector $\{FV_{int\_norm}, C\}$ to adjust the trainable parameters for tuning the output close to the target values. For training, Levenberg-Marquardt back propagation algorithm [48] was used with mean squared error as the cost function. The stopping criteria was set by defining the number of epochs. MLPNN was tuned by providing performance goal(MSE), learning rate (η), momentum(m) and regularization parameter (λ) to generate the optimized classifier model.

2) RADIAL BASIS FUNCTION NEURAL NETWORK

RBFNN is a three layered network with input, hidden and output layers. The network is robust and converges at a faster rate. RBFNN uses Gaussian activation function from the class of radial functions in the hidden layer which is monotonic with respect to the distance from the center. The Gaussian function with center μ and spread/radius is represented by

$$G(x) = exp(-\frac{(FV_{int\_norm} - \mu)^2}{\sigma^2}) \qquad (8)$$

For the known input FV_{int_norm}, the hidden layer computes the distance between the input and centers of the basis functions $\|(FV_{int\_norm} - \mu)\|$, where $\|.\|$ is the L2 norm and applies RBF function. The computation is continued in the output layer that predicts the class label of the sample. The RBFNN is provided with $\{FV_{int\_norm}, C\}$ for training. During training the learnable parameters of the network, spread of the activation function and the weights connecting hidden layer to output layer, are tuned to obtain the best model.

The effectiveness of the models are now tested using the samples from testing set with unseen samples. For the testing sample, the features $FV_{int\_norm\_test}$ are extracted and is fed to classifier model to predict the class label c_{test}. where $c_{test} = f(FV_{int\_norm\_test}; \{FV_{int\_norm}, C\},$ MLPNN/RBFNN parameters) and $c_{test} = \{c_i\}_{i=1,2........7}$

IV. RESULTS AND DISCUSSION

To evaluate the performance of the proposed method, two experiments (i) subject dependent FER and (ii) subject independent FER were performed on JAFFE and KDEF databases.

A. SUBJECT DEPENDENT FER

The first experiment was conducted to investigate the combination of ZM based shape descriptors and Law's texture features for efficient facial expression representation and discrimination. This was experimented on the JAFFE and KDEF databases separately using the hold out method. As per the method, two disjoint sets were created for training and testing the classifier model.

1) CASE (I)

For all the images of the training set the SPZM and texture features LTexM were computed. The SPZM are extracted by considering the ZM orders $n = 1,2,3$ with the associated repetition factors as presented in equation (4). The SPZM shape features of length 84 are integrated with LTexM and normalized to form $FV_{int\_norm\_1}$ with dimension 1×102. The labeled and normalized feature vector is provided to the neural network classifiers for training the model that can further predict the expression label for the unseen sample.

The MLPNN was trained by providing the labeled training dataset to learn the relationship between the input feature vectors and the class labels by modifying the weights and biases. The network was tuned appropriately by setting the number of epochs limit to 1000 and cost function close to zero. The network converges and provide the best model for $\eta = 0.5$, $m = 0.95$ and $\lambda = 0.6$. The model is tested by providing the facial images from the testing set. The ability of the model in classifying all the facial expressions is assessed by computing the performance metrics from the confusion matrix that include classification accuracy(CA), true negative rate (TNR), false acceptance rate(FAR) and false rejection rate (FRR). Table 1 shows the confusion matrix.

Where, True Positive (TP): true accept.

True Negative (TN): True reject

TABLE 1. Confusion Matrix.

Actual facial expression	Predicted facial expression	
	TP	FN
	FP	TN

TABLE 2. Confusion matrix of MLPNN experimented with JAFFE database.

Angry		Disgust		Fear		Happy	
9	1	5	6	10	2	9	3
1	65	3	62	3	61	3	61

Neutral		Sad		Surprise	
9	1	5	6	10	0
2	64	2	63	0	66

TABLE 3. Performance metrics computed from testing dataset for MLPNN with $FV_{int\_norm\_1}$.

Emotions	JAFFE				KDEF			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	97.36	98.48	1.51	1.13	77.5	85.8	14.16	8.34
Disgust	88.15	95.38	4.61	7.24	74.28	81.66	18.33	7.39
Fear	93.42	95.31	4.68	1.9	85.71	95.8	4.16	10.13
Happy	92.1	95.31	4.68	3.22	77.5	84.5	15.4	7.1
Neutral	96.05	96.96	3.03	0.92	81.07	91.25	8.75	10.18
Sad	89.47	96.92	3.07	7.46	77.5	88.3	11.6	10.9
Surprise	100	100	0	0	83.57	96.25	6.6	9.83

False Positive (FP): False accept.

False negative (FN): False reject.

Here TP and TN represent the correct facial expression classifications whereas FP and FN are miss classifications, For example, a happy face falsely recognized as sad or angry.

From the result of the testing process, confusion matrices [49] are framed for each expression for both databases as presented in Table 2. From the attributes of the confusion matrix, the performance metrics are calculated.

$$\text{Classification accuracy} = \frac{TP + TN}{TP + TN + FP + FN} \quad (9)$$

$$\text{TNR/Specificity} = \frac{TN}{TN + FP} \quad (10)$$

$$\text{FAR} = \frac{FP}{TN + FP} \quad (11)$$

$$\text{FRR} = \frac{FN}{FN + TP} \quad (12)$$

The performance metrics computed for all the expressions of the testing dataset from JAFFE and KDEF are shown in Table 3.

The experiment was conducted to prove the discriminating ability of the integrated feature set in the accurate classification of facial expressions. From the results, it is observed that the proposed method delivers more improved classification accuracy. Also, it can be observed that the MLPNN provides high CA (greater than 90%) for the expressions angry, fear, happy, neutral and surprise for the facial images from JAFFE dataset and CA greater than 80% for fear, neutral and surprise from KDEF dataset.

(a)

(b)

FIGURE 9. Plot of FAR and FRR metrics provided by testing (a) JAFFE and (b) KDEF databases for MLPNN with $FV_{int\_norm\_1}$. A - Angry, D - Disgust, F - Fear, H - Happy, N - Neutral, SA - Sad, SU - Surprise.

Further, the specificity of the system was also measured through TNR. A value, TNR=1(100%), indicates the best specificity. A higher value of TNR close to one produces less false-positive results. From the results of the JAFFE database, it is observed that the average TNR is 96.90%, so the average prediction of false positives is only 3.1%. Similarly, for KDEF dataset, the average TNR is 89%.

In the next experiment, the system was tested for type-I and type-II errors which represent false match and false non-match respectively. FAR indicates type-I error whereas FRR indicates type-II error. Ideally, the two errors are inversely proportional to each other. False acceptance and false rejection recognize the facial expression and assign an emotion class label to which it does not belong to. These two are undesirable and may affect the success rate of a FER system. In some applications like psychological disorder detection, it may lead to false diagnosis and if a device has to take some actions based on human emotions. The plots shown in Fig. 9(a) and Fig. 9(b) display FAR and FRR of both databases. These plots aid in analyzing the relationship between FAR and FRR for each facial expression.

Fig. 9 notifies the inverse relationship between the errors which ensures that the proposed framework returns the least possible errors for identifying the emotions represented by the facial expressions. Thus the proposed system provided an average FAR of 3.08% and FRR 3.12% for JAFFE dataset. Similarly, the two measures are 11.28% and 9.12% respectively for KDEF dataset which is desirable for a FER system to contribute a higher performance rate.
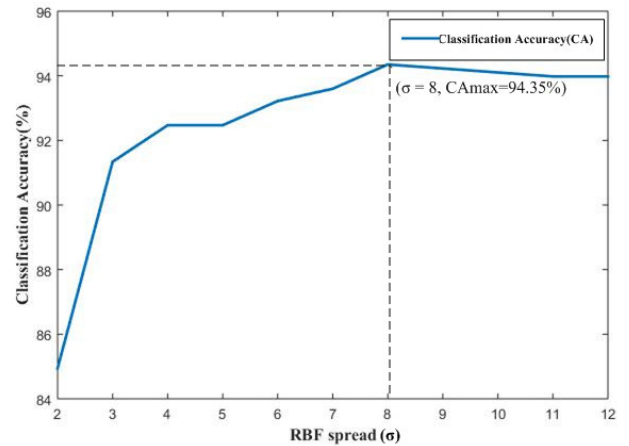
FIGURE 10. Plot of Classification accuracy for variation in RBF spread for $FV_{int\_norm\_1}$.

Later, the RBFNN was trained and tested with the training and testing sets from JAFFE and KDEF datasets. For the training process, the cost function, sum squared error(SSE), was set and the spread(σ) of the radial activation function was varied from $\sigma = 2$ in incremental steps of 1 to improve the performance. With each variation in σ, the accuracy of the model was noted. The network converged for $\sigma = 8$. Fig. 10 shows the accuracy of the model with respect to spread of RBF obtained for JAFFE database.

From the Fig. 10, it is noted that the maximum accuracy of 94.35% is obtained for $\sigma = 8$. With the optimized value, the performance metrics are computed for all the expressions of the testing datasets and are presented in Table 4.

TABLE 4. Performance metrics computed from testing dataset for RBFNN with $FV_{int\_norm\_1}$.

Emotions	JAFFE				KDEF			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	97.36	98.48	1.51	1.13	74.28	82.08	17.91	7.81
Disgust	89.47	93.84	6.15	4.38	73.92	81.25	18.75	7.33
Fear	90.78	95.31	4.68	4.54	82.14	92.5	7.5	10.36
Happy	96.05	98.43	1.56	2.39	80.35	91.66	8.33	11.32
Neutral	98.68	98.48	1.51	0	79.64	89.58	10.41	9.95
Sad	93.42	96.92	3.07	3.51	83.21	96.25	3.75	13.04
Surprise	94.73	95.45	4.54	0.73	85.35	97.08	2.91	11.74

The analysis of the performance measures provided by RBFNN indicates that the expressions angry, happy, neutral were classified with CA greater than 96% from JAFFE dataset. Similarly the expressions fear, happy, sad and surprise were classified accurately with CA greater than 80% from KDEF dataset. Similarly, the model also provides the specificity (average) of 96.70% and 90.05% for JAFFE and KDEF datasets respectively which is an indication of lower false positives. The evaluation of type-I and type-II errors indicates the performance of the model in providing the least error values, FAR of 3.28% and FRR 2.38%, for JAFFE dataset. The two measures are 9.93% and 10.22% respectively for KDEF dataset.

TABLE 5. Performance metrics computed from testing dataset for MLPNN with $FV_{int\_norm\_2}$.

Emotions	JAFFE				KDEF			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	96.05	96.96	3.03	0.92	86.42	92.5	7.5	6.08
Disgust	90.78	95.38	4.61	4.61	85.71	90.79	9.2	5.09
Fear	88.15	95.31	4.68	7.17	91.07	97.91	2.08	6.85
Happy	96.05	96.87	3.12	0.83	88.92	95.41	4.5	6.58
Neutral	97.36	98.48	1.51	1.13	83.92	89.16	10.8	5.28
Sad	93.42	93.84	6.15	0.43	82.5	85	15	2.5
Surprise	98.68	98.48	1.51	0	89.64	95.83	4.16	6.2

TABLE 6. Performance metrics computed from testing dataset for RBFNN with $FV_{int\_norm\_2}$.

Emotions	JAFFE				KDEF			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	98.68	100	0	1.32	88.92	94.58	5.41	5.61
Disgust	93.42	96.92	3.07	3.51	88.21	92.91	7.08	4.71
Fear	92.1	96.87	3.12	4.78	89.64	95	5	5.36
Happy	97.4	98.43	1.56	1.04	92.14	97.5	2.5	5.36
Neutral	96.05	96.92	3.95	0	86.42	90.41	9.58	4
Sad	96.05	98.46	1.53	2.44	86.42	88.33	11.66	1.92
Surprise	97.36	96.96	2.64	0	90.35	95.83	4.16	5.49

2) CASE (II)

The experiment on subject dependent FER was continued with the next set of feature vectors. At this point, the texture features remain the same as considered in case (i) whereas the SPZM are captured by varying the ZM order from $n = 2$ to 4. The variation in n with the corresponding repetition factors resulted in $[|A_{20}|, |A_{22}|, |A_{31}|, |A_{33}|, |A_{40}|, |A_{42}|]$ which are utilized in equation (4) to provide SPZM of length 126. The order of ZM was varied to explore the competency of higher order ZM in presenting best shape features. Now these features are combined with LTexM features to form integrated feature vector $FV_{int\_norm\_2}$ with dimension 1×144.

The $FV_{int\_norm\_2}$ framed for the training datasets of JAFFE and KDEF are provided to both MLPNN and RBFNN classifiers. The classifiers were trained by varying the trainable parameters of the network to improve the performance. The MLPNN parameters optimized for $\eta = 0.5$, $m = 0.95$ and $\lambda = 0.6$ and RBFNN converged for $\sigma = 18$. The models were later evaluated for their performance by providing the samples of the testing data set. The measures computed on testing dataset with $FV_{int\_norm\_2}$ for both MLPNN and RBFNN are presented in Table 5 and Table 6.

The result and the analysis of the performance measures obtained from the second set of feature vectors $FV_{int\_norm\_2}$ implies improvement in the performance of the FER system. The average classification accuracies of 94.35% & 95.86% for JAFFE and 94.35% & 95.86% were achieved from MLPNN and RBFNN classifiers respectively. Another noteworthy improvement is in terms of type-I and type-II errors. The results clearly indicate that the average type-I error for all the expressions is reduced by 30% and type-II error by 36.84%. On the similar terms, CA of 86% and 88.87% was observed.

The results presented by MLPNN and RBFNN for both $FV_{int\_norm\_1}$ and $FV_{int\_norm\_2}$ were also examined to check for the discrimination ability. The assessment
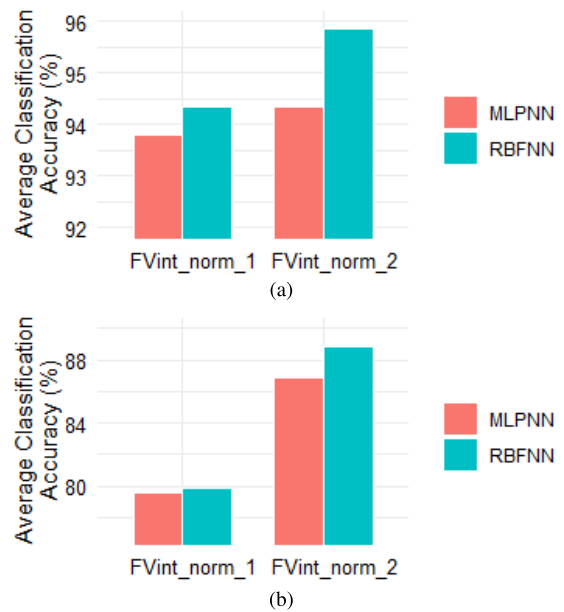

FIGURE 11. Average Classification Accuracy of the seven expressions (a) JAFFE and (b) KDEF databases.

TABLE 7. Performance comparison.

Methods	Average Classification/ Recognition accuracy (%)	
	JAFFE Dataset	KDEF Dataset
Histogram of oriented gradient(HOG) with SVM [50]	76.19	80.95
SURF with logistic regression[51]	-	74.05
VGG-16 with transfer learning, ResNet50 with transfer learning[52]	-	73.6, 76
DeepPCA with ELM[53]	-	83
Auxiliary Models [54]	95	-
LBP[55]	93.8	-
Adapted gradient LPQ[56]	92.59	-
Geometric features with modified HMM[57]	84.7	-
Histogram of oriented gradient(HOG) in wavelet domain[58]	71.43	-
Exemplar based[59]	92.53	-
Deep Convolutional Neural Network[60]	-	59.64
Kernel collaboration representation and deep subspace learning[61]	68.80	80.20
Modified LDP Features with the Optimization based Multi-SVNN Classifier[62]	96.00	-
Proposed method (RBFNN)	95.86	88.87

signified the performance of RBFNN that can be noted from the Fig. 11.

Later, the analysis is extended to compare the performance of the proposed method with state of the art facial expression recognition techniques and the same is presented in Table 7.

A comparative evaluation was carried out considering different image descriptors for the JAFFE and KDEF databases. The Table 7 provides the average classification accuracy which implies that the proposed method is comparable with the state of the art techniques.

TABLE 8. Performance metrics computed for subject independent FER with $FV_{int\_norm\_1}$.

Emotions	MLPNN				RBFNN			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	85.06	98.48	1.51	13.43	73.83	79.34	20.65	5.52
Disgust	76.16	95.38	4.61	19.23	80.64	94.5	5.49	13.87
Fear	71.49	95.31	4.68	23.83	74.29	82.06	17.93	7.78
Happy	70.56	95.31	4.68	24.76	81.77	92.53	7.06	11.17
Neutral	75.7	96.96	3.03	21.27	77.57	88.04	11.95	10.48
Sad	74.29	96.92	3.07	22.64	72.69	81.56	18.03	9.28
Surprise	82.24	93.47	6.52	11.24	82.24	89.69	10.32	7.44

TABLE 9. Performance metrics computed for subject independent FER with $FV_{int\_norm\_2}$.

Emotions	MLPNN				RBFNN			
	CA(%)	TNR(%)	FAR(%)	FRR(%)	CA(%)	TNR(%)	FAR(%)	FRR(%)
Angry	75.23	86.95	13.05	11.72	79.43	93.29	6.7	13.87
Disgust	69.15	78.02	21.98	8.87	75.23	86.81	13.18	11.59
Fear	79.43	90.76	9.23	11.34	78.03	88.04	11.95	10.02
Happy	74.76	84.69	15.31	9.93	73.83	81.42	18.57	7.6
Neutral	80.84	92.39	7.61	11.55	72.89	77.17	22.82	4.29
Sad	74.29	83.06	16.94	8.77	85.04	98.9	1.09	13.87
Surprise	75.23	84.23	15.77	9	81.77	92.93	2.17	16.06

B. SUBJECT INDEPENDENT FER

In the second experiment, the suitability of the extracted and integrated feature sets was checked for subject independent FER. To evaluate the performance, the combination of both JAFFE and KDEF databases was considered. The integrated features $FV_{int\_norm\_1}$ and $FV_{int\_norm\_2}$ were extracted from the facial images of both KDEF and JAFFE databases. The feature vectors of the KDEF dataset were labeled to form the training set while the JAFFE dataset forms the testing set. The performance metrics obtained with MLPNN and RBFNN are shown in Table 8 and Table 9.

The results were analyzed to find that the model with $FV_{int\_norm\_1}$ and MLPNN delivered a CA greater than 80 percentage for the angry and surprise expressions but for disgust, happy and surprise RBFNN delivers. Similarly, for the feature set $FV_{int\_norm\_2}$, an accuracy greater than 80 percentage was obtained for neutral expression with MLPNN and sad & surprise expressions with RBFNN. Also the average specificity is found to be greater than 85 percentage for both the integrated feature vectors for MLPNN and RBFNN that reduces the false positives. The metrics FAR and FRR presented in Tables 8 and 9 suggest that the proposed method has better generalization capability. The proposed method is able to identify the true facial expression independent to the subject. It is observed that the recognition of facial expression is improved by considering higher order ZM ($FV_{int\_norm\_2}$) and RBFNN performs better as compared to MLPNN in identifying the subject independent facial expression.

The descriptors were also significant in discriminating the emotions across the datasets (subject independent).

C. ABLATION STUDY

An additional experiment was conducted using ablation study to prove the effectiveness of integrating the Law's texture features with SPZM features in enhancing the performance of both subject dependent and independent FER.
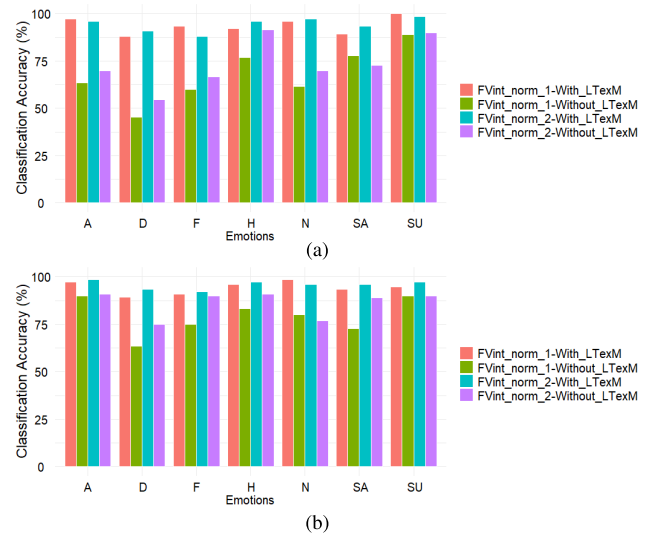
FIGURE 12. Plot of classification accuracies of all the expressions testing JAFFE dataset for subject dependent FER (a) MLPNN and (b) RBFNN.
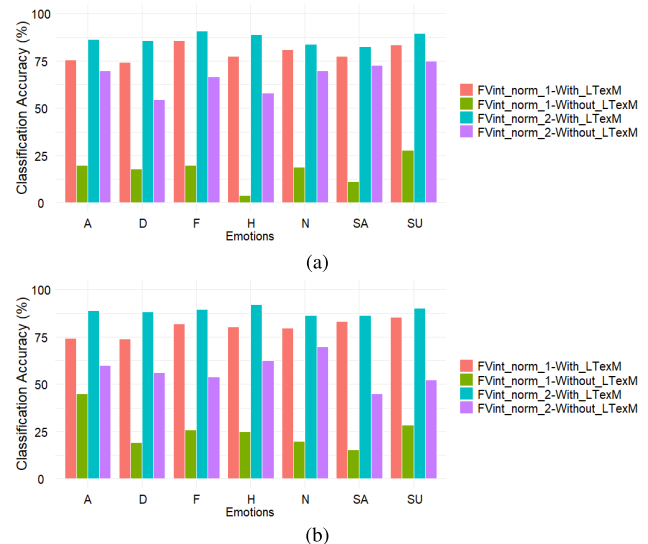
FIGURE 13. Plot of classification accuracies of all the expressions testing KDEF dataset for subject dependent FER (a) MLPNN and (b) RBFNN.

For ablation study, the Law's texture features were removed from the integrated feature set retaining only the SPZM features in both training set and testing set. $(FV_{int\_norm\_i}) = \{SPZM\}_{j=1(n=1,2,3), j=2(n=2,3,4)}$

Now the reduced training feature set with the class labels were provided to both MLPNN and RBFNN classifiers to create the model. For training the neural network the same criteria and tuning parameters were retained as mentioned in section III.D. The trained model is then tested with a reduced testing feature set. The performance is evaluated for recognition accuracy to investigate how the removal of texture features affects the performance of the subject dependent and independent FER. The results of the ablation study are displayed in Fig. 12, 13 and 14.

The classification accuracies shown in Fig. 12,13 and 14 obtained with the two neural network classifiers for both
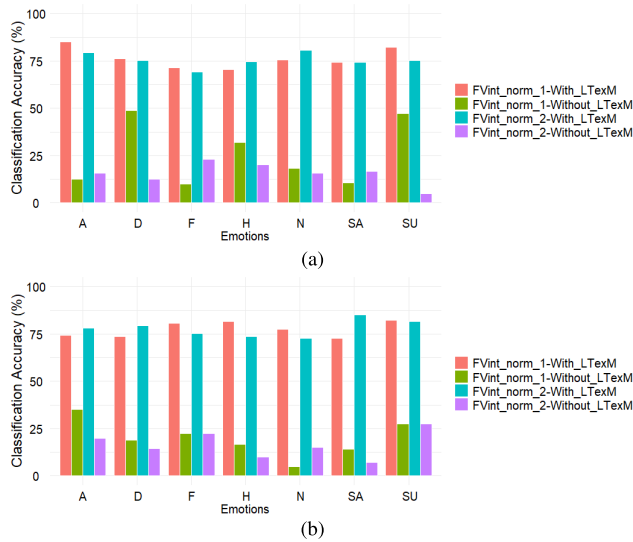
FIGURE 14. Plot of classification accuracies of all the expressions testing JAFFE dataset for subject independent FER (a) MLPNN and (b) RBFNN.

integrated feature vectors indicate that the combination of texture and shape descriptors are appropriate to represent the facial expressions for recognition followed by classification. From this study, It is noted that the integration of texture features contributes significantly in improving the recognition accuracy. It can also be noted that the RBFNN generalizes well by providing better performance in FER.

V. CONCLUSION

This paper presents a FER method using the combination of Law's texture features and ZM based shape descriptors. The proposed method uses feed forward neural networks MLPNN and RBFNN algorithms to learn the relationship between the features and the class labels of the emotions. The results proved that the proposed method has a good class discrimination ability and performed well with $FV_{int\_norm\_2}$ where the texture features were integrated with SPZM obtained from higher order ZM. The RBFNN classifier presented a noticeable result with an average recognition accuracy of 95.86% and 88.87% on the JAFFE and KDEF datasets respectively for subject dependent FER using RBFNN. The model classified the angry, happy and surprise expressions with good recognition rate whereas the performance for disgust, neutral and sad expressions is quite small in comparison. The method is also able to perform well even in the presence of facial orientations. Subsequently, the proposed method has better generalization capability in identifying the facial expressions across datasets for subject independent FER. The framed feature set proved to be accurate in capturing the facial expression patterns and thus returned less number of miss classifications by providing least possible type-I and type-II errors.

Further, the effectiveness of combined features can be extended for better understanding and recognition of the micro expressions. These expressions last only for a fraction

of a second on the face and that actually represent the true feeling of an individual. The proposed method will be revised for both spatial and motion information of the image sequence to describe the key characteristics from the dynamic facial expressions for classification.

REFERENCES

[1] P. Ekman and W. V. Friesen, *Facial Action Coding System: Investigator's Guide*. Palo Alto, CA, USA: Consulting Psychologists Press, 1978.

[2] X. Fan and T. Tjahjadi, "A dynamic framework based on local zernike moment and motion history image for facial expression recognition," *Pattern Recognit.*, vol. 64, pp. 399–406, Apr. 2017.

[3] Y.-I. Tian, T. Kanade, and J. F. Cohn, "Recognizing action units for facial expression analysis," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 23, no. 2, pp. 97–115, Feb. 2001.

[4] Z. Zhang, M. Lyons, M. Schuster, and S. Akamatsu, "Comparison between geometry-based and Gabor-wavelets-based facial expression recognition using multi-layer perceptron," in *Proc. 3rd IEEE Int. Conf. Autom. Face Gesture Recognit.*, Apr. 1998, pp. 454–459.

[5] A. L. A. Ramos, B. G. Dadiz, and A. B. G. Santos, "Classifying emotion based on facial expression analysis using Gabor filter: A basis for adaptive effective teaching strategy," in *Computational Science and Technology*. Singapore: Springer, 2020, pp. 469–479.

[6] E. Owusu, Y. Zhan, and Q. R. Mao, "A neural-AdaBoost based facial expression recognition system," *Expert Syst. Appl.*, vol. 41, no. 7, pp. 3383–3390, Jun. 2014.

[7] A. Hernandez-Matamoros, A. Bonarini, E. Escamilla-Hernandez, M. Nakano-Miyatake, and H. Perez-Meana, "Facial expression recognition with automatic segmentation of face regions using a fuzzy based classification approach," *Knowl.-Based Syst.*, vol. 110, pp. 1–14, Oct. 2016.

[8] G. P. Hegde, M. Seetha, and N. Hegde, "Kernel locality preserving symmetrical weighted Fisher discriminant analysis based subspace approach for expression recognition," *Eng. Sci. Technol., Int. J.*, vol. 19, no. 3, pp. 1321–1333, 2016.

[9] T. Ojala, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *IEEE Trans.Pattern Anal. Mach. Intell.*, vol. 24, no. 7, pp. 971–987, Aug. 2002.

[10] M. J. Cossetin, J. C. Nievola, and A. L. Koerich, "Facial expression recognition using a pairwise feature selection and classification approach," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2016, pp. 5149–5155.

[11] X. Wei, H. Wang, G. Guo, and H. Wan, "A general weighted multi-scale method for improving LBP for face recognition," in *Proc. Int. Conf. Ubiquitous Comput. Ambient Intell.*, Dec. 2014, pp. 532–539.

[12] M. Guo, X. Hou, Y. Ma, and X. Wu, "Facial expression recognition using ELBP based on covariance matrix transform in KLT," *Multimedia Tools Appl.*, vol. 76, no. 2, pp. 2995–3010, 2017.

[13] X. H. Guo, X. Zhang, C. Deng, and J. Wei, "Facial expression recognition based on independent component analysis," *J. Multimedia*, vol. 8, no. 4, pp. 402–409, 2013.

[14] S. Shah and D. Shah, "Partition-based face recognition using LDP and SVM," *J. Inf. Commun. Technol. Robot. Appl.*, vol. 8, no. 1, pp. 37–43, 2017.

[15] Y. Wu and W. Qiu, "Facial expression recognition based on improved local ternary pattern and stacked auto-encoder," *AIP Conf. Proc.*, vol. 1864, Jul. 2017, Art. no. 020131.

[16] B. R. Ilyas, B. Mohammed, M. Khaled, A. T. Ahmed, and A. Ihsen, "Facial expression recognition based on DWT feature for deep CNN," in *Proc. 6th Int. Conf. Control, Decis. Inf. Technol. (CoDIT)*, Apr. 2019, pp. 344–348.

[17] Z. Wang, Q. Ruan, and G. An, "Facial expression recognition using sparse local Fisher discriminant analysis," *Neurocomputing*, vol. 174, pp. 756–766, Jan. 2016.

[18] H. K. Meena, K. K. Sharma, and S. D. Joshi, "Effective curvelet-based facial expression recognition using graph signal processing," *Signal, Image Video Process.*, vol. 14, no. 2, pp. 241–247, Mar. 2020.

[19] M. Dahmane and J. Meunier, "Emotion recognition using dynamic grid-based HoG features," *Proc. Int. Conf. Autom. Face Gesture*, Mar. 2011, pp. 884–888.

[20] M. M. F. Donia, A. A. A. Youssif, and A. Hashad, "Spontaneous facial expression recognition based on histogram of oriented gradients descriptor," *Comput. Inf. Sci.*, vol. 7, no. 3, pp. 31–37, Jul. 2014.

[21] D. Ghimire and J. Lee, "Geometric feature-based facial expression recognition in image sequences using multi-class AdaBoost and support vector machines," *Sensors*, vol. 13, no. 6, pp. 7714–7734, Jun. 2013.

[22] A. Sánchez, J. V. Ruiz, A. B. Moreno, A. S. Montemayor, J. Hernández, and J. J. Pantrigo, "Differential optical flow applied to automatic facial expression recognition," *Neurocomputing*, vol. 74, no. 8, pp. 1272–1282, Mar. 2011.

[23] C. Singh, N. Mittal, and E. Walia, "Face recognition using zernike and complex zernike moment features," *Pattern Recognit. Image Anal.*, vol. 21, no. 1, pp. 71–81, Mar. 2011.

[24] M. Saaidia, N. Zermi, and M. Ramdani, "Facial expression recognition using neural network trained with zernike moments," in *Proc. 4th Int. Conf. Artif. Intell. with Appl. Eng. Technol.*, Dec. 2014, pp. 187–192.

[25] M. Ahmady, R. Ghasemi, and H. Rashidy Kanan, "Local weighted pseudo zernike moments and fuzzy classification for facial expression recognition," in *Proc. 13th Iranian Conf. Fuzzy Syst. (IFSC)*, Aug. 2013, pp. 1–4.

[26] R. Jafri and H. R. Arabnia, "A survey of face recognition techniques," *J. Inf. Process. Syst.*, vol. 5, no. 2, pp. 41–68, 2009.

[27] V. G. Mahesh, A. N. Raj, and Z. Fan, "Invariant moments based convolutional neural networks for image analysis," *Int. J. Comput. Intell. Syst.*, vol. 10, no. 1, pp. 936–950, 2017.

[28] 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, Jan. 2018.

[29] K. Zhang, Y. Huang, Y. Du, and L. Wang, "Facial expression recognition based on deep evolutional spatial-temporal networks," *IEEE Trans. Image Process.*, vol. 26, no. 9, pp. 4193–4203, Sep. 2017.

[30] Y. Liu, X. Yuan, X. Gong, Z. Xie, F. Fang, and Z. Luo, "Conditional convolution neural network enhanced random forest for facial expression recognition," *Pattern Recognit.*, vol. 84, pp. 251–261, Dec. 2018.

[31] S. Xie and H. Hu, "Facial expression recognition using hierarchical features with deep comprehensive multipatches aggregation convolutional neural networks," *IEEE Trans. Multimedia*, vol. 21, no. 1, pp. 211–220, Jan. 2019.

[32] K. Li, Y. Jin, M. W. Akram, R. Han, and J. Chen, "Facial expression recognition with convolutional neural networks via a new face cropping and rotation strategy," *Vis. Comput.*, vol. 36, no. 2, pp. 391–404, Feb. 2020.

[33] H. Zhang, B. Huang, and G. Tian, "Facial expression recognition based on deep convolution long short-term memory networks of double-channel weighted mixture," *Pattern Recognit. Lett.*, vol. 131, pp. 128–134, Mar. 2020.

[34] X. Wang, X. Wang, and Y. Ni, "Unsupervised domain adaptation for facial expression recognition using generative adversarial networks," *Comput. Intell. Neurosci.*, vol. 2018, Jul. 2018, Art. no. 7208794.

[35] H. Yang, Z. Zhang, and L. Yin, "Identity-adaptive facial expression recognition through expression regeneration using conditional generative adversarial networks," in *Proc. 13th IEEE Int. Conf. Autom. Face Gesture Recognit.*, May 2018, pp. 294–301.

[36] *The Japanese Female Facial Expression (JAFFE) Dataset*. Accessed: Feb. 23, 2016. [Online]. Available: http://www.kasrl.org/jaffe.html

[37] D. Lundqvist, A. Flykt, and A. Öhman, "The Karolinska directed emotional faces (KDEF)," Dept. Clin. Neurosci., Psychol. Sect., CD ROM From, Karolinska Inst., Solna, Sweden, 1998, vol. 91, no. 630, p. 2.

[38] A. Khotanzad and Y. H. Hong, "Invariant image recognition by zernike moments," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 12, no. 5, pp. 489–497, May 1990.

[39] A. Tahmasbi, F. Saki, and S. B. Shokouhi, "Classification of benign and malignant masses based on zernike moments," *Comput. Biol. Med.*, vol. 41, no. 8, pp. 726–735, Aug. 2011.

[40] V. G. Mahesh and A. N. J. Raj, "Invariant face recognition using Zernike moments combined with feed forward neural network," *Int. J. Biometrics*, vol. 7, no. 3, pp. 286–307, 2015.

[41] V. Mahesh, A. Raj, and T. Celik, "Silkworm cocoon classification using fusion of zernike moments-based shape descriptors and physical parameters for quality egg production," *Int. J. Intell. Syst. Technol. Appl.*, vol. 16, no. 3, pp. 246–268, 2017.

[42] B.-D. Liu, J. Meng, W.-Y. Xie, S. Shao, Y. Li, and Y. Wang, "Weighted spatial pyramid matching collaborative representation for remote-sensing-image scene classification," *Remote Sens.*, vol. 11, no. 5, p. 518, Mar. 2019.

[43] S. Lazebnik, C. Schmid, and J. Ponce, "Spatial pyramid matching," *Object Categorization, Comput. Human Vis. Perspect.*, vol. 3, no. 4, p. 5, 2009.

[44] K. I. Laws, "Texture energy measures," in *Proc. Image Understanding Workshop*, Nov. 1979, pp. 47–51.

[45] O. Faust, U. R. Acharya, K. M. Meiburger, F. Molinari, J. E. W. Koh, C. H. Yeong, P. Kongmebhol, and K. H. Ng, "Comparative assessment of texture features for the identification of cancer in ultrasound images: A review," *Biocybern. Biomed. Eng.*, vol. 38, no. 2, pp. 275–296, 2018.

[46] A. S. Setiawan, Elysia, J. Wesley, and Y. Purnama, "Mammogram classification using Law's texture energy measure and neural networks," *Procedia Comput. Sci.*, vol. 59, pp. 92–97, Jan. 2015.

[47] S. Dash and U. R. Jena, "Texture classification using steerable pyramid based Laws' masks," *J. Electr. Syst. Inf. Technol.*, vol. 4, no. 1, pp. 185–197, May 2017.

[48] M. T. Hagan and M. B. Menhaj, "Training feedforward networks with the marquardt algorithm," *IEEE Trans. Neural Netw.*, vol. 5, no. 6, pp. 989–993, 1994.

[49] J. Han, J. Pei, and M. Kamber, *Data Mining: Concepts and Techniques*. Amsterdam, The Netherlands: Elsevier, 2011.

[50] S. K. Eng, H. Ali, A. Y. Cheah, and Y. F. Chong, "Facial expression recognition in JAFFE and KDEF Datasets using histogram of oriented gradients and support vector machine," in *Proc. IOP Conf. Ser., Mater. Sci. Eng.*, Nov. 2019, vol. 705, no. 1, pp. 1–7.

[51] Q. Rao, X. Qu, Q. Mao, and Y. Zhan, "Multi-pose facial expression recognition based on SURF boosting," in *Proc. Int. Conf. Affect. Comput. Intell. Interact. (ACII)*, Sep. 2015, pp. 630–635.

[52] Z. Elgarrai, O. E. Meslouhi, M. Kardouchi, and H. Allali, "Robust facial expression recognition system based on hidden Markov models," *Int. J. Multimedia Inf. Retr.*, vol. 5, no. 4, pp. 229–236, Nov. 2016.

[53] K. Rujirakul and C. So-In, "Histogram equalized deep PCA with ELM classification for expressive face recognition," in *Proc. Int. Workshop Adv. Image Technol. (IWAIT)*, Jan. 2018, pp. 1–4.

[54] Y. Wang, Y. Li, Y. Song, and X. Rong, "Facial expression recognition based on auxiliary models," *Algorithms*, vol. 12, no. 11, p. 227, Oct. 2019.

[55] M. Pietikäinen and A. Hadid, "Texture features in facial image analysis," in *Proc. Int. Workshop Biometric Person Authentication*. Berlin, Germany: Springer, Oct. 2005, pp. 1–8.

[56] S. Kherchaoui and A. Houacine, "Facial expression identification using gradient local phase," *Multimedia Tools Appl.*, vol. 78, no. 12, pp. 16843–16859, Jun. 2019.

[57] M. Rahul, "Facial expression recognition using geometric features and modified hidden Markov model," *Int. J. Grid Utility Comput.*, vol. 10, no. 5, pp. 488–496, 2019.

[58] S. Nigam, R. Singh, and A. K. Misra, "Efficient facial expression recognition using histogram of oriented gradients in wavelet domain," *Multimedia Tools Appl.*, vol. 77, no. 21, pp. 28725–28747, Nov. 2018.

[59] N. Farajzadeh and M. Hashemzadeh, "Exemplar-based facial expression recognition," *Inf. Sci.*, vols. 460–461, pp. 318–330, Sep. 2018.

[60] Y. I. Colon, M. Q. Hill, C. J. Parde, and C. D. Castillo, "Facial expression information in deep convolutional neural networks trained for face identification," *J. Vis.*, vol. 19, no. 10, p. 93, 2019.

[61] Z. Sun, Z.-P. Hu, R. Chiong, M. Wang, and W. He, "Combining the kernel collaboration representation and deep subspace learning for facial expression recognition," *J. Circuits, Syst. Comput.*, vol. 27, no. 8, Jul. 2018, Art. no. 1850121.

[62] I. Michael Revina and W. R. Sam Emmanuel, "Face expression recognition with the optimization based multi-SVNN classifier and the modified LDP features," *J. Vis. Commun. Image Represent.*, vol. 62, pp. 43–55, Jul. 2019.

VIJAYALAKSHMI G. V. MAHESH received the B.E. degree in electronics and communication engineering from Bangalore University, India, in 1999, the M.Tech. in digital communication and networking from Visvesvaraya Technological University, in 2005, and the Ph.D. degree from the Vellore Institute of Technology, Vellore, India. She is currently working as an Associate Professor with the BMS Institute of Technology and Management, Bengaluru, India. Her research interests include pattern recognition, machine learning, image processing, and deep learning.

CHENGJI CHEN was born in Guangdong, China. He received the Bachelor of Engineering degree, in 2019. He is currently pursuing the master's degree in engineering with Shantou University, Shantou, China. His research interests include image processing and facial expression recognition.

ALEX NOEL JOSEPH RAJ (Member, IEEE) received the B.E. degree in electrical engineering from Madras University, India, in 2001, the M.E. degree in applied electronics from Anna University, in 2005, and the Ph.D. degree in engineering from the University of Warwick, in 2009. From October 2009 to September 2011, he was a Design Engineer with Valeport Ltd., Totnes, U.K. From March 2013 to December 2016, he was a Professor with the Department of Embedded Technology, School of Electronics Engineering, Vellore Institute of Technology, Vellore, India. Since January 2017, he has been with the Department of Electronic Engineering, College of Engineering, Shantou University, China. His research interests include deep learning, signal and image processing, and FPGA implementations.

VIJAYARAJAN RAJANGAM (Member, IEEE) received the Bachelor of Engineering degree in electronics and communication from the University of Madras, in 1998, and the master's degree in applied electronics from Madurai Kamarajar University, in 1999. He did his Ph.D. research in image fusion with Anna University, in 2015. He is currently working as a Faculty Member with the Centre for Healthcare Advancement, Innovation and Research, School of Electronics Engineering, VIT University, Chennai. He has been in academics for over 20 years and has published his research in various reputed journals and conferences. His research interests include image fusion, bio-cryptography, emotion classification using signal and image processing, and deep learning for image analysis.

PALANI THANARAJ KRISHNAN (Member, IEEE) received the Ph.D. degree from the Faculty of Information and Communication Engineering, Anna University, Chennai, in 2018. He is currently working as an Assistant Professor with the Department of Electronics and Instrumentation Engineering, St. Joseph's College of Engineering, Chennai. He has developed deep learning algorithms for performing image classification of medical images for disease diagnosis. He has published his works in many reputed and refereed journals indexed in Web of Science and Scopus. His research interests include image processing, advanced signal processing, image segmentation, machine learning, and deep learning.

● ● ●

WILEY

An integrated feature frame work for automated segmentation of COVID-19 infection from lung CT images

Deepika Selvaraj[1] | Arunachalam Venkatesan[1] |
Vijayalakshmi G. V. Mahesh[2] | Alex Noel Joseph Raj[3]

[1]Department of Micro and Nano
Electronics, School of Electronics
Engineering, Vellore Institute of
Technology, Vellore, India

[2]Department of Electronics and
Communication, BMS Institute of
Technology and Management, Bangalore,
India

[3]Key Laboratory of Digital Signal and
Image Processing of Guangdong Province,
Department of Electronic Engineering,
College of Engineering, Shantou
University, Shantou, China

Correspondence
Alex Noel Joseph Raj, Key Laboratory of
Digital Signal and Image Processing of
Guangdong Province, Department of
Electronic Engineering, College of
Engineering, Shantou University, Shantou
China.
Email: jalexnoel@stu.edu.cn

Funding information
The Research Start-Up Fund Subsidized
Project of Shantou University, China,
Grant/Award Number: NTF17016

Abstract
The novel coronavirus disease (SARS-CoV-2 or COVID-19) is spreading across
the world and is affecting public health and the world economy. Artificial
Intelligence (AI) can play a key role in enhancing COVID-19 detection. How-
ever, lung infection by COVID-19 is not quantifiable due to a lack of studies
and the difficulty involved in the collection of large datasets. Segmentation is a
preferred technique to quantify and contour the COVID-19 region on the lungs
using computed tomography (CT) scan images. To address the dataset prob-
lem, we propose a deep neural network (DNN) model trained on a limited
dataset where features are selected using a region-specific approach. Specifi-
cally, we apply the Zernike moment (ZM) and gray level co-occurrence matrix
(GLCM) to extract the unique shape and texture features. The feature vectors
computed from these techniques enable segmentation that illustrates the
severity of the COVID-19 infection. The proposed algorithm was compared
with other existing state-of-the-art deep neural networks using the Radiopedia
and COVID-19 CT Segmentation datasets presented specificity, sensitivity, sen-
sitivity, mean absolute error (MAE), enhance-alignment measure (EM_φ), and
structure measure (S_m) of 0.942, 0.701, 0.082, 0.867, and 0.783, respectively.
The metrics demonstrate the performance of the model in quantifying the
COVID-19 infection with limited datasets.

KEYWORDS
artificial intelligence, computed tomography image, deep neural network, feature extraction,
limited training points, segmentation, Zernike moment

1 | INTRODUCTION

The novel coronavirus disease-2019 (COVID-19) is a lung
infection caused by Severe Acute Respiratory Syndrome
(SARS) and is transmitted from person to person easily.
The first human COVID-19 case has been reported in
Wuhan city, China in December 2019.[1] This disease was
traced during the sewage water study from Milan and
Turin in Italy before China identified the first case and

the related information was shared in a recent article.[2]
The World Health Organization (WHO) has reported
10.59 million COVID cases across 213 countries and terri-
tories, as of June 2020. A large percentage of the popula-
tion at present is affected by this disease. The recovery
and mortality rates are 54.22% and 4.91% approximately.
All over the world, researchers from biomedical depart-
ments are involved in an effort to find the effective vac-
cine for the COVID-19.[1] However, early detection is very

important for choosing the right type of treatment to control the disease and avoid the spread. AI has been crucial in helping and serving the automated diagnostics and public health surveillance to analyse the severity of the COVID-19 as well as in protecting the people from the virus spread.[3]

In the human body, the lungs are the major organs of the respiratory system. The disease in the lung will affect the entire air circulation system and may even result in death. The other major diseases that affect the lungs are lung cancer, tuberculosis, pneumonia.[4,5] Also, some microorganisms such as fungus, bacteria, and virus may cause illness which requires proper medical treatment to cure the infection/disease. Likewise, the infection caused due to the COVID-19 virus requires the same. The most commonly used COVID-19 tests are RT-PCR and antigen testing. Reverse transcription polymerase chain reaction (RT-PCR) is the standard COVID-19 test used to detect the viral nucleic acid that provides the genetic information of the virus. The test requires nasopharyngeal and throat swabs collected from the infected person. The duration of the test is too long and the result of RT-PCR is affected by sampling errors and low viral load.[6] The antigen test is designed to detect the spike protein that is responsible for facilitating the entry of virus into the human cell. Though the test is fast, its sensitivity is poor.

An alternate detection method for COVID-19 is based on medical imaging techniques that identify the infected location and infection rate. Medical imaging records the information in the lung region using chest radiography (X-rays) and Computed-Tomography (CT) scan.[7-10] Compared to the X-rays, a CT scan is mostly preferred for detecting the radiographic features of COVID-19 from the three-dimensional view of the lung's region.[11] Further, the CT scans can be used to give two-dimension views: axial view, coronal view, and sagittal view for the proper COVID infection detection. CT scan imaging of lungs gives good sensitivity to examine the COVID-19 infected region even before the medical symptoms occur.[12]

The spread of the infection due to COVID-19 in the lungs are identified through Ground glass opacification (GGO) and pulmonary consolidation phases. GGO is identified in the early stage of the infection whereas pulmonary consolidation indicates the final stage of the disease. These phases are observed from the CT scan images of the COVID cases.[13-15] The qualitative rate of the disease indicated in CT scans provides important information to guide against COVID-19.

Rajinikanth et al.[16] suggested that infection visibility analysis using CT scans gives better capability and reliability in the detection of the disease using the two-dimensional (coronal) view. The dataset used is obtained from the Radiopedia database.[17] The experimental analysis and the simulation of the model were performed in a MATLAB environment.

Accurate segmentation of radiographic features is important to quantify the disease in CT scan images. Segmentation of medical images needs marking manually by experts. As the number of COVID cases is increasing rapidly, it is difficult to label the infected region manually. Thus, automatic segmentation is needed from the CT scans to detect the infected region.[18] From the above consideration, our contributions are as following:

1. We propose a Deep Neural Network (DNN) model to detect and segment the COVID-19 from the axial view CT scan image from Radiopedia and COVID-19 CT Segmentation database. Our model provides better performance with limited number of training points.
2. Using the region-specific approach, instead of taking all the training points or images,[19] limited training points are chosen from the infected region as well as the background region of the CT scan image. Before selecting the training points, for each image, the shape and texture descriptors are extracted using Zernike moments and GLCM, respectively.
3. The performance of the proposed method is evaluated using the standard metrics for the different test images and evaluated metrics are compared with the existing models.[19-24]

The methodology needs enhancement techniques for CT scans of the COVID-19 dataset, feature extraction and appropriate classifier model to detect and segment the infected region. The rest of the paper is arranged as follows. Section 2 discusses the related works on the COVID-19 using AI techniques. Section 3 gives a detailed explanation and workflow of methodology and implementation. Sections 4 and 5 outline the experimental design, results and discussion. Section 6 provides the conclusion.

2 | RELATED WORKS ON THE COVID-19 USING AI

A comprehensive list of work for COVID-19 image-based AI techniques are found in Maga et al.[18] and Fu et al.[25] Developing a deep network gives more benefits for automatic and fast segmentation of the medical images.[26] Camouflaged object detection (COD) is introduced to identify the embedded object with their surroundings.[27] COD is beneficial in medical image applications such as lung infection segmentation. Ali et al.[28] proposed the integration approach consisting of unsupervised machine

learning (self-organizing map), dimensionality reduction (principal component analysis) and computational classification (Adam Deep Learning) to present a better classification performance. Particularly, the U-Net model works well in many of the segmentation tasks for medical images.[29] There are a few applications that have adopted U-Net for liver, heart, and multi-organ segmentation.[30-32] Chen et al.[33] has proposed a deep learning model for multiple regions auto segmentation for COVID-19 and used the aggregated residual transformations to extract the features from the CT image. Contrastive learning is proposed to train the encoder module, which provides the expressive feature information from the publicly available CT image COVID-19 dataset.[34] The Deep Learning model requires a lesser number of image samples for training the dataset to provide an automatic classification of COVID-19 infected images. Also, the authors concluded that the contrast learning method achieves better performance than ResNet-50. Narin et al.[35] have analysed the ROC result of three Convolutional Neural Network (CNN) models such as ResNet-50, InceptionV3, and Inception-ResNetV2 using a chest X-ray image. Shi et al.[36] have discussed pre-processing the COVID CT images using location-specific feature extraction and infection Size Aware Random Forest method (iSARF) to distinguish the infection level and then classification using random forest. The iSARF performed well by providing Sensitivity, Specificity and Accuracy of 0.907, 0.833 and 0.879, respectively. Yujin et al.[37] proposed, the patch-based CNN using a novel probabilistic Grad-CAM saliency map with a limited number of training points from the chest X-rays. Wang et al.[38] proposed an open-source CNN model called COVIDNet, which is trained and tested by the COVIDx dataset. It is useful to identify the COVID-19 infection from chest X-rays and has achieved a sensitivity of 80%.

Lung Infection Segmentation Deep Network (Inf-Net) used to detect the COVID-19 infection in the lungs, from the CT image slices where a parallel partial decoder is adapted to extract the high-level features.[19] Later implicit reverse attention and explicit edge attention are used to enhance the features. Charmaine et al.[39] have presented the deep learning model using the location-attention oriented approach to calculate the probability of COVID-19 infection region. Chest CT dataset is used for training the deep learning system.

A novel semi-supervised shallow learning model including Parallel Quantum-Inspired Self-supervised Network (PQIS-Net) with Fully-Connected (FC) layers for automatic segmentation of COVID-19 CT image is proposed in Konar et al.[40] The patch-based classification was applied to the segmented images for the diagnosis of COVID-19 using the two publicly available datasets. The efficiency (F1-score and AUC) of the PQIS-Net was compared with pre-trained convolutional based models.

The truncated VGG-19 model was proposed to analyse the COVID-19 CT scans. TheVGG-16 model was used to extract features from the CT images using fine-tuning.[41] The feature selection was achieved through Principal Component Analysis (PCA) and the classification was done using four different classifier methods such as deep convolutional neural network (CNN), Bagging Ensemble with support vector machine (SVM), Extreme Learning Machine (ELM) and Online sequential ELM on 208 test images. From the different classifiers, Bagging Ensemble with SVM have achieved the following: an accuracy of 95.7%, AUC0.958 and F1-score 95.3%.

The three-phase COVID-19 CT image detection model is introduced in Ahuja et al.[42] The modules are as follows: (a) Augmentation using stationary wavelets; (b) COVID-19 detection using pre-trained CNN model such as ResNet18, ResNet50, ResNet101, and SqueezeNet; (c) Abnormality localization in CT scan images. The experimental analysis showed that the pre-trained ResNet18 transfer learning model has given a better classification accuracy of 99.82% for training and 99.4% for testing.

A pipeline model is presented in Dey et al.[43] with several sub-modules to classify the segmented region from the COVID-19 images. First, the COVID-19 CT images are segmented using the using Social-Group-Optimization and Kapur's Entropy thresholding, followed by K-means clustering and morphology-based segmentation. Next, a classification module is introduced to classify the segmented region. Here PCA based fusion technique is used to fuse the features and then fused features are trained with different classifiers such as Random Forest, k-Nearest Neighbours (KNN), Support Vector Machine with Radial Basis Function, and Decision Tree. Experimental results indicated an accuracy of 91% and 87% with Morphology-based segmentation and kNN classifier. A summary of the various techniques available for COVID-19 screening is presented in Table 1.

Wang et al.[44] have proposed the weakly-supervised deep learning (UNet and DeCoVNet) using 3-Dimensional CT volumes to identify COVID-19.The literature indicates the successful evaluation of COVID-19 infection using deep learning and traditional machine learning methods by various researchers. But the study also clarifies that some of the regions are not evaluated in the classifier layer due to lack of training data samples and inefficient feature representation. The detailed design flow of the proposed framework is explained in the next section.

TABLE 1 Literature summary of related techniques for COVID-19 screening

Reference	Technique	Dataset	Findings
33	Aggregated residual transformations to learn a robust and expressive feature representation.	110 COVID-19 CT image of size 512×512 are collected from 60 patients.	The model gives 0.95 precision and 0.89 accuracy. 12.7% of accuracy and 14.5% of precision are improvement with U-Net model.
34	Contrastive learning method.	MedSeg dataset:110 CT COVID-19 images. COVID-19 CT dataset: 349 CT COVID-19 images 397 Non-COVID images.	ResNet-50 performance gives accuracy of 0.868, recall of 0.872, AUC of 0.931.
35	ResNet50, InceptionV3 and InceptionResNetV2.	Chest X-ray 100 images: 50 non-infected image and 50 infected images. All images are 224×224 pixel size.	ResNet50 model provides good classification performance with 98% accuracy than other two proposed models (97% accuracy for InceptionV3 and 87% accuracy for Inception-ResNetV2).
36	An infection Size Aware Random Forest method (iSARF).	COVID-19 images are taken from 1658 patients.	iSARF yielded sensitivity of 0.907, specificity of 0.833, and accuracy of 0.879.
37	A patch-based convolutional neural network approach.	COVID-19 dataset: 180 infected image of size 224×224.	Global approach gives 70.7% of accuracy, 60.6% of precision, 60.1% of recall, 59% of F1-score. Local approach gives 88.9% of accuracy, 83.4% of precision, 85.9% of recall, 84.4% of F1-score.
19	COVID-19 Lung Infection Segmentation Deep Network (Inf-Net) with a parallel partial decoder method is to aggregate the high-level features.	Dataset consists of 100 axial CT images from different COVID-19 patients. 45 CT images randomly selected as training, 5 CT images for validation, and the remaining 50 images for testing.	Inf-Net achieves 0.692 of sensitivity, 0.943 of specificity, 0.725 of sensitivity, 0.960 of specificity.
40	Semi-supervised shallow neural network model using Parallel self-supervised neural network model (PQIS-Net).	Variable size 2482 lung CT scans: 1252 COVID-19 infected CT image and 1230 Non-infected COVID image. 20 labelled COVID-19 CT image of size 512×512 which includes infection region masks, lung masks (left and right) and lung-infection pair masks.	Accuracy, precision, f1 score, and AUC are 0.931, 0.890, 0.826 and 0.982, respectively.
41	Truncated VGG-19 with fine-tuning method. Classification is done with four different classifier method such as deep convolutional neural network (CNN), Bagging Ensemble with support vector machine (SVM), Extreme Learning Machine (ELM) and Online sequential ELM.	CT Scan Dataset: 344 COVID-19 images, 358 Non-COVID-19 images.	Best performing classifier Bagging Ensemble with SVM achieves 95.7% accuracy, 95.8% precision, 0.958 AUC and 95.3% F1-score.
42	Transfer Learning method with argumentation: ResNet18, ResNet 50, ResNet101, SqueezeNet.	COVID-19:349 CT scan images and Non-COVID-19 CT: 397 CT scan images.	ResNet18 achieves the best performance model. Training accuracy is 99.82%, validation accuracy is 97.32% and testing accuracy is 99.4%.

TABLE 1 (Continued)

Reference	Technique	Dataset	Findings
43	Social-group-optimization with Kapur's entropy thresholding and then followed by k-means clustering and morphology-based segmentation.	COVID-19 dataset consists of 400 numbers of greyscale lung CT image (200 Non-infected image and 200 infected images).	Morphology-based segmentation and KNN gives more than 91% and 87% accuracy.

FIGURE 1 Block diagram for the flow of binary classification for computed tomography image
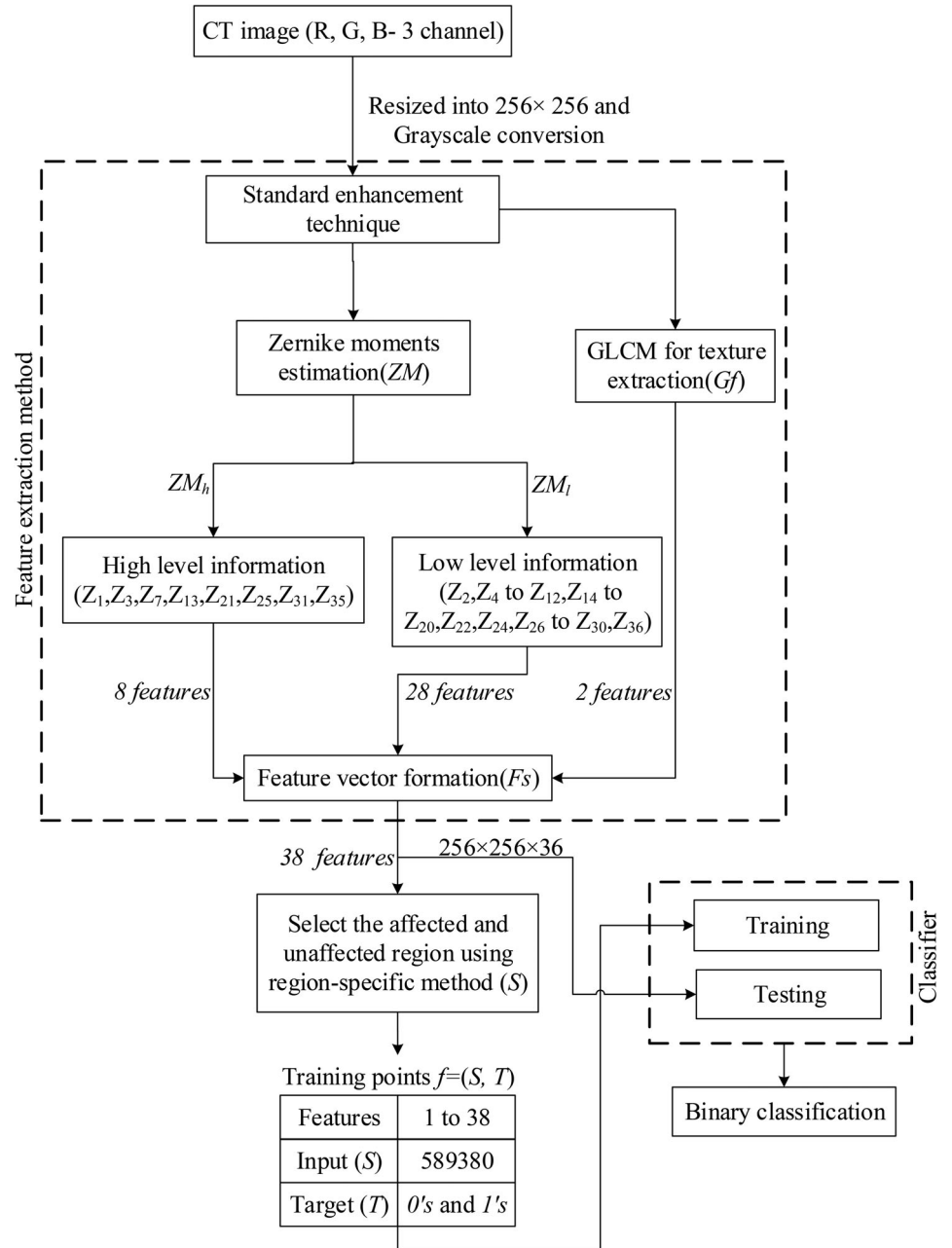
3 | METHODOLOGY

The Datasets used in the proposed framework consists of 80 images of size 512×512 pixels (3 channels). The segmentation process of each CT image involves (a) pre-processing, (b) extracting the robust and sensitive features, (c) training the region-specific region, and (d) segmenting the region of interest (infected region) from the background. The block diagram of the proposed methodology is as shown in Figure 1.

3.1 | Pre-processing and feature extraction methods

The pre-processing is based on the standard enhancement approach[45] called adjust image intensity values (imadjust). This approach describes the expressive information from the COVID-19 CT image by adjusting the contrast of the GGO and pulmonary consolidation region. The infected pixels are brightened visually as shown in Figure 2. Also, the Contrast Improvement Index (CII)[46] metrics are analysed to highlight the importance of pre-processing on the COVID dataset. The image quality (pixel intensity) is improved by 30% to 40% when compared to the original CT image as shown in Table 2. CII is computed based on Equations (1) and (2).

$$CII = \frac{C_{enhanced}}{C_{original}} \qquad (1)$$

$$C = \frac{I_f - I_b}{I_f + I_b} \qquad (2)$$

where $C_{enhanced}$ and $C_{original}$ are the contrast of the enhanced and original images respectively, C is the average value of the contrast in the enhanced or original image. I_f and I_b are the mean value for foreground and background contrast of the image. A higher CII value indicates improvement in image quality.

The enhanced images present significant details for extracting the features. The extraction of features is a description of each pixel of the CT image that contains descriptive information in the form of a stack of the vectors (feature vector). The feature vector helps to find out the output class (infected COVID-19 region and background region). The feature vector considered for the work is the integration of shape and texture features. The texture features are derived from the gray level co-occurrence matrix and shape descriptor features are extracted from the Zernike moment.
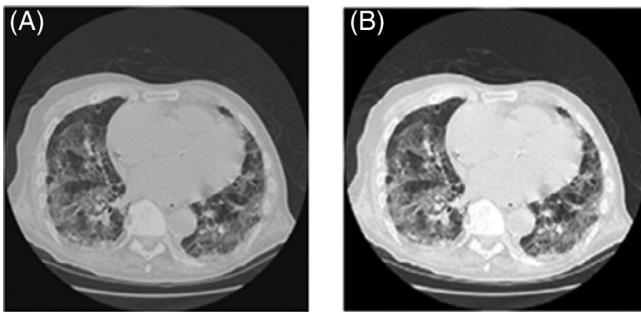
FIGURE 2 A, Original computed tomography image; B, Image after enhancement

3.2 | Zernike moment (ZM) based shape descriptor features and gray level co-occurrence matrix (GLCM) for texture features

Based on the infection stage/level, the infected region of COVID-19 CT images has dissimilarity in shapes, sizes and structures of the geometry. The shape descriptor is good to define the infected pixel information range from its background region. The shape parameters can be extracted using descriptor based on Zernike and Hu moments.[47,48] For better representation, the shape descriptor should possess the properties such as (a) provide low redundancy and higher discrimination ability. (b) Rotation, scale and translation invariance, and (c) present coarse to finer detail representation.

Hu moments require extensive computation power to calculate higher order moments. Therefore, we choose the ZM based shape descriptor for representing shape features of the COVID-19 lungs image. The ZM is calculated for every pixel of the pre-processed image by sliding a window of size 5×5 with a stride of 1 along the column and row. For the computation of ZMs, different polynomial orders (n) and the corresponding repetition factor (m) were considered for the function $I(x, y)$. Next, we compute the Zernike moments, by projecting $I(x, y)$ onto the set of complex Zernike polynomials as given in Equation (3).

$$Z_{nm} = \frac{m+1}{\pi} \sum_x \sum_y I(x,y) P_{n,m}^*(x,y), x^2 + y^2 \leq 1 \qquad (3)$$

where, $P_{nm}(x, y)$ is Zernike polynomial, $I(x, y)$ is the image of size 5×5 pixels, n is a non-negative integer, m is an integer such that $0 \leq |m| \leq n$, and $n - |m|$ is even. The function $P_{nm}^*(x, y)$ is the complex conjugate of the orthogonal basis function $P_{nm}(x, y)$ defined as:

$$P_{nm}(x,y) = P_{n,m}(r,\theta) = R_{n,m}(r)e^{jm\theta} \qquad (4)$$

where, $r = \sqrt{x^2 + y^2}$ and $\theta = \tan^{-1}\frac{y}{x}$, with $0 \leq \theta \leq 2\pi$, $j = \sqrt{-1}$.

The radial polynomial $R_{nm}(r)$ is defined as:

$$R_{nm}(r) = \sum_{q=0}^{n-|m|} (-1)^q \frac{(2n+1-q)!}{q!(n-|m|-s)!(n+|m|+1-s)!} r^{n-q} \qquad (5)$$

The Zernike moment obtained from the Equation (3) is a complex number represented as shown in Equation (6):

TABLE 2 CII values for the pre-processed COVID-19 CT images

Number of images	Contrast for the original CT image	Contrast for the pre-processed image	CII for the pre-processed image
1	0.5984	0.8166	1.3646
2	0.4264	0.7348	1.7234
3	0.1650	0.3279	1.9872
4	0.4808	0.6698	1.3931
5	0.4846	0.9497	1.9599

Note: Default CII value for the original image is 1.000.

$$Z_{nm} = R_{ZM} + jI_{ZM} \tag{6}$$

$$|Z_{nm}| = \sqrt{R_{ZM}^2 + I_{ZM}^2} \tag{7}$$

Equation (7) is the magnitude of the Zernike moment $|Z_{nm}|$ that denotes the ZM shape descriptor for specified n and m. Thus, a total number of $36|Z_{nm}|$ were obtained with n ranging over 1 to 10 and matching combinations of n, m with m such that $0 \le |m| \le n$, and $n - |m|$ is even for every pixel. The magnitude of ZM calculated for every pixel to find the ZM feature maps are as displayed in Figure 3.

From the observation of Zernike features, we find that all the 36 different $|Z_{nm}|$ provide unique information about the infected and background region. To analyse the Zernike robust feature (ZM_h) and Zernike sensitive feature (ZM_l), each point is taken from both the regions with 36 ZM for different n, m and are plotted as indicated in Figure 4 and the separation of the ZM_h and ZM_l is indicated in Equations (8) and (9).

$$ZM_h = \{|Z_{0,0}|, |Z_{2,0}|, |Z_{4,0}|, |Z_{6,0}|, |Z_{8,8}|, |Z_{8,8}|, |Z_{10,0}|, |Z_{10,8}|\} \tag{8}$$

$$ZM_l = \{|Z_{1,1}|, |Z_{2,12}| \text{ to } |Z_{5,5}|, |Z_{6,2}| \text{ to } \\ |Z_{7,7}|, |Z_{8,2}|, |Z_{8,6}|, |Z_{9,1}| \text{ to } |Z_{9,9}|, |Z_{10,10}|\} \tag{9}$$

From the different combinations of ZM feature extraction image, we plot the intensity range for the set of pixels for the 1st image of the dataset. The graphical plot presented in Figure 5 represents number of pixels in x-axis and ZM features in y-axis. The above observation indicates that the Zernike features from background information are dense at the bottom region and the infected region information mostly occurs above a specific threshold.

To make the binary classification more accurate, we add an extra texture feature computed using the GLCM.

The texture feature is more important in medical imaging (CT image) for extracting the GGO and pulmonary consolidation features.[49] GLCM consists of the relationship of the different angles between image pixels. If the gray level co-occurrence matrix obtained from an image be denoted as $q = [q(r, s | d, \theta)]$. In this point, GLCM is used to estimate the features of rth pixel frequency with the feature of sth pixel frequency by the length ($d = 1$) and direction ($\theta = 0$). The computation of the texture feature follows the same procedure as that of ZM, that is, computing the texture features for every pixel by moving a window of 5×5 with a stride of 1 over the pre-processed image.

From GLCM, texture features such as entropy, variance, contrast, correlation, dissimilarity, energy and homogeneity are calculated. Among all these features, we considered the GLCM variance and contrast as shown in Equations (10) and (11). These two features are combined to form GLCM features Gf that gives proper discrimination of infected region and background region from the CT image.

$$\text{Contrast} = \sum_{r=0}^{N-1} \sum_{s=0}^{N-1} q(r,s)(r-s)^2 \tag{10}$$

$$\text{Variance} = \sum_{r=0}^{N-1} \sum_{S=0}^{N-1} q(r,s)(r-\mu)^2 \tag{11}$$

$$Gf_i = [\text{variance}_i, \text{contrast}_i] \tag{12}$$

where, i refers to the number of training images, 'μ' denotes mean and 'N' denotes the image size, Gf_i refers the GLCM features (2 features). The intensity for the GLCM feature images (contrast and variance) is plotted as shown in Figure 6. It represents number of pixels in x-axis and GLCM features in y-axis. Here the red points denote infected region and green points denote background respectively. The observation gives clear distinction between the background region and infected region.
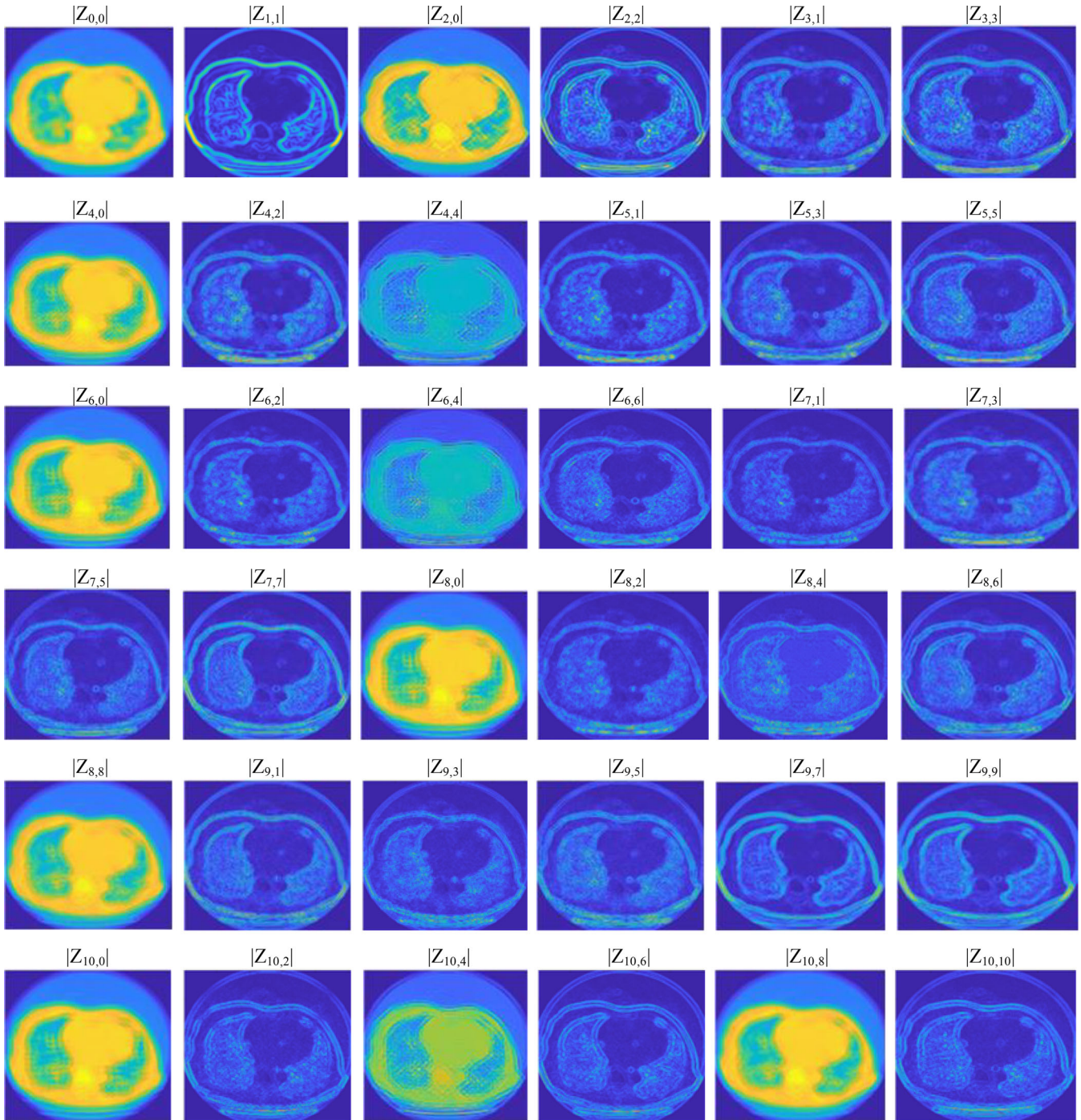
FIGURE 3 Zernike moment features with different combination of n and m [Color figure can be viewed at wileyonlinelibrary.com]

Now the shape features computed from the ZM and the texture features obtained from the GLCM are integrated to form a feature vector of 38 features as indicated in Equation (13).

$$Fs_i = \{ZM_{h,i}, ZM_{l,i}, Gf_i\} \qquad (13)$$

The next section discusses about the formation of training dataset from Fs_i features.

4 | EXPERIMENTAL DESIGN

4.1 | COVID-19 dataset

COVID-19 dataset (axial view) is collected from References 17 and 50 to evaluate our implementation. The previous implementations have used X-rays and CT image datasets for segmentation and classification of the infected COVID regions.[50,51] The COVID-19 CT dataset

used in this work contains 80 images of size 512×512 pixels (3 channel) along with the ground truth image (binary image). Ground truth images are useful for the identification of targets from the input dataset. From the dataset, 30 images are selected for training the network, the remaining 50 images are used for testing the trained network for performance evaluation.

4.2 | Training and testing dataset

COVID-19 CT images contain additional background pixels (without infected region) other than infected pixels. Using the whole image as the training data gives testing results with less accuracy that reduce the overall performance. It should be able to classify the background region from the infected region during the testing. Hence, a balanced training dataset is required for the both infected region If and background Bf pixels along with all the features from the training images.

The works[19,33,35] have used the whole image as training data. The training result has delivered a performance greater than 90% as compared to the overall testing accuracy. However, as explained in Fan et al.[19] the overall sensitivity of the test dataset is reduced by not selecting the true positive target properly (infected pixels). Hence, we chose properly infected pixels and background pixels from the extracted images.

4.3 | Region-specific for formation of training points

The dataset of training points is manually selected from the infected and background pixels from the 30 training images (after pre-processing and Fs_i) by using the region-specific approach. The set of features (S_i) is formed by marking the specific region (infection region (If) and background region (Bf)) for each feature from the training image dataset as shown in Figure 7.

We chose the If region with 62 pixels and Bf region as 456 pixels. This feature helps to classify the image pixels as infected or not from COVID-19. The algorithm for region-specific approach is explained below:

```
Input: Fs_i = {Z_{h,i}, Z_{l,i}, Gf_i}; //i refers to
number of training images
Output: If; //Infection region feature.
Bf; // Background region features from Fs_i.
S_i; // Input feature set for training.
Start
    Initialize S_{i,k} = [];
```
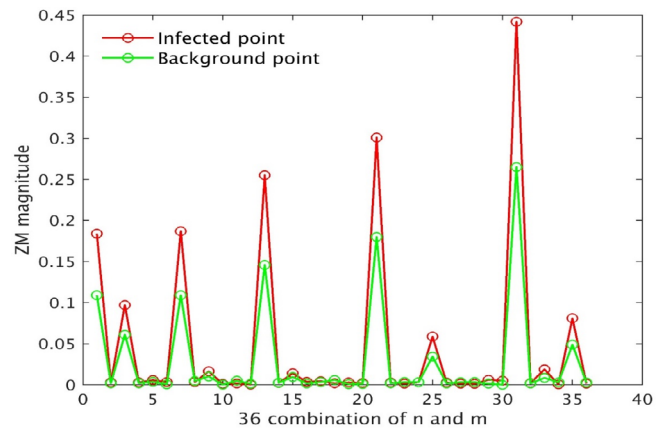

FIGURE 4 Experimental analysis of low-level and high-level feature from Zernike moment (total 36 Zernike feature, red – infected point, green – background point) [Color figure can be viewed at wileyonlinelibrary.com]

```
Initialize j = 518; // Number of training
point in each features
    for i = 1 to 30 // Number of training
images
    for m = 1 to 38 // Number of features
    If = Fs_i[j-517:j-456] // Vectorized
    format of Infected region points
    Bf = Fs_i[j-517:j-62] // Vectorized
    format of Background region points
    M_i = [If, Bf];
    S_i = [S_i, M_i] // Training points
    features from 30 training image
    end
    end
Stop
```

Region selection is helpful in getting balanced image information between background and infected pixels. The selected pixels are rearranged into the vector points for training. The total training points consist of 589 380 pixels with 519 371 background pixels and 70 009 infected pixels. In the total training points, each image contains 19 646 points of 38 features. A large number of points are taken into COVID infected region to reduce the misclassification of testing image features.

During the training process, the feature vector is labelled using the ground truth as shown with T_i as the target vector represented in binary form.

$$f_{i,k} = [S_i, T_i] \quad (14)$$

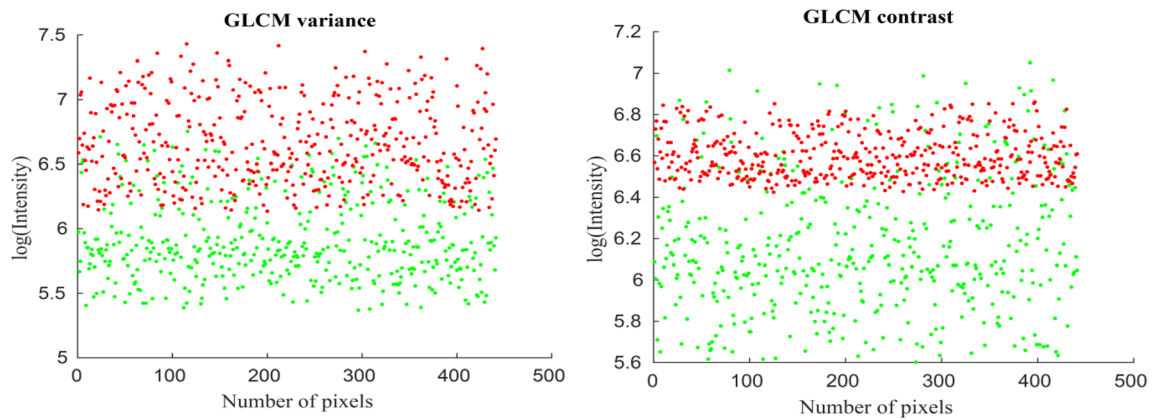$$T_i = [\text{class} 1_i, \text{class} 2_i] \quad (15)$$
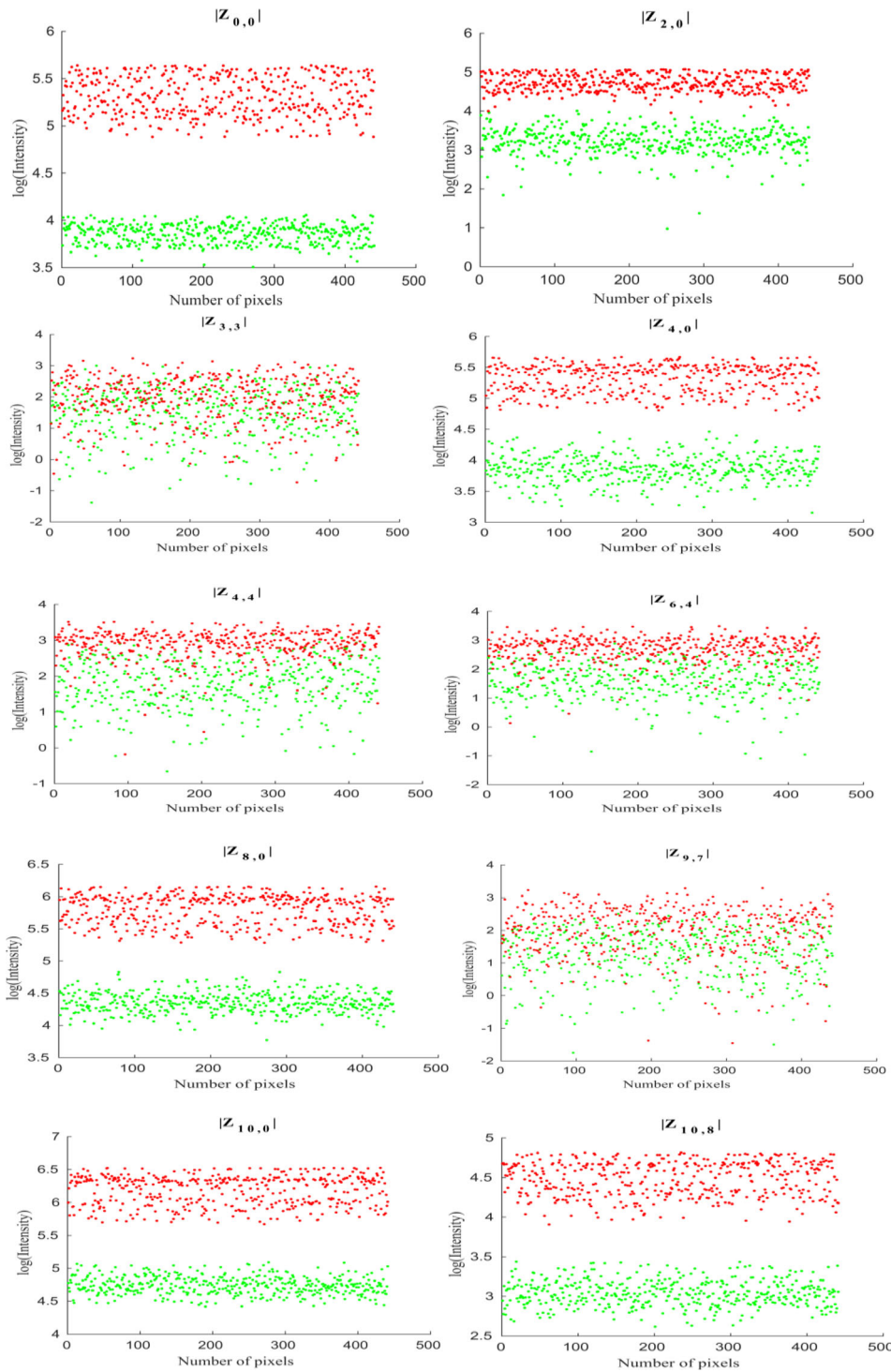
FIGURE 5 Plots for Zernike moment features from specific region of the training image (Red points represent background and Green points represents infected pixels) [Color figure can be viewed at wileyonlinelibrary.com]

here $f_{i,k}$ is the formed training dataset, S is the input features, T is target for the corresponding input features, i refers to the number of training images, k refers to the number of training points ($k = 1$ to $589\,380$ points from all features of Fs_i). From the T, class $1 = $ '1' and *class $2 = $ '0'* represents an infected pixel, *class $1 = $ '0'* and *class $2 = $ '1'* represents the background pixel.

Furthermore, to select a suitable classifier for the proposed work, the performance of different classifiers such

as Support Vector Machine (SVM), Deep Neural Network (DNN), Decision Tree (DT), Logistic Regression (LR) and Gaussian Naive Bayes (GNB) are compared using the training data points. The performance of classifiers is evaluated from the confusion matrix using the standard metrics: Area Under Curve (AUC), Specificity, Sensitivity and Accuracy as presented in Table 3. From Table 3, it is observed that the DNN classifier outperforms the other classifier in terms of Accuracy, AUC and Sensitivity for

FIGURE 6 Plots for GLCM features from specific region of the training image (Red points represents background and Green points represents infected pixels) [Color figure can be viewed at wileyonlinelibrary.com]

(A)

(B)
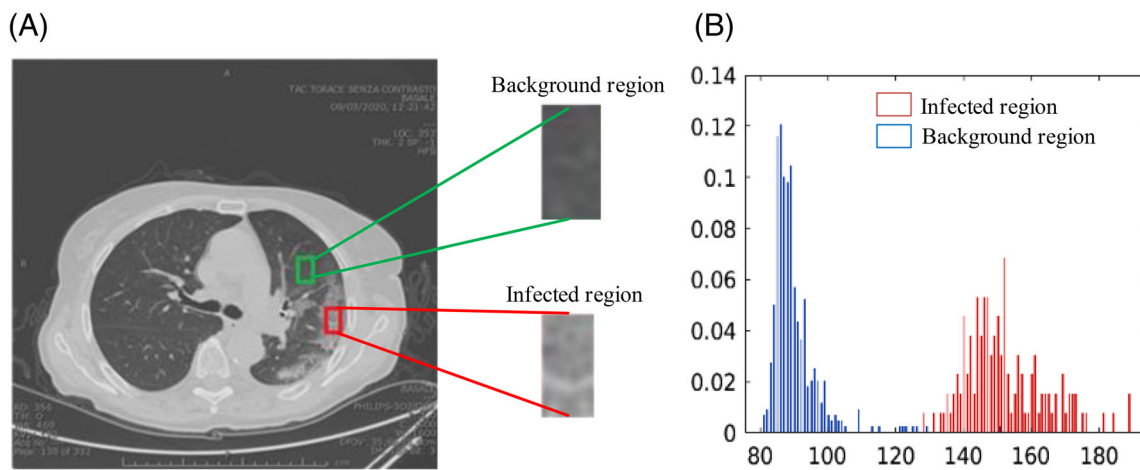
FIGURE 7 A, Specific region of background (*Bf*) and infection features (*If*) are separated out for training. B, Histogram plot for *Bf* (blue colour) and *If* (brown colour) [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 3 Evaluation of training points metrics for different classifiers

Different classifiers	AUC	Sensitivity	Specificity	Accuracy (%)
SVM	0.8723	0.5549	0.8988	88.6
GNB	0.9107	0.4025	0.9735	83.3
LR	0.8187	0.5211	0.8950	88.2
DT	0.94	0.756	0.9593	93
DNN	**0.9427**	**0.7678**	**0.9285**	**93.8**

Note: The significance of the bold numbers are highlighted results of the proposed algorithm.

the given training points and corresponding confusion matrix for the $f_{i,k}$ for DNN has been presented in Table 4. AUC is estimated from the ROC graph for the different classifiers that are as plotted in Figure 8.

4.4 | Binary classification using DNN

The set of ZM and GLCM features extracted using the methodology as explained in Section 3 form the feature set. The corresponding ground truth data is considered as a target. A supervised learning method is implemented by training a Deep Neural Network (DNN). Our architecture is identical to the basic feed-forward network with multiple hidden layers. This work considers the architecture where the size of the input layer is 38 neurons (38 features), three hidden layers with 58 neurons per layer and binary classification output layer as shown in Figure 9.

The total number of learnable parameters (weights and biases) in the neural network is 9224. The first layer

TABLE 4 Confusion matrix of training points for DNN classifier

		T_i		
		Infected pixels	Background pixels	Accuracy
Training output	Infected pixels	46 469 (7.9%)	12 684 (2.2%)	93.8%
	Background pixels	23 540 (4.0%)	506 687 (85.9%)	
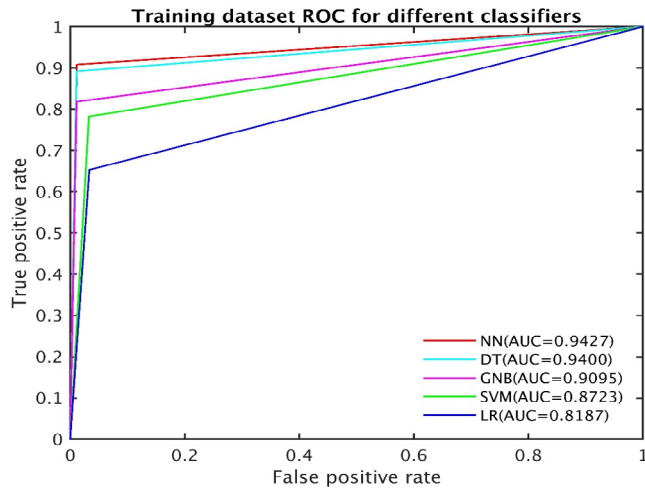
FIGURE 8 ROC for different classifiers [Color figure can be viewed at wileyonlinelibrary.com]
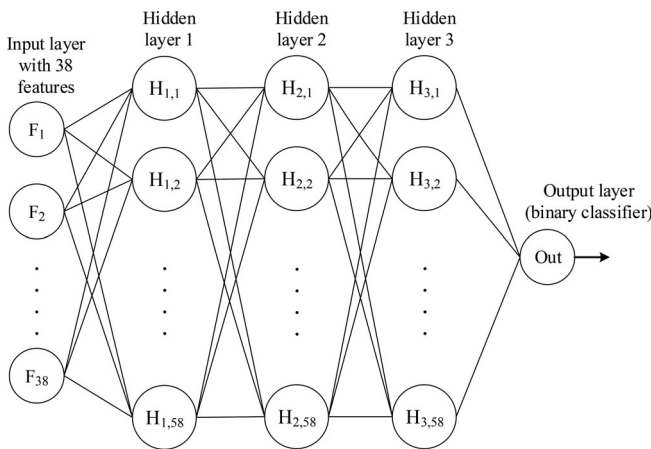
FIGURE 9 Deep neural network (DNN) architecture

has 2262 (38 × 58 + 58), the second layer 3422 (58 × 58 + 58), the third layer 3422 (58 × 58 + 58) and the final layer with 118 (58 × 2 + 1). The loss function (cross-entropy loss) converges by modifying the parameters using Scaled Conjugate Gradient (SCG) training function. The three hidden layers are followed with the activation function called tan-sigmoid and the output layer involves the softmax operation that gives the probability of two outputs. The probability of the first neuron in the output layer represents the lungs infected region and the second
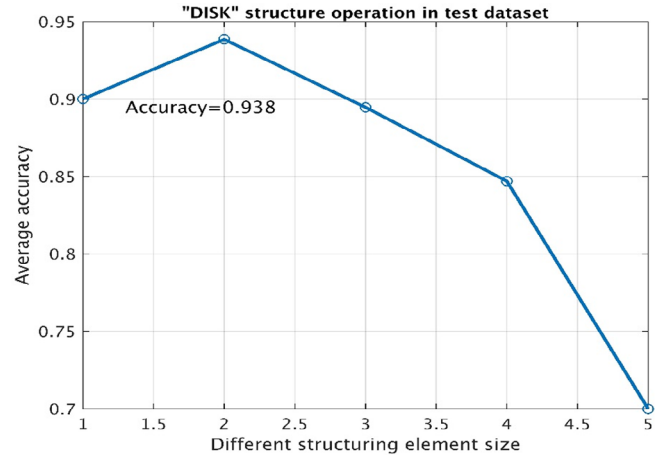
FIGURE 10 Test dataset plot for morphological structuring element vs average accuracy [Color figure can be viewed at wileyonlinelibrary.com]

neuron is complementary of the first probability called background region. The training required 8 minutes and 49 seconds to get the best error performance of 0.1367 at 100 epochs.

5 | RESULTS AND DISCUSSION

The trained DNN classifier is evaluated for the test image (50 images) from the COVID-19 dataset. After the pre-processing and Fs_i, the test vector is formed for each test image of size 38 × 256 × 256. All the experimental analysis is done in MATLAB R2020a run on Intel Corei3-2330 M CPU. The feature generation process takes around 250 minutes (about 4 hours) for the test set of 50 images and testing time for each test image is around 4 to 7 minutes.

The challenges of automatic segmentation from the CT image are.

1. CT Images are varying in terms of texture, shape, and position across different cases and consolidation is small, which may result in a false negative output from a whole CT image.
2. Due to the low contrast and blurriness within the GGO of the infection area, there is a significant

difficulty in the identification of the contours in the CT images.

3. Artificial intelligence (AI) model requires a greater number of training data to train the model and more human/ specialist resources are required to collect and annotate the ground truths of the COVID-19 dataset.

As explained in Section 3, the feature vector of 38 features is given to the trained DNN model as input. The output of the classifier gives the probability score for two classes (infected and background). This probability score decides and falls on particular class 1 or class 0. The probability vector output is reshaped into the image size 256×256. Further, the image of size 256×256 is
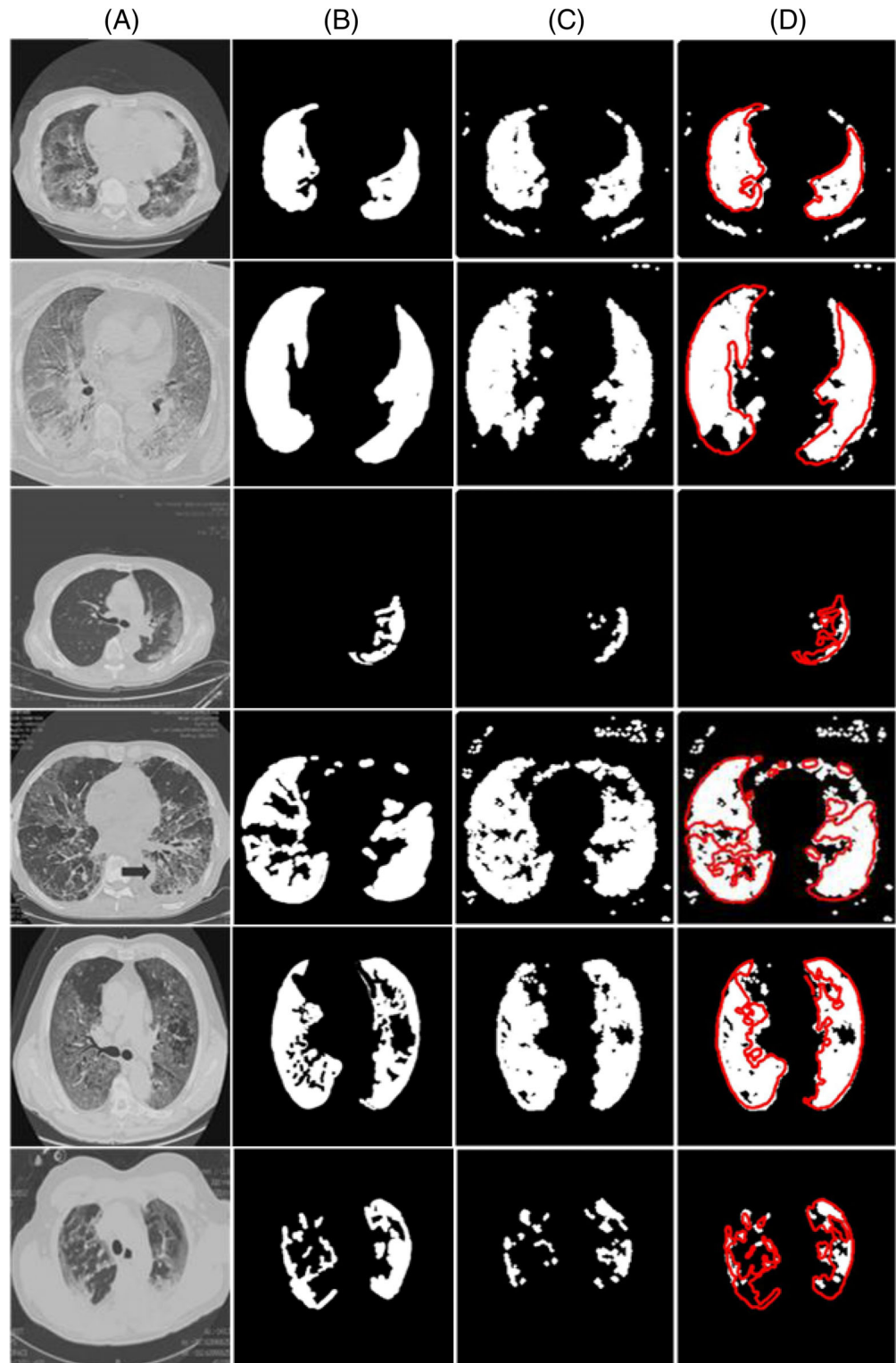
FIGURE 11 Segmentation results for the COVID-19 dataset. A, Column shows computed tomography axial view input image; B, Column shows ground truth image; C, Column shows the output from deep neural network model and D, Column shows overlay the GT contour on output image for visual understanding [Color figure can be viewed at wileyonlinelibrary.com]

TABLE 5 Performance evaluation for the test dataset (36 ZM features)

| Number of test images | Accuracy (A) = $\frac{TP+TN}{TP+TN+FP+FN}$ | IOU score | | Mean IOU = $\frac{C_0+C_1}{2}$ | Sensitivity or Recall (R) = $\frac{TP}{TP+FN}$ | Specificity = $\frac{TN}{TN+FP}$ | F1 score = $2\times\frac{P\times R}{P+R}$ | Precision (P) = $\frac{TP}{TP+FP}$ |
		Class 0 (C_0) = $\frac{TN}{TN+FP+FN}$	Class 1 (C_1) = $\frac{TP}{TP+FP+FN}$					
1	83.5	0.393	0.815	0.604	0.414	0.98	0.564	0.884
2	85.6	0.337	0.843	0.59	0.381	0.969	0.505	0.747
3	88.8	0.459	0.936	0.697	0.490	0.992	0.629	0.879
4	91.5	0.166	0.928	0.547	0.177	0.994	0.285	0.736
5	89.3	0.333	0.925	0.629	0.371	0.987	0.5	0.765
6	82.2	0.382	0.881	0.631	0.453	0.967	0.552	0.708
7	97.7	0.058	0.895	0.476	0.060	0.99	0.110	0.583
8	86.3	0.086	0.872	0.479	0.101	0.977	0.16	0.375
9	83	0.66	0.922	0.791	0.711	0.984	0.600	0.915
10	83.3	0.238	0.824	0.531	0.285	0.955	0.385	0.590

TABLE 6 Confusion matrix for test image 1

| | | GT | | |
		Infected pixels	Background pixels	Accuracy
Testing output	Infected pixels	14 025 (21.4%)	2228 (3.4%)	92.30%
	Background pixels	2818 (4.3%)	46 465 (70.9%)	

compared with the ground truth image to detect and segment the COVID-19 regions and also evaluate the performance of the model.

To make the output binary image effective, the morphological operation is applied to the output binary image. Different morphological structuring element produces different test image results. Hence, we find the suitable structuring element with a proper size for all the test images to achieve the overall average accuracy and selecting the structuring element that gives high accuracy. The average accuracy vs morphological structuring element graph for COVID-19 CT image test data is shown in Figure 10. From the plot, it is determined that "imclose" disk structuring element of size '2' reaches the maximum accuracy. The test dataset results are displayed in Figure 11. In Figure 11d column, GT contours are overlaid on the output image which gives visual understanding of the non-distinguishable pixels. This arises due to the poor lightening between infected pixels and background pixels from the original image as shown in Figure 11 (3rd-row image) and also the thin occurrence of GGO and consolidation in Figure 11 (6th-row image). From the ablation study Table 5, 36 ZM shape descriptor training feature points alone do not achieve better results in Sensitivity and Mean IOU. Shape descriptor features do not capture the proper infected region (True Positive) information from the test image. so that, 2 GLCM texture features are added with 36 ZM features and trained the DNN model with 38 features. Evaluated metrics achieve better results with the combination of ZM and GLCM features (38 features) compared to the 36 features results. We have analysed each test image performance by standard metrics[39] such as Accuracy (A), Sensitivity or Recall (R), Specificity, F1 score, Dice index, Precision (P), Mean IOU and score using the confusion matrix as represented in Table 6.

Additionally, three more important metrics are Enhance-alignment Measure (EM_φ), Mean Absolute Error (MAE)[38] and Structure measure (S_m)[52] are used for performance evaluation. Enhance-alignment Measure (EM_φ) is a newly proposed metric. It is useful to estimate the global and local similarities of two binary images as presented in Equation (16). where, W represents width, H represents the height of GT, F represent the final predicted output, (x, y) represents each pixel coordinate of GT and F. φ denotes the enhanced alignment matrix.

$$EM_\varphi = \frac{1}{W\times H}\sum_{x}^{W}\sum_{y}^{H}\varphi(F(x,y),GT(x,y)) \quad (16)$$

Mean Absolute Error (MAE), measures the error between the F and GT of each pixel in the image as

TABLE 7 Performance evaluation for the random 10 test image from COVID-19 dataset (36 ZM and 2 GLCM features)

Number of test images	Accuracy (A) $= \frac{TP+TN}{TP+TN+FP+FN}$	IOU score Class 0 (C_0) $= \frac{TN}{TN+FP+FN}$	Class 1 (C_1) $= \frac{TP}{TP+FP+FN}$	Mean IOU $= \frac{C_0+C_1}{2}$	Sensitivity or Recall (R) $= \frac{TP}{TP+FN}$	Specificity $= \frac{TN}{TN+FP}$	F1 score $= 2 \times \frac{P \times R}{P+R}$	Precision (P) $= \frac{TP}{TP+FP}$	Dice index	MAE	EM_φ	S_m
1	0.9230	0.90	0.73	0.815	0.832	0.95	0.847	0.86	0.8028	0.1003	0.9033	0.8363
2	0.9237	0.907	0.68	0.793	0.876	0.93	0.815	0.76	0.7369	0.0945	0.9277	0.8431
3	0.9610	0.95	0.66	0.805	0.723	0.98	0.795	0.88	0.8121	0.1025	0.8512	0.8017
4	0.9447	0.94	0.54	0.74	0.835	0.95	0.705	0.61	0.7923	0.0678	0.7778	0.8411
5	0.9553	0.95	0.56	0.75	0.600	0.99	0.720	0.906	0.6479	0.0863	0.8122	0.7501
6	0.9579	0.95	0.74	0.845	0.882	0.96	0.851	0.82	0.7488	0.0703	0.9366	0.8416
7	0.9788	0.97	0.3	0.62	0.330	0.98	0.432	0.66	0.673	0.0394	0.9646	0.6475
8	0.9394	0.938	0.31	0.624	0.346	0.98	0.473	0.81	0.6366	0.0808	0.7627	0.6103
9	0.8771	0.848	0.61	0.729	0.828	0.89	0.759	0.70	0.6665	0.1624	0.8422	0.7391
10	0.9248	0.910	0.67	0.79	0.852	0.94	0.704	0.76	0.6945	0.1128	0.8962	0.8212

indicated in Equation (17). Structure measure (S_m), measures the structural similarity between the GT mask and prediction map as represented in Equation (18). Where $\alpha \in [0, 1]$, the default setting $\alpha = 0.5$ represents the balance factor between region-aware similarity (S_r) and object-aware similarity (S_o). The performance of each test image is evaluated by metrics as listed in Table 7.

$$MAE = \frac{1}{W \times H} \sum_{x}^{W} \sum_{y}^{H} |F(x,y) - GT(x,y)| \quad (17)$$

$$S_m = \alpha * S_o + (1-\alpha) * S_r \quad (18)$$

We have compared our method with some of the existing works that employ deep network for the segmentation of the COVID-19 infected region from the CT image dataset as presented in Table 8. In Ronneberger et al,[20] standard U-Net model is evaluated using 45 training images. The performance of the outcome is low in terms of sensitivity, specificity, MAE, Dice and EM_φ. The modernized version of the U-Net model is described in References 21–24 which resulted in a poor evaluation of the Dice index, sensitivity compared to our developed method. In our method, MAE is quite higher than the Semi-Inf-Net and Inf-Net model reported in Fan et al.[19] EM_φ, S_m and Sensitivity are moderately 3%, 2.12%, 3.3% lesser compared to Semi-Inf-Net model. However, the overall average comparison of our test image result is much improved and superior with a limited number of training points as shown in Table 9. Also, Table 9 shows a performance improvement with our method in each of the metrics compared to other existing works. The proposed model improved the sensitivity by 31.2%, 10%, 6.5%, 18%, 4.3% and 1.3% compared with U-Net model, Attention-UNet, Gated-UNet, Dense-UNet, U-Net++, and Inf-Net respectively. Likewise, the percentage improvements of other metrics in relation to other state-of-the-art is presented in Table 9.

TABLE 8 Comparative performance evaluation of other existing methods

Existing reference	Methods	Sensitivity	Specificity	Dice	MAE	EM_φ	S_m
Ronneberger et al[20]	U-Net[a]	0.534	0.858	0.439	0.186	0.625	0.622
Oktay et al[21]	Attention-UNet[a]	0.637	0.744	0.583	0.112	0.739	0.744
Schlemper et al[22]	Gated-UNet[a]	0.658	0.725	0.623	0.102	0.814	0.725
Li et al[23]	Dense-UNet[a]	0.594	0.655	0.515	0.184	0.662	0.655
Zhou et al[24]	U-Net++[a]	0.672	0.722	0.581	0.120	0.720	0.722
Deng-Ping Fan et al[19]	Inf-Net	0.692	0.781	0.682	0.082	0.838	0.781
Deng-Ping Fan et al[19]	Semi-Inf-Net	0.725	0.800	0.739	0.064	0.894	0.800
Our method	DNN (specific background and infected region)[b]	0.701	0.942	0.757	0.082	0.867	0.783

[a]Models are analysed in Reference 19 using 45 training images and 50 test images.[50]
[b]589 824 training point from 30 images and 20 test images,[17] 30 test images.[19]

TABLE 9 Percentage improvement compared between our model with existing work

Models	Sensitivity	Specificity	Dice	MAE	EM_φ	S_m
U-Net	31.2%	9.79%	72.43%	44%	38.7%	25.8%
Attention-UNet	10.04%	26.61%	29%	73%	17.3%	5.2%
Gated-UNet	6.53%	29.93%	21%	80%	6.5%	8%
Dense-UNet	18.01%	43.8%	46%	44.4%	30.9%	19%
U-Net++	4.31%	30.47%	30%	68%	20.4%	6%
Inf-Net	1.30%	20.6%	10.9%	—	3.4%	0.25%
Semi-Inf-Net	—	17.75%	2.4%	—	—	—
Avg. improvement with our method	10.19%	25.5%	30.2%	44.2%	16.7%	9.17%

Note: —, Unavailability of data.

6 | CONCLUSION

In this paper, we propose a DNN model for COVID-19 detection from the lung's CT axial view image. The proposed method has adopted the ZM for shape features, GLCM for texture feature and specific- region for selecting the training points, precisely to extract the unique information from the CT image. Our design gives an effective detection of the COVID-19 infection in the lungs and an appropriate tool for radiologists to define the infection stage/percentage. Our model provided better results with limited number of training points. The average test dataset performance reached 70%, 94%, 86% and 78% in terms of Sensitivity, Specificity, EM_{φ} and S_m, respectively. With the proposed method, the average performance is improved in terms of Sensitivity, Specificity, Dice and EM_{φ} by 10.9%, 26%, 24.7%, and 16.5%, respectively in comparison to other popular deep networks such as U-Net, Gated-UNet, Dense-UNet, U-Net++, Inf-Net, Semi-UNet. Moreover, these popular deep networks need a larger number of images for training dataset to maintain the model performance. From the results and higher evaluation metrics, it is evident that the proposed model performed significantly well with much smaller number of training points in the training dataset. The limitation of our model is the difficulty in detecting the GGO from the poor contrast CT images which require additional features and enhancement methods to provide more detailed information. In future, the inner structure information can be extracted using unique texture features and added to the training dataset to improve the performance of the DNN model and also to overcome the limitations.

DATA AVAILABILITY STATEMENT

COVID-19 CT image dataset that support the findings of this study are openly available in https://radiopaedia.org/articles/COVID-19-3?lang=us.Codes are deposited into the public github database https://github.com/deepikas517/Auto-segment-COVID-19-CT-images.git. Remaining materials and supporting data of this article are included within the article.

ORCID

Deepika Selvaraj https://orcid.org/0000-0003-0348-6627
Arunachalam Venkatesan https://orcid.org/0000-0002-7797-4749
Alex Noel Joseph Raj https://orcid.org/0000-0003-1505-3159

REFERENCES

1. Shi F, Wang J, Shi J, et al. Review of Artificial intelligence techniques in imaging data acquisition, segmentation and diagnosis for COVID-19. *IEEE Reviews in Biomedical Engineering*. 2020;1-1. https://doi.org/10.1109/rbme.2020.2987975.
2. La Rosa G, Iaconelli M, Mancini P, Bonanno Ferraro G, Veneri C, Bonadonna L, Lucentini L, Suffredini E. First detection of SARS-CoV-2 in untreated wastewaters in Italy. *Science of The Total Environment*. 2020;736:139652. https://doi.org/10.1016/j.scitotenv.2020.139652.
3. Fong SJ, Dey N, Chaki J. *Artificial Intelligence for Coronavirus Outbreak*. Singapore: Springer; 2020. https://doi.org/10.1007/978-981-15-5936-5
4. Syrjala H, Broas M, Ohtonen P, Jartti A, Pääkkö E. Chest magnetic resonance imaging for pneumonia diagnosis in outpatients with lower respiratory tract infection. *European Respiratory Journal*. 2017;49(1):1601303. https://doi.org/10.1183/13993003.01303-2016.
5. Zech JR, Badgeley MA, Liu M, Costa AB, Titano JJ, Oermann EK. Variable generalization performance of a deep learning model to detect pneumonia in chest radiographs: A cross-sectional study. *PLOS Medicine*. 2018;15(11):e1002683. https://doi.org/10.1371/journal.pmed.1002683.
6. Xie X, Zhong Z, Zhao W, Zheng C, Wang F, Liu J. Chest CT for Typical Coronavirus Disease 2019 (COVID-19) Pneumonia: Relationship to Negative RT-PCR Testing. *Radiology*. 2020;296(2):E41-E45. https://doi.org/10.1148/radiol.2020200343.
7. Lee JH, Kim DI, Cho MK. Computed tomography apparatus and method of controlling X-ray by using the same, ed: Google Patents, 2017. https://patents.google.com/patent/US9655584B2/en
8. Forthmann P, Pfleiderer G. Augmented display device for use in a medical imaging laboratory. Google Patents, 2019.
9. Jensen VT. Method and system of acquiring images with a medical imaging device, ed: Google Patents, 2009. https://patents.google.com/patent/US20060293582
10. Scheib S. Dosimetric end-to-end verification devices, systems, and methods. U.S. Patent 9,643,029, issued May 9, 2017. https://patents.google.com/patent/US20150085993
11. Li Y, Xia L. Coronavirus Disease 2019 (COVID-19): Role of Chest CT in Diagnosis and Management. *American Journal of Roentgenology*. 2020;214(6):1280-1286. https://doi.org/10.2214/ajr.20.22954.
12. Salehi S, Abedi A, Balakrishnan S, Gholamrezanezhad A. Coronavirus Disease 2019 (COVID-19): A Systematic Review of Imaging Findings in 919 Patients. *American Journal of Roentgenology*. 2020;215 (1):87-93. https://doi.org/10.2214/ajr.20.23034.
13. Huang C, Wang Y, Li X, et al. Clinical features of patients infected with 2019 novel coronavirus in Wuhan, China. *The Lancet*. 2020;395(10223):497-506. https://doi.org/10.1016/s0140-6736(20)30183-5.
14. Wang L, Wang Y, Ye D, Liu Q. Review of the 2019 novel coronavirus (SARS-CoV-2) based on current evidence. *International Journal of Antimicrobial Agents*. 2020;55(6):105948. https://doi.org/10.1016/j.ijantimicag.2020.105948.

15. Li K, Fang Y, Li W, Pan C, Qin P, Zhong Y, Liu X, Huang M, Liao Y, Li S. CT image visual quantitative evaluation and clinical classification of coronavirus disease (COVID-19). *European Radiology*. 2020;30(8):4407-4416. https://doi.org/10.1007/s00330-020-06817-6.

16. Rajinikanth V, Dey N, Raj ANJ, Hassanien AE, Santosh KC, Raja N. Harmony-search and otsu based system for coronavirus disease (COVID-19) detection using lung CT scan images. *arXiv preprint arXiv:2004.03431*;2020.

17. https://radiopaedia.org/articles/COVID-19-3?lang=us. Last Accessed date April 5, 2020.

18. Gaál G, Maga B, Lukács A. Attention u-net based adversarial architectures for chest x-ray lung segmentation. *arXiv preprint arXiv:2003.10304*; 2020. https://arxiv.org/pdf/2003.10304.pdf

19. Fan DP, Zhou T, Ji GP, et al. Inf-net: automatic COVID-19 lung infection segmentation from CT images. *IEEE Trans Med Imaging*. 2020;39:2626-2637. https://doi.org/10.1101/2020.04.22.20074948

20. Ronneberger O, Fischer P, Brox T. U-net: convolutional networks for biomedical image segmentation. *International Conference on Medical Image Computing and Computer-Assisted Intervention*. Cham: Springer; 2015, October:234-241. https://doi.org/10.1007/978-3-319-24574-4_28

21. Oktay O, Schlemper J, Folgoc LL, Lee M, Heinrich M, Misawa K, Glocker B. Attention u-net: Learning where to look for the pancreas. *arXiv preprint arXiv:1804.03999*; 2018. https://arxiv.org/pdf/1804.03999.pdf

22. Schlemper J, Oktay O, Schaap M, Heinrich M, Kainz B, Glocker B, Rueckert D. Attention gated networks: Learning to leverage salient regions in medical images. *Medical Image Analysis*. 2019;53:197-207. https://doi.org/10.1016/j.media.2019.01.012.

23. Li X, Chen H, Qi X, Dou Q, Fu C-W, Heng P-A. H-DenseUNet: Hybrid Densely Connected UNet for Liver and Tumor Segmentation From CT Volumes. *IEEE Transactions on Medical Imaging*. 2018;37(12):2663-2674. https://doi.org/10.1109/tmi.2018.2845918.

24. Zhou Z, Siddiquee MMR, Tajbakhsh N, Liang J. Unet++: A nested u-net architecture for medical image segmentation. *Deep Learning in Medical Image Analysis and Multimodal Learning for Clinical Decision Support*. Cham: Springer; 2018:3-11. https://doi.org/10.1007/978-3-030-00889-5_1.

25. Fu H, Fan DP, Chen G, Zhou T. COVID-19 imaging-based AI research collection. 2020. https://github.com/HzFu/COVID19_imaging_AI_paper_list.

26. Angulakshmi M, Lakshmi Priya GG. Automated brain tumour segmentation techniques- A review. *International Journal of Imaging Systems and Technology*. 2017;27(1):66-77. https://doi.org/10.1002/ima.22211.

27. Fan DP, Ji GP, Sun G, Cheng MM, Shen J, Shao L. Camouflaged object detection. *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. Seattle, WA, USA, USA: IEEE; 2020:2777-2787. https://doi.org/10.1109/CVPR42600.2020.00285

28. Ali MNY, Sarowar MG, Rahman ML, Chaki J, Dey N, Tavares JMRS. Adam Deep Learning With SOM for Human Sentiment Classification. *International Journal of Ambient Computing and Intelligence*. 2019;10(3):92-116. https://doi.org/10.4018/ijaci.2019070106.

29. Shan F, Gao Y, Wang J., Shi W, Shi N, Han M, Shi Y. Lung infection quantification of covid-19 in CT images with deep learning. *arXiv preprint arXiv:2003.04655*; 2020. https://arxiv.org/ftp/arxiv/papers/2003/2003.04655.pdf

30. Ye C, Wang W, Zhang S, Wang K. Multi-Depth Fusion Network for Whole-Heart CT Image Segmentation. *IEEE Access*. 2019;7:23421-23429. https://doi.org/10.1109/access.2019.2899635.

31. Liu Z, Song Y-Q, Sheng VS, Wang L, Jiang R, Zhang X, Yuan D. Liver CT sequence segmentation based with improved U-Net and graph cut. *Expert Systems with Applications*. 2019;126:54-63. https://doi.org/10.1016/j.eswa.2019.01.055.

32. Dong X, Lei Y, Wang T, Thomas M, Tang L, Curran WJ, Liu T, Yang X. Automatic multiorgan segmentation in thorax CT images using U-net-GAN. *Medical Physics*. 2019;46(5):2157-2168. https://doi.org/10.1002/mp.13458.

33. Chen X, Yao L, Zhang Y. Residual attention U-net for automated multi-class segmentation of COVID-19 chest CT images. *arXiv preprint arXiv:2004.05645*; 2020. https://arxiv.org/pdf/2004.05645.pdf

34. Chen X, Yao L, Zhou T, Dong J, Zhang Y. Momentum contrastive learning for few-shot covid-19 diagnosis from chest CT images. *arXiv preprint arXiv:2006.13276*; 2020. https://arxiv.org/pdf/2006.13276.pdf

35. Narin A, Kaya C, Pamuk Z. Automatic detection of coronavirus disease (covid-19) using x-ray images and deep convolutional neural networks. *arXiv preprint arXiv:2003.10849*; 2020. https://arxiv.org/ftp/arxiv/papers/2003/2003.10849.pdf

36. Shi F, Xia L, Shan F, Wu D, Wei Y, Yuan H, Shen D. Large-scale screening of covid-19 from community acquired pneumonia using infection size-aware classification. *arXiv preprint arXiv:2003.09860*; 2020. https://arxiv.org/ftp/arxiv/papers/2003/2003.09860.pdf

37. Oh Y, Park S, Ye JC. Deep Learning COVID-19 Features on CXR Using Limited Training Data Sets. *IEEE Transactions on Medical Imaging*. 2020;39(8):2688-2700. https://doi.org/10.1109/tmi.2020.2993291.

38. Wang L, Wong A. COVID-Net: a tailored deep convolutional neural network Design for Detection of COVID-19 cases from chest X-ray images. *arXiv preprint arXiv:2003.09871*; 2020. https://arxiv.org/pdf/2003.09871.pdf

39. Butt C, Gill J, Chun D, Babu BA. RETRACTED ARTICLE: Deep learning system to screen coronavirus disease 2019 pneumonia. *Applied Intelligence*. 2020;1-7. https://doi.org/10.1007/s10489-020-01714-3.

40. Konar D, Panigrahi BK, Bhattacharyya S, Dey N. Auto-Diagnosis of COVID-19 Using Lung CT Images with Semi-Supervised Shallow Learning Network. *Research Square*. 2020. https://doi.org/10.21203/rs.3.rs-34596/v1.

41. Singh M, Bansal S, Ahuja S, Dubey RK, Panigrahi BK, & Dey N. Transfer Learning Based Ensemble Support Vector Machine Model for Automated COVID-19 Detection Using Lung Computerized Tomography Scan Data. *Research Square*. 2020. https://doi.org/10.21203/rs.3.rs-32493/v1.

42. Ahuja S, Panigrahi BK, Dey N, Rajinikanth V, Gandhi TK. Deep transfer learning-based automated detection of COVID-19 from lung CT scan slices. *Applied Intelligence*. 2020. https://doi.org/10.1007/s10489-020-01826-w.

43. Dey N, Rajinikanth V, Fong SJ, Kaiser MS, Mahmud M. Social group optimization–assisted Kapur's entropy and morphological

segmentation for automated detection of COVID-19 infection from computed tomography images. *Cognitive Computation*. 2020;12(5):1011-1023. https://doi.org/10.1007/s12559-020-09751-3.

44. Wang X, Deng X, Fu Q, Zhou, Feng J, Ma H, Liu W, Zheng C. A weakly-supervised framework for COVID-19 classification and lesion localization from chest CT. *IEEE Transactions on Medical Imaging*. 2020;39(8):2615–2625. https://doi.org/10.1109/tmi.2020.2995965.

45. Chaki J, Dey N. *A beginner's guide to image preprocessing techniques*. Boca Raton: CRC Press; 2018. https://doi.org/10.1201/9780429441134

46. Subramani B, Veluchamy M. MRI brain image enhancement using brightness preserving adaptive fuzzy histogram equalization. *International Journal of Imaging Systems and Technology*. 2018;28(3):217-222. https://doi.org/10.1002/ima.22272.

47. Raj ANJ, Mahesh VG. Zernike-moments-based shape descriptors for pattern recognition and classification applications. *Advanced Image Processing Techniques and Applications*. Hershey: IGI Global; 2017:90-120. https://doi.org/10.4018/978-1-5225-2053-5.ch004

48. Adapa D, Joseph Raj AN, Alisetti SN, Zhuang Z, Ganesan K, Naik G. A supervised blood vessel segmentation technique for digital Fundus images using Zernike Moment based features. *PLOS ONE*. 2020;15(3):e0229831. https://doi.org/10.1371/journal.pone.0229831.

49. Barstugan M, Ozkaya U, Ozturk S. Coronavirus (covid-19) classification using ct images by machine learning methods. *arXiv preprint arXiv:2003.09424*; 2020. https://arxiv.org/ftp/arxiv/papers/2003/2003.09424.pdf

50. COVID-19 CT segmentation dataset. https://medicalsegmentation.com/covid19/. Accessed April 11, 2020.

51. Cohen JP, Morrison P, Dao L, Roth K, Duong TQ, Ghassemi M. Covid-19 image data collection: Prospective predictions are the future. *arXiv preprint arXiv:2006.11988*; 2020. https://arxiv.org/pdf/2006.11988.pdf

52. Fan DP, Cheng MM, Liu Y, Li T, Borji A. Structure-measure: a new way to evaluate foreground maps. *Proceedings of the IEEE International Conference on Computer Vision*. Venice, Italy: IEEE; 2017:4548-4557.

A Survey on Dental Imaging for Building Classifier to Benefit the Dental Implant Practitioners

Shashikala J[1]

Research Scholar, Department of Electronics and
Communication, JAIN (Deemed-to-be University),
Assistant Professor, BMS Institute of Technology &
Management, Bengaluru, Karnataka, India

Thangadurai N[2]

Professor and Research Coordinator, Department of
Electronics and Communication, Faculty of Engineering and
Technology, Centre for Emerging Technologies, JAIN
(Deemed-to-be University), Bengaluru, Karnataka, India

Abstract—**Endo-osseous implants are considered an ideal dental fixture. It is becoming the preferred choice of the edentulous patient to rehabilitate toothlessness because of their aesthetic and functional outcome. Despite the successful surgery and implant placement, complications occur, which may be related to several factors, like operative assessment, treatment planning, patient-related factors, surgical procedures, and surgeons' experience. Comprehensive radiological assessment plays a vital role in clinical analysis for better treatment planning, avoiding complications, and increasing the Implant's success rate. However, despite the variety of dental imaging, choosing the right imaging technology has become difficult for clinical experts. The investigative survey conducted in this paper aims to determine the correlation between different imaging modalities, their essential role in implant therapy. This review extensively discussed which types of computational operations applied to image modalities in the existing literature address various noises and other relevant issues. These study findings reveal significant issues with various dental imaging modalities and provide an understanding to bridge all existing research gaps towards building cost-effective classification and predictive models for accurate dental treatment planning and higher implant success rates.**

Keywords—*Dental implant; complication; implant failure; dental imaging; pre-processing*

I. INTRODUCTION

It is said that the mouth is a mirror of health that reflects the health condition of a person, or in other words, it is a cautionary system for disease. The mouth consists of both teeth and gums; their health condition is significant for oral health as poor oral health leads to various fatal diseases, too [1]. Apart from the fatal diseases, an issue of Edentulism-(toothlessness) is found in both kids, adults, and old aged people due to respective reasons leads to the inconvenience of chewing the food so, poor nutritional intake and as a result a poor health condition [2]. However, the stage of Edentulism also creates an issue of the hollowness of speaking-(pronunciation) along with other discomforts [3]. The traditional treatments of bridging and dentures were adopted for a long time as this was only a choice of treatment for the condition of Edentulism. However, the modern technique, namely dental implant surgery, is gaining popularity as an alternative solution to meet the deficiency of natural teeth by artificial tooth replacement [4].The dental implant procedure is based on the conception of direct contact between bone and the

metal implant-(osseointegration), whose ultimate objective is to restore all the functional and aesthetic aspects [5]. An extensive planning and clinical examination performed by the dentist before surgery as a minor causality may cause serious harm to the patient. Therefore, an implant's success depends on several mutual factors, like implant region, bone quality, medical history of patient, skills, and the surgeon or dentist's experience. One of the significant challenges in dental implantation practice is the complex surgical procedures, which require preoperative and postoperative evaluation for achieving a higher success rate in dental implants [6]. The preoperative evaluation includes various factors such as the patient's general health conditions, bone quality, alveolar bone axis, and transplant site. The postoperative evaluation is carried out after the implantation to prevent any bias and risk of failure. Medical imaging technology plays a crucial role in the preoperative evaluation process. It provides the patient's anatomical details for the dental Implant-based on the maxillofacial structure and the two-dimensional geometric projection, helping clinical experts decide whether the implant surgery is suitable for the patient [7]. A systematic radiographic evaluation can provide an effective direction for precise positioning, which has important clinical significance in terms of accuracy and functional and aesthetic effects of the Implant [8]. Many imaging techniques are used in clinical dentistry practices, including conventional radiographic images and Computed Tomography (CT) for preoperative assessment and analysis of the complex jawbone structures. However, each imaging modality has some advantages and limitations too. Therefore, choosing the most suitable imaging method for dental implants is still tricky in dental practices. Another major issue is that the dental imaging is mostly associated with the poor image quality and superimposition factors that need to be process with an effective image enhancement and pre-processing techniques. In order to make the dental image representation more explanatory, several studies on dental image analysis have been conducted using digital image pre-processing methods. The proposed study aims to determine the prevalence of digital imaging modalities in dental implants and how they can help improvise the dental implant success rate. Therefore, this paper conducts a review analysis to highlight the importance of various imaging modalities and pre-processing techniques to explore the research gap. The rest of the sections of this paper are organized as follows: Section II discusses the background highlighting complications in the implant procedure and dental implant failures. This section also

discusses how to improvise the dental implant success rate by prediction using image analysis. Section III presents a thorough analysis of what kinds of dental imaging are used and for what purposes. A comprehensive analysis is conducted on dental imaging modalities to highlight their importance and limitations in this section. Section IV presents an analysis of the current state-of-the-art, observing the trends towards adopting radiography and cone-beam computed tomography to avoid anatomical structures critical to dental implant surgery. Section V discussion and perspective are presented. In this, significant research direction based on evidential proofs, open research issues, and inferences is explored to develop effective predictive models to benefit dental implant practitioners. Finally, the overall contribution of this paper is concluded in Section VI.

II. STUDY BACKGROUND

Success and failure are two critical terms in dental implantology. The term implant success can be an ideal clinical setting, meaning that the Implant is into the jawbone and functions well and pleasingly. The term implant failure refers to the loss of osseointegration. Another statement is that it is an initial instance at which the Implant's efficacy, evaluated quantitatively, drops below a cut-off value or specified level [9].Dental implants may fail for various reasons, with the scope that distinguishes between complications and implant failures. This study uses the term implant failure, which is the complete loss of osseointegration, and the severity of implants that require to be removed from the implant site. In order to avoid any form of ambiguity, the study made a distinction between discussing implant failure and complications. Implant complications can be stated as an event that requires quantifiable clinical attention, and if such measures are not taken, the outcome of the implant therapy may be impaired. Implant complications may be caused due to poor patient selection, inadequate pre-assessment of the patient. Also, the degree of complications that are difficult to control may lead to implant failure. Therefore, through the proper patient selection and treatment planning, surgical fixing of implants can provide long-lasting functional and aesthetic restoration to the Edentulous Patient. Various studies have attempted to identify and quantify the rate of dental implant-related complications. However, to date, no single standard system for classifying dental implant-related complications. The authors in [10] discussed specific categories of complications related to dental implants. Existing studies [11-14] suggested the classification of complications associated with implant therapy considering all factors and causes. Other studies [15-16] considered the classification based on the particular phase of implant treatment that they tend to occur. The work carried out in [11] and [6] performed a classification of complications based on surgical, bone loss, implant loss, peri-implant soft tissue mechanical factor, and aesthetic/phonetic factor. In [13], the authors discussed the classification of dental implant complications, mechanical, technical, and biological. Classification of Surgical complications, Biological complications, and Restorative complications is carried out in [14]. The existing work of [15][16] discusses surgical complication based on three factors viz. i) implant treatment associated (wrong angulation, the judgment of improper

implant-site, and lack of Communication among dental disciplines), ii) anatomy associated (nerve injury, bleeding, Sinus membrane complication, and devitalization of adjacent teeth), and iii) procedure associated-(Mechanical complication), lack of stability, mandibular fracture, aspiration, and ingestion. The authors in [17] discussed reversible complications are obstructions that are either temporary or easily fixed.

A proper surgical procedure analysis, including careful radiograph analysis, is significant to reduce the possibility of any implant complications and dental implant failure. Closer evaluation of dental radiographs helps to establish an appropriate treatment strategy for implant patients [18]. Several reviews and remarks have been given since the past few decades that described the significance of imaging techniques in dental disciplines [19-20]. Dental imaging plays a major-role in implant procedures to determine comprehensive information about the patient's maxillo-facial area to understand whether the surgical procedure is suitable for the patient. However, the role of imaging is not limited to determining only the maxillo-facial area but also at different stages of the treatment processes, leading to the ease of surgical practice towards achieving higher success in dental implants [21-22].Imaging in dental treatment stage-1 subjected to patient diagnosis and clinical analysis conducted before implant surgery. Imaging evaluation assists the dentist in making a clinical decision and effective treatment planning based on past radiographs, medical history, and new radiographs evaluations that determine bone angulations, quality of bone, the critical structure of the maxillo-facials, presence of disease, and analysis of the implant site. In dental treatment phase-2, the role of imaging is to care about surgical intervention by assessing the surgical site and implant position during and after surgery and estimates the duration required for healing. Phase-3 of dental treatment begins after the intra-operative assessment and continues until the Implant remains in the jaw. At this stage, dental imaging helps determine the care plan. If any changes or complications are noted during this period, the necessary clinical steps are taken to prevent any possibility of the risk of failure.

However, despite the variety of dental imaging, choosing an appropriate imaging technique has become a challenging task for clinical experts. Each imaging modality is associated with certain advantages and limitations. One of the major issues encountered in the dental image is the poor image quality due to poor contrast, uneven illumination, low resolution, and noise inclusion during the dental image acquisition process. In order to avoid any ill-effect, the radiation is kept low while taking the dental X-ray. The dental x-ray constructed at low-radiation generates very poor-quality images with lower contrast and brightness, causing visibility differences during analysis. The specific noises during the radiography cause degradation to the dental image. [23]. Therefore, an effective mechanism should be implemented to enhance the quality of the image that can provide a significant clinical analysis in dental implant surgical procedures. The criteria that need to be considered as follows:

- The dental image must provide cross-sectional interpretations that describe the spatial relationships between internal structures.

- The dental radiograph should not be compromised with distortion to a greater extent. However, the smallest distortion can be considered with a predictable average error to obtain a quantified analysis and precise measurement.

- It should provide an accurate description of bone density and cortical plate thickness to achieve the initial equilibrium and stability in the Implant.

- Radiography must provide higher dimensional accuracy in implant treatment procedure that includes analysis of implant placement site, pre-existing pathological condition of the patient, and evaluation alveolar thickness.

- The imaging tool should be available/provided at a reasonable price, and radiology doses should be as little as possible.

At present, there is various research works carried out towards dental imaging. However, there is always an impediment towards accurate diagnosis when it comes to medical image processing, as it demands a higher degree of accuracy. Hence, the prime statement of the problem of the proposed study is "To explore the strength and effectiveness of existing methodologies associated with dental imaging approach with respect to classification." The next section discusses about the different dental imaging modalities highlighting their advantage and limitations.

III. DENTAL IMAGING MODALITIES

In this modern era, a variety of imaging technologies are widely used in the dental field. The traditional implant practitioners depend on 2D radiography. The advancement in imaging technology provided a 3D imaging technique, which offers advanced clinical evaluation in dental implants [24] and [25]. This section presents the adoption of verities of imaging modalities and their uses in different dental implant disciplines.

A. Conventional Imaging Modalities

Two-dimensional conventional imaging aims to complement the clinical analysis in dental implants by gaining a deep understanding of the internal teeth structure and alveolar bone. The different conventional imaging modalities are illustrated as follows:

- *Periapical Radiography-*It offers a systematic detail about the anatomical structures like teeth and surrounding tissues around the implant site. It is used for preoperative assessment to understand the implant area's structure, vertical height, and bone quality. However, these imaging modalities may be difficult to adopt due to accurate instrument positioning support's unavailability. This imaging technique is associated with distortion and magnification, limiting the quantified bone quality assessment, and suffers from providing accurate spatial relationships between internal overlapping dental structures [26].

- *Cephalometric Radiography-* This helps to capture the image of the head with the mandible in a lateral view to examine the associations between teeth, jaw, and the remaining part of the facial skeleton. This technique outlines the geometrical structure of the anterior alveolar region. The limitation is that it only displays cross-sectional images of bones associated with low magnification and overlapping issues [27].

- *Panoramic Radiography-* It is an x-ray radiography image that captures the entire mouth structure in a single image representation using a tomographic technique. It visualizes both maxillary-(upper jaw) and mandibular-(lower jaw) dental curves and supporting structures. It is mostly adopted as an initial screening x-ray image to assess dental and bone support, identify affected teeth, and the condition of dental implants. This imaging technique is primarily used in the preoperative assessment to depict jaws in a single radiograph film or a charge-coupled device image receptor [28]. The distinct advantage of the panoramic imaging technique is that it offers a low patient radiation dose and is cost-effective in terms of time and computation complexity. It involves easy functioning and takes little time to capture the entire image of dentition in a single film or image receptor. Like other conventional radiography imaging techniques, it also has some limitations. Since this imaging technique is an extra oral technique, it does not provide delicate anatomy than periapical radiographs. It suffers from the issues like geometric distortion, superimposition, and magnification. Some other problems, like positioning error and technical/ processing error during panoramic radiography [29].

- *Digital Radiography-* It is direct digital radiography carried using several functional units that includes x-ray–sensitive plates, sensors, mechanism of dividing it into electronic segments, and transferred to a computer to present and store the image. Compared to conventional imaging modalities, direct digital radiography offers good image quality with very little radiation. Few studies have mentioned that the overall reduction of radiation dose is up to 80% [30] and about 50% to 70% radiation reduction in intra oral and extra oral digital imaging [31]. Direct digital radiography has reduced processing time; images can be obtained immediately during the surgical procedure. Since this image is stored and processed in a computer, it can be manipulated with software programs to obtain enhanced visualization and accurate measurement. However, one of the significant disadvantages of digital radiography techniques is that the localization of sensors in the implant site sometimes becomes very challenging due to sensor size and positioning of the connecting cord.

Various conventional imaging modalities are discussed above. The limitation of conventional technology is that it encounters the superimposition of overlapping structures. The overlapping structure is caused due to the depiction of three-dimensional maxillofacial structures onto a two-dimensional

image plane, which results in the loss of spatial information that complicates the identification of objects of interest. The next sub-section discusses the applications and advantages of advanced imaging modalities in dental treatment.

B. Advanced Imaging Modalities

The conventional imaging modalities provide evidence for routine dentistry practices. Advanced imaging mechanisms are needed to demonstrate more information, complex diagnostics, and dental implant treatment plans. Hence, several techniques have changed the diagnosis and treatment planning strategies of dentistry. Some advanced dental imaging modalities are given below:

- *Computerized Axial Tomography (CAT)* - This is a unique X-ray imaging mechanism named computed tomography (CT), displaying the detailed images of the patient's anatomy with hard-and-soft tissues of the maxillofacial region. The CT uses multiple X-rays to construct a two-dimensional maxillofacial region and is converted into a three-dimensional image through processing. CT can obtain multiple, cross-sectional image-(slices) and generate high-contrast resolution images without suffering from superimposition and noise issues [32]. CT scans used to determine the quality of bones and the arrangement of teeth that cannot be efficiently obtained by the periapical imaging technique. CT identifies the diseases and immediacy of critical structures where implants are placed with the differentiation of tissues for analysis. The limitation of CT radiographs is that it has higher radiation exposure, high scan cost, and may not provide a good view of the small fissure resulting in false-negative readings [33].

- *Magnetic Resonance Imaging (MRI)* - MRI includes radio waves and adopts hydrogen atoms 'behaviour within a large magnetic field to look at body regions and generates an MR image of the internal structure. MRI represents soft tissue differences with high contrast sensitivity, which makes it advantageous over CT imaging. MR images can distinguish minor alveolar ducts and the contours between cortical bone and cancellous bone, thus obtaining necessary information about the maximum implant length, angle, and stability [34]. The MRI in the dental implant procedure seems to be an effective mechanism for 3-D imaging as it avoids the radiation risk of CT imaging. The adoption of MRI depends on the specific use conditions for an accurate diagnosis. The MRI achieves a flexible acquisition plane without changing image quality and resolution. However, MRI is susceptible to artifacts, distortion, and signal loss due to high magnetic susceptibility materials, while dental amalgam has little effect [35].

- *Cone Beam Computed Tomography (CBCT)* – CBCT is a variation of conventional CT. The application of CBCT is mainly for carrying diagnosis and planning of surgery in dental implants. One scan can produce many images of the area-of-interest. CBCT involves the mechanism of a cone-shaped-X-ray-beam moving around the patient to produce a large number of 2D views of ROI, and it is then converted into a 3D view using a cone-beam algorithm. CBCT in dentistry offers a high-resolution representation of bone and teeth, giving a spatial relationship between the adjacent structures. CBCT is used to evaluate osseous disease and identify jaw bone infections and diseases that help perform risk-free surgery, i.e. complications (pain and swelling) [36]. CBCT includes fast scanning procedures associated with lower radiation dose, lower scan cost and DICOM compatibility and has reduced metal product interference than other methods [37-38]. The limitation of CBCT is that it has a limited contrast range, gives fewer details of internal soft tissues, and has a large noise factor and artifacts.

C. Primary Findings

All the imaging methods have a vital role in dentistry applications. The conventional 2D and advanced 3D radiographs provide necessary information for dental treatment and Implant, while a dental digital panoramic image can offer a clinical diagnosis of the jawbone. The significance of digital panoramic imaging is that it has a low radiation dose and shorter exposure time [39]. But intraoral imaging has issues like low image quality, variable magnification, and ghost images. The superposition of the upper cervical spine is the main limitation of panoramic X-ray photography [40], and osseointegration cannot be detected due to overlapping issues [41]. Hence, it is limited to preoperative diagnostic, leading to implant failure [42]. Hence, implantation surgery may compromise the health of nearby soft tissues and cells [43]. The use of CT and CBCT is described in [44-45] over 2D radiographs to assess complex structures like the maxillary sinus. However, the limitations of these imaging modalities are i) not available in many local hospitals due to higher cost and multi-disciplinary technical requirements. The researchers also informed that the patients were exposed to higher radiation doses when CT examination is done than of 2D digital imaging and CBCT examination. Some research works also compared CBCT and digital panoramic imaging to assess the bone height towards planning treatment in different dental implant phases [46] and revealed that digital panoramic is self-sufficient to describe the incisor area but lacks in the canine area. Also, [47] have performed a comparison of error estimation and found CBCT has a better result, which holds a low average preoperative assessment error in the maxillary area than the digital panoramic imaging technique.

IV. STATE-OF-THE-ART REVIEWS

This section presents a review study on the state-of-the-art in the context of digital radiographs adopted in dental implant surgery. Digital radiography is cost-effective and is used in dental radiography. The study (Choi et al. [48]) investigates the impact of enhancement over periapical radiographs by considering three pre-processing techniques for diagnostics. The outcome gives quality differences between the processed image and the input image. A work of (Hao et al. [49]) considered denoising CBCT dental images where improved non-local means filtering is applied [49]. The outcomes demonstrated in terms of PSNR and MSE. The segmentation

operation over digital radiograph image is performed in (Cunha et al.) for the accurate visualization of dental Implant and crestal bone line [50]. A contrast enhancement over Digitized film-based panoramic dental image using the CLAHE-Rayleigh is found in the study of (Suprijanto et al.). The study outcome shows that this method has achieved better performance in terms of PSNR [51]. The authors in the study of (Yin et al.) have used approaches of noise filtering technique for CBCT image based on thresholding mechanisms and wavelet transform [52]. (Mortaheb and Rezaeian) introduces an automated dental CT image approach for identifying the vertical structure and arrangement of the teeth [53]. The study (Lamecker et al.) focuses on automated segmentation operation for Computer-assisted craniomaxillo facial surgery using cone-beam volumetric tomography-(CBVT) dental image [54]. A noise that occurred by positioning error in the digital panoramic dental image is considered in the work of (Amiri and Moudi et al.) and (Kandan and Kumar), which achieves better visualization of the roots of maxillary teeth in the digital radiograph [55-56]. The work carried out by (Naik et al.) used the histogram equalization technique for enhancing the overall visualization of the digital radiographs for accurate analysis of the bone structure and quality [57]. The authors (Kamezawa et al.) used a multiple noise filtering approach for CBCT imaging for exposure radiation dose reduction in an automated guided patient positioning system [58]. An edge enhancement-based pre-processing technique is applied on panoramic X-Ray in the study (Jufriadif et al.) to detect proximal caries [59]. The work of (Supriyanti et al.) used a point processing mechanism for contract stretching of a digital panoramic dental image [60]. In the study of (Khatter et al.), the authors have applied a multi-scale retinex mechanism over CBCT to perform a precise assessment of root canal anatomy for endodontic therapy [61]. An image pre-processing I2I scheme based on neural network architecture is adopted in the research work of (Zhao et al.), which considers generative adversarial networks (GAN) to suppress ring artifacts [62]. Mean-shift algorithm-based image segmentation is adopted in the study of (Gunawan et al.). The authors have identified a fuzzy region in the segmented image and performed fuzzy merging processes based on similarity measurement [63]. A work towards brightness preserving in dental digital periapical images using entropy and histogram analysis is found in the study of (Qassim et al.) [64]. A most recent research work carried out by (Abdallah et al.) [65] have used Anisotropic filtering to eliminate noise, and Contrast Limiting Adaptive Histogram Equalization (CLAHE) to enhance contrast, and sharpness of the dental panoramic image.

V. DISCUSSION AND PERSPECTIVE

Several radiographic modalities were described with their respective features and limitations. Each has its applicability in respective dental conditions to assist the dentists in planning, evaluation, and implant treatment. A precise strategy can reduce the surgical complexity and postoperative complications and lead to higher success considering both aesthetic and functional aspects. Therefore, suitable radiographic selection plays an important role, and the advanced 3D radiograph technique provides all the functional utilities compared to the conventional radiograph technique. Due to the cost factor, digital, panoramic radiography is in

wide use. However, advanced imaging modalities like (MRI, CT, and CBCT) provide better visualization and compatibility with analysis tools so that many complementary and significant information for successful dental implant planning is made available. The MRI facilitates precise localization of the complex structures and useful when the differentiation of soft tissue analysis is requiring, but it carries artifacts like geometric distortion. CT imaging is more suitable for the analysis of bone quantity and quality because it can quickly cover the expanded anatomical area and generate images with reduced noise caused by the patient's movement. The advanced and recent modality, namely, Cone Beam CT (CBCT), offers fast data acquisition of the complete field of view with minimal radiation exposure. It is useful in the diagnosis and Endodontic treatment. In all the above discussed, dental imaging modalities suffer image quality degradation due to various factors like superimposition, geometric distortion, loss of signal, contrast, motion artifacts, and positioning errors that cause challenges during interpretation. The efficient pre-processing techniques can enhance image quality; thereby, significant interpretations for accurate treatment planning in the pre-assessment phase during surgery can be achieved. The post-surgery complication can be avoided to illuminate the possibilities of implant failure.

A systematic review of existing research literature with these imaging modalities is inferred, used while proposing models for segmentation of ROI and classification of complex anatomical structures of the oral region. This paper potentially identifies the trend of the pre-processing techniques adopted and also found that both 2-D dental radiographs and CBCT are advantageous over other modalities. It is recommended that adopting 2D dental imaging with an efficient pre-processing technique for enhancement will be a better choice in implant treatment planning and surgical process until CBCT matures. In the future, CBCT with efficient pre-processing for enhancement and noise filtering may provide a way better path towards an effective modality for successful dental implantation.

A. Research Gap

Based on the above discussion and review analysis, the significant open research problem is highlighted as follows:

No standard open-source dataset is available for the analysis of CBCT. In most research works, the dataset was either collected from the hospitals or considered based on the experimental setup. It has also been seen that few research works have considered dental image data from internet sources.

- Most image enhancement techniques are in the transform domain so that some artifacts may appear in the output image. As a result, it may lead to over-enhancement and issues related to the edge of the image.

- Lack of novelty is analyzed in most of the existing literature subjected to dental image pre-processing tasks. Most of the existing research works follow a similar pattern towards pre-processing the medical image. An improvement and optimization mechanism should be considered.

- The research works towards a predictive model have also not focused on the computational complexity associated with their prediction model for classification of the anatomical structure in preoperative assessment for the Dental Implant.

- Analysis of dental Images based on consideration of suitable parameters is missing in the existing literature. In order to perform effective image analysis, researchers must Analysis and evaluate image quality based on the HSV feature and statistics error metrics like Peak-Signal-to-noise-ratio, MSE-(Mean square error), SNR-(Signal to noise ratio), CNR-(Contrast to noise ratio), SD-(Spectral Distance) and SSIM-(Structural Similarity index.).

- Standard benchmarking is also missing in most of the existing image pre-processing methods.

VI. Conclusion

A dental implant is a complicated procedure that involves multi-disciplinary activities for treatment and surgical planning. Appropriate knowledge and understanding of the complexity and evaluation of implant failure factors is crucial for dental practitioners. Apart from this, digital imaging analysis is critical stage clinicians need to understand the technical parameters. However, equally, it is essential to manipulate these dental radiographs using a suitable pre-processing mechanism to know the potential factors associated with each stage of implant treatment. This paper has presented an investigative review analysis of different complications factors, various dental imaging modalities, and state-of-art pre-processing techniques. Finally, the proposed survey also explored the significant issues in the existing literature and discussed the significant point of highlighting the open research problem. Therefore, the proposed review works provide an effective future research direction for establishing predictive models with effective pre-processing schemes to benefit dental implant practitioners.

References

[1] Clark, Danielle, and Liran Levin. "In the dental implant era, why do we still bother saving teeth?." Dental Traumatology 35, no. 6 (2019): 368-375.

[2] Andersson, Lars, Jens O. Andreasen, Peter Day, Geoffrey Heithersay, Martin Trope, Anthony J. DiAngelis, David J. Kenny et al. "Guidelines for the Management of Traumatic Dental Injuries: 2. Avulsion of Permanent Teeth." Pediatric dentistry 37, no. 6 (2015).

[3] Emami, Elham & de Souza, Raphael &Kabawat, Marla &Feine, Jocelyne. (2013). The Impact of Edentulism on Oral and General Health. International journal of dentistry. 2013. 498305. 10.1155/2013/498305.

[4] Alajlan, Abdulrahman, AryafAlhoumaidan, AbeerEttesh, and Mazen Doumani. "Assessing Knowledge and Attitude of Dental Patients regarding the Use of Dental Implants: A Survey-Based Research." International journal of dentistry 2019 (2019).

[5] Oh, Ji-hyeon. "Recent advances in dental implants." Maxillofacial plastic and reconstructive surgery 39, no. 1 (2017): 33.

[6] Bryce, G., D. I. Bomfim, and G. S. Bassi. "Pre-and postoperative management of dental implant placement. Part 2: management of early-presenting complications." British dental journal 217, no. 4 (2014): 171.

[7] Gupta, Sarika, Neelkant Patil, Jitender Solanki, Ravinder Singh, and Sanjeev Laller. "Oral implant imaging: a review." The Malaysian journal of medical sciences: MJMS 22, no. 3 (2015): 7.

[8] Jayadevappa, Busnur Shilpa, G. S. Kodhandarama, and S. V. Santosh. "Imaging of dental implants." Journal of Oral Health Research 1, no. 2 (2010): 50-62.

[9] Esposito M, Hirsch JM, Lekholm U, Thomsen P. Biological factors contributing to failures of osseointegrated oral implants (I) success Criteria and Epidemiology. Eur J Oral Sci 1998;106:527-51;

[10] Hanif, Ayesha, Saima Qureshi, Zeeshan Sheikh, and Haroon Rashid. "Complications in implant dentistry." European journal of dentistry 11, no. 01 (2017): 135-140.

[11] Goodacre CJ, Kan JY, Rungcharassaeng K. Clinical complications of osseointegrated implants. J Prosthet Dent 1999;81(5):537-552.

[12] Goodacre CJ, Bernal G, Rungcharassaeng K, Kan JYK. Clinical complications with implants and implant prostheses. J Prosthet Dent 2003;90(2):121-132.

[13] Hanif, Ayesha, Saima Qureshi, Zeeshan Sheikh, and Haroon Rashid. "Complications in implant dentistry." European journal of dentistry 11, no. 01 (2017): 135-140.

[14] Guo, Q., R. Lalji, A. V. Le, R. B. Judge, D. Bailey, W. Thomson, and K. Escobar. "Survival rates and complication types for single implants provided at the Melbourne Dental School." Australian dental journal 60, no. 3 (2015): 353-361.

[15] Misch K, Wang H. Implant surgery complications: etiology and treatment. Implant Dent 2008;17(2):159-168.

[16] Misch, Kelly, and Hom-Lay Wang. "Implant surgery complications: etiology and treatment." Implant dentistry 17, no. 2 (2008): 159-168.

[17] (60) Park SH, Wang HL. Implant reversible complications: classification and treatment. Impl Dent 2005;14:211-220.

[18] Gupta, Sarika, Neelkant Patil, Jitender Solanki, Ravinder Singh, and Sanjeev Laller. "Oral implant imaging: a review." The Malaysian journal of medical sciences: MJMS 22, no. 3 (2015): 7.

[19] Vandenberghe, Bart. "The digital patient–Imaging science in dentistry." Journal of dentistry 74 (2018): S21-S26.

[20] Satpathy, Anurag, Rajeev Ranjan, SubhashreePriyadarsini, Somesh Gupta, Piyush Mathur, and Monalisa Mishra. "Diagnostic Imaging Techniques in Oral Diseases." In Medical Imaging Methods, pp. 59-95. Springer, Singapore, 2019.

[21] Vandenberghe, Bart. "The digital patient–Imaging science in dentistry." Journal of dentistry 74 (2018): S21-S26.

[22] Chandak, Shruti, Arjit Agarwal, Ashutosh Kumar, Rajul Rastogi, Pawan Joon, Asif M. Wani, and Yuktika Gupta. "Comparative Study of DENTA Scan and Radiography for Preoperative Assessment of Dental Implants." Annals of International Medical and Dental Research 4, no. 1 (2018): 26.

[23] Rahmi-Fajrin H, Puspita S, Riyadi S, Sofiani E. Dental radiography image enhancement for treatment evaluation through digital image processing. Journal of clinical and experimental dentistry. 2018 Jul;10(7):e629.

[24] White, Stuart C., and Michael J. Pharoah. "The evolution and application of dental maxillofacial imaging modalities." Dental Clinics of North America 52, no. 4 (2008): 689-705.

[25] Ríos-Santos, José V., Cristina Ridao-Sacie, Pedro Bullón, Ana Fernández-Palacín, and Juan J. Segura-Egea. "Assessment of periapical status: a comparative study using film-based periapical radiographs and digital panoramic images." Med Oral Patol Oral Cir Bucal 15, no. 6 (2010): e952-6.

[26] Tanwani, HemlataBhagwan, Sheetal Sameer Potnis, Sandesh Satish Baralay, and Sameer Sidagouda Patil. "Comparison of conventional and digital cephalometric analysis: A pilot study." Journal of Dental and Allied Sciences 3, no. 2 (2014): 80.

[27] Choi, Bo-Ram, Da-Hye Choi, Kyung-Hoe Huh, Won-Jin Yi, Min-Suk Heo, Soon-Chul Choi, Kwang-Hak Bae, and Sam-Sun Lee. "Clinical image quality evaluation for panoramic radiography in Korean dental clinics." Imaging science in dentistry 42, no. 3 (2012): 183-190.

[28] Jayachandran, Sadaksharam. "Digital imaging in dentistry: A review." Contemporary clinical dentistry 8, no. 2 (2017): 193.

[29] Langland OE, Langlais RP, Preece JW. Principles of dental imaging. 2nd ed. Philadelphia: Lippincott Williams & Wilkins, 2002: 285.

[30] Farman AG, Farman TT. Extraoral and panoramic systems. Dent Clin North Am 2000; 44: 257-272,

[31] Aggarwal V, Logani A, Shah N. The evaluation of computed tomography scans and ultrasounds in the differential diagnosis of periapical lesions.

[32] Shah, Naseem, Nikhil Bansal, and Ajay Logani. "Recent advances in imaging technologies in dentistry." World journal of radiology 6, no. 10 (2014): 794.

[33] Niraj, Lav Kumar, Basavaraj Patthi, Ashish Singla, Ritu Gupta, Irfan Ali, Kuldeep Dhama, Jishnu Krishna Kumar, and Monika Prasad. "MRI in dentistry-a future towards radiation free imaging–systematic review." Journal of clinical and diagnostic research: JCDR 10, no. 10 (2016): ZE14.

[34] Mendes, Silwan, Carin A. Rinne, Julia C. Schmidt, Dorothea Dagassan-Berndt, and Clemens Walter. "Evaluation of magnetic resonance imaging for diagnostic purposes in operative dentistry—a systematic review." Clinical Oral Investigations (2019): 1-11.

[35] Bornstein, Michael M., William C. Scarfe, Vida M. Vaughn, and Reinhilde Jacobs. "Cone beam computed tomography in implant dentistry: a systematic review focusing on guidelines, indications, and radiation dose risks." International journal of oral & maxillofacial implants 29 (2014).

[36] Gupta, Jyoti, and Syed Parveez Ali. "Cone beam computed tomography in oral implants." National journal of maxillofacial surgery 4, no. 1 (2013): 2.

[37] Bornstein, Michael M., Keith Horner, and Reinhilde Jacobs. "Use of cone beam computed tomography in implant dentistry: current concepts, indications and limitations for clinical practice and research." Periodontology 2000 73, no. 1 (2017): 51-72.

[38] Jacobs, Reinhilde, Benjamin Salmon, Marina Codari, Bassam Hassan, and Michael M. Bornstein. "Cone beam computed tomography in implant dentistry: recommendations for clinical use." BMC Oral Health 18, no. 1 (2018): 88.

[39] Suomalainen A, PakbaznejadEsmaeili E, Robinson S. Dentomaxillofacial imaging with panoramic views and cone beam CT. Insights Imaging 2015; 6: 1-16.

[40] Tang Z, Liu X, Chen K. Comparison of digital panoramicradiography versus cone beam computerized tomography for measuring alveolar bone. Head Face Med 2017; 13: 2.

[41] Isidor F. Clinical probing and radiographic assessment in relation to the histologic bone level at oral implants in monkeys. Clin Oral Implants Res 1997; 8: 255-64.

[42] Greenstein G, Cavallaro J, Romanos G, Tarnow D. Clinical recommendations for avoiding and managing surgical complications associated with implant dentistry: a review. J Periodontol 2008; 79: 1317-29.

[43] Tang Z, Liu X, Chen K. Comparison of digital panoramic radiography versus cone beam computerized tomography for measuring alveolar bone. Head Face Med 2017; 13: 2.

[44] Hassan B, Jacobs R. Cone beam computed tomography - 3D imaging in oral and maxillofacial surgery. Eur Med Imaging Rev 2008; 1: 38-40.

[45] Kopecka D, Simunek A, Streblov J, Slezak R, Capek L. Measurement of the interantral bone in implant dentistry using panoramic radiography and cone beam computed tomography: a human radiographic study. West Indian Med J 2014; 63: 503-9.

[46] Renton T, Dawood A, Shah A, Searson L, Yilmaz Z. Post-implant neuropathy of the trigeminal nerve. A case series. Br Dent J 2012; 212: E17.

[47] Angelopoulos C, Thomas S, Hechler S, Parissis N, Hlavacek M. Comparison between digital panoramic radiography and conebeam computed tomography for the identification of the mandibular canal as part of presurgical dental implant assessment. J Oral Maxillofac Surg 2008; 66: 2130-5.

[48] Choi, Jin-Woo, Won-Jeong Han, and Eun-Kyung Kim. "Image enhancement of digital periapical radiographs according to diagnostic tasks." Imaging science in dentistry 44, no. 1 (2014): 31-35.

[49] Hao, Jia, Li Zhang, Liang Li, and Kejun Kang. "An improved non-local means regularized iterative reconstruction method for low-dose dental CBCT." In 2012 IEEE Nuclear Science Symposium and Medical Imaging Conference Record (NSS/MIC), pp. 3422-3425. IEEE, 2012.

[50] Cunha, Pedro, Miguel A. Guevara, Ana Messias, Salomão Rocha, Rita Reis, and Pedro MG Nicolau. "A method for segmentation of dental implants and crestal bone." International journal of computer assisted radiology and surgery 8, no. 5 (2013): 711-721.

[51] Juliastuti, E., and LusiEpsilawati. "Image contrast enhancement for film-based dental panoramic radiography." In 2012 International Conference on System Engineering and Technology (ICSET), pp. 1-5. IEEE, 2012.

[52] Yin, Yong, Gang Yu, Hongjun Wang, Zhi Liu, and Dengwang Li. "CBCT image denoising based on multi-scale wavelet transform." In 2010 3rd International Conference on Biomedical Engineering and Informatics, vol. 1, pp. 150-153. IEEE, 2010.

[53] Mortaheb, Parinaz, Mehdi Rezaeian, and Hamid Soltanian-Zadeh. "Automatic dental CT image segmentation using mean shift algorithm." In 2013 8th Iranian Conference on Machine Vision and Image Processing (MVIP), pp. 121-126. IEEE, 2013.

[54] Lamecker, Hans, Stefan Zachow, Antonia Wittmers, Britta Weber, H. Hege, B. Isholtz, and Michael Stiller. "Automatic segmentation of mandibles in low-dose CT-data." International Journal of Computer Assisted Radiology and Surgery 1 (2006): 393.

[55] AsadiAmiri, Sekine, and Ehsan Moudi. "Image quality enhancement in digital panoramic radiograph." Journal of AI and Data Mining 2, no. 1 (2014): 1-6.

[56] Kandan, R. Somas, A. John, and S. Kumar. "An improved contrast enhancement approach for panoramic dental x-ray images." ARPN J Eng App Sci 10 (2015): 1897-1901.

[57] Naik, Anjali, Shubhangi Vinayak Tikhe, and S. D. Bhide. "Histogram Equalization for Class-Identification of Dental Disease Using Digital Radiography." In International Conference on Business Administration and Information Processing, pp. 144-151. Springer, Berlin, Heidelberg, 2010.

[58] Kamezawa, Hidemi, KatsutoshiShirieda, Hidetaka Arimura, Noboru Kameda, and Masafumi Ohki. "An approach of exposure dose reduction of cone-beam computed tomography in an image guided patient positioning system by using various noise suppression filters." In 2014 Joint 7th International Conference on Soft Computing and Intelligent Systems (SCIS) and 15th International Symposium on Advanced Intelligent Systems (ISIS), pp. 1475-1780. IEEE, 2014.

[59] Na'am J, Harlan J, Madenda S, Santony J, Suharinto C. Detection of proximal caries at the molar teeth using edge enhancement algorithm. International Journal of Electrical and Computer Engineering. 2018 Oct 1;8(5):3259.

[60] Supriyanti R, Setiadi AS, Ramadhani Y, Widodo HB. Point Processing Method for Improving Dental Radiology Image Quality. International Journal of Electrical and Computer Engineering (2088-8708). 2016 Aug 1;6(4).

[61] Khatter, Ashish, Anita Thakur, and Nitya Reddy. "CBCT Image Feature Enhancement for Endodontic Therapy." In 2019 6th International Conference on Signal Processing and Integrated Networks (SPIN), pp. 293-296. IEEE, 2019.

[62] Zhao, Shuyang, Jianwu Li, and QirunHuo. "Removing ring artifacts in CBCT images via generative adversarial network." In 2018 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP), pp. 1055-1059. IEEE, 2018.

[63] Gunawan W, Arifin AZ, Indraswari R, Navastara DA. Fuzzy Region Merging Using Fuzzy Similarity Measurement on Image Segmentation. International Journal of Electrical & Computer Engineering (2088-8708). 2017 Dec 1;7(6).

[64] Qassim, Hassan M., Nasseer M. Basheer, and Mazin N. Farhan. "Brightness preserving enhancement for dental digital X-ray images based on entropy and histogram analysis." J Appl Sci Eng 22 (2019): 187-94.

[65] Yousif Mohamed, Nouf H Abuhadi, and Maryam Hasan Hugri"Enhancement of Dental X-rays Images Using Image Processing Techniques"Journal of Research in Medical and Dental Science 2021, Volume 9, Issue 2, Page No: 12-16.
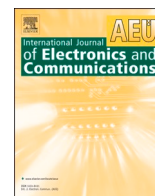
Regular paper

Circularly polarized ultra–wide band filtering antenna with controllable band-notch for wireless communication system

Prashant Ranjan [a,*], Amit Kumar [b]

[a] Department of Electronics and Communication Engineering, University of Engineering and Management Jaipur, Rajasthan, India
[b] Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bengaluru, India

ARTICLE INFO

ABSTRACT

A UWB filtering antenna (or filtenna) with a band-notched characteristic has been proposed in this paper. A square-shaped-stepped patch UWB antenna and an inverted T–shaped bandpass filter based integrated filtering antenna (IFA) with band-notch characteristic has been designed and fabricated. The size of the proposed filtenna is the same as the UWB patch antenna. The UWB patch antenna has a bandwidth from 3 GHz to more than 14 GHz. A UWB filter having a notched band is proposed and integrated with the primarily offered square-shaped– stepped patch antenna to design an IFA for UWB wireless communication system. The filter is placed on the feeding position of the patch antenna, and the output of the filter is coupled with the radiating patch to achieved the desired result. Proposed filtenna achieves circular polarization, minimum of 14 dB out of band gain rejection, and a notched band from 5.55 to 6.98 GHz, which avoids the interference from Dedicated Short Range Communication (DSRC) applications. A wide axial-ratio bandwidth (AXBW) from 3.1 to 5.5 GHz and 7–9.8 GHz along with a maximum gain of 6.8 dBi within the operating frequency range of the proposed filtenna has been achieved.

1. Introduction

Filters and antennas are among the essential elements of an ultra-wideband (UWB) wireless communication device. Modern UWB antennas should have a small size to make them suitable for small portable devices. Planar microstrip antennas printed on the circuit board are a better option for many wireless communications. Microstrip antennas are ideal for UWB applications because they are featured with a low profile, lightweight, inexpensive, and ease of fabrication. Planar antennas with different shapes and different types are widely reported as they demonstrate the fundamental features related to UWB technology [1].

There are chances to receive undesired adjacent frequency components by the antenna, which will affect the receiver performance. A bandpass filter with notch band characteristic is needed to suppress the unwanted signals. On the other hand, it is increasingly popular to integrate the bandpass filter and antenna into a single module to reduce the insertion loss as well as the size of the system. As a result, some filtering antennas are reported having a filter-like frequency response [2–5]. The microstrip antenna is integrated vertically with differential

resonators [2], open stubs, and parasitic strips [3], shorting vias [4], and the multi stub feed [5] for creating band-notch characteristics. Microstrip patch filtering antennas [6–9], filtering horn antennas [10–11], filtering planar ultra-wideband (UWB) antennas [12–14], substrate integrated waveguide (SIW) filtennas [15–16], and filtering slot antennas [17] are reported in the existing literature. In some cases filter based on stub-loaded open-loop resonator [18–19] have been used to create multi-band and band-notches. These designs are based on the filter synthesis approach. This method can provide a satisfying filtering response, but antenna gain and radiation of antenna degraded due to the interference and insertion loss of the filter. Another approach is to design the filtering antenna by introducing simple parasitic elements or resonators into the radiator or its feeding circuit. Since there is no requirement of other specific filtering circuits, the resultant designs have lower insertion loss and a more compact size. Many filtering antennas are designed using this approach, but most of them are unidirectional, and no out of band gain rejection characteristic.

This paper proposes an ultra-wideband filtering antenna having band-notched characteristics and an omnidirectional radiation pattern. To avoid the interference between the dedicated short-range

---

* Corresponding author.
E-mail address: prashant.ranjan@uem.edu.in (P. Ranjan).
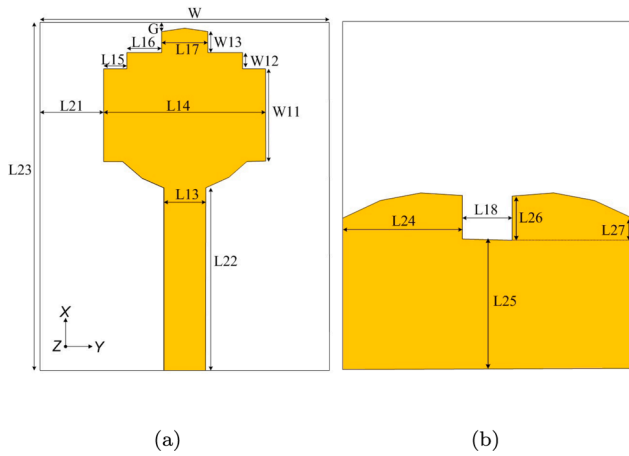
Fig. 1. Design structure of the proposed UWB antenna (a) Top view (b) Bottom view.

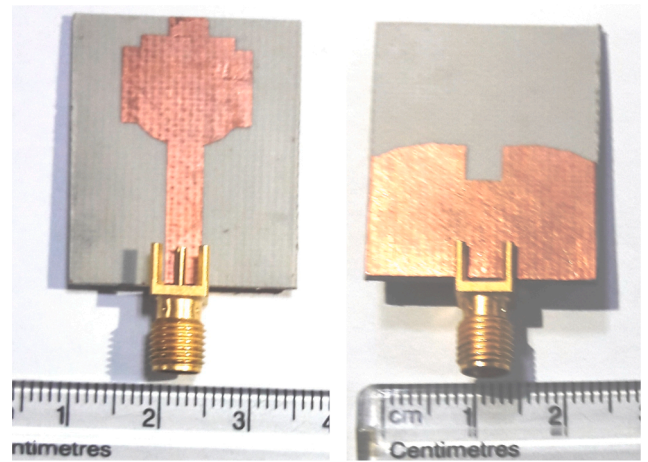(a)   (b)

Fig. 2. Fabricated design of proposed UWB antenna (a) Top view (b) Bottom view.

Table 1
Parameter values of the proposed UWB antenna.

Parameters	Values (mm)	Parameters	Values (mm)
L_{13}	3.6	L_{14}	14
L_{15}	2	L_{16}	3
L_{17}	4	L_{18}	4.3
L_{21}	6.5	L_{22}	17.7
L_{23}	32	L_{24}	10.35
L_{25}	11.14	L_{26}	3.86
L_{27}	2.07	W	27
W_{11}	8	W_{12}	1.4
W_{13}	2.2	W_{14}	11.14
G	0.8	–	–

communications (5.9 GHz is assigned for DSRC applications) and UWB systems, a notch band filter in UWB systems is necessary. However, the use of filters in the UWB system increases the complexity and overall size. A choice to overcome this problem is to create a band-notched in the UWB antenna. But it cannot provide an out of band gain rejection and circular polarization. Thus for out of band gain rejection and circular polarization, proposed a band-notched filtering antenna. A band-notched filter is used by replacing the feed line of the proposed square–shaped stepped patch UWB antenna. The size of the proposed filtenna is the same as the square–shaped stepped patch UWB antenna. This paper is organized in four sections. Section–2 includes the design of a square-shaped stepped patch UWB antenna. In this section, the fabrication of antenna and measured results are also presented. In section–3, an inverted T-shaped notch band UWB filter [20] and their results are discussed and presented. In section–4, we proposed a UWB filtering antenna with a notched band by integrating the UWB antenna to the notch band filter. Simulated and measured results of the proposed filtering antenna with a notched band are discussed and presented. In the last section, the conclusion of the paper is presented.

In this proposed Filtenna an UWB antenna has been integrated with notched band UWB bandpass filter based on inverted T-shaped Multiple Mode Resonator (MMR) to achieve controllable one band-notched characteristics along with an added advantage of circular polarization unlike many others proposed antenna on a single integrated platform.

2. Proposed a square-shaped stepped patch UWB antenna without integration of BPF

The design structure of the proposed antenna is shown in Fig. 1. The printed circuit board of Rogers 4350 having $\varepsilon_r = 3.38$ and $\tan\delta = 0.0037$ is used to design the antenna, where ε_r is relative dielectric constant and $\tan\delta$ is loss tangent of the material. The thickness of the substrate material is 1.524 mm. The size of the antenna is $32 \times 27 \times 1.524$ mm³. ANSYS HFSSv15 [21] software is used to simulate and analyze the antenna. The proposed design consists of the microstrip-fed square-shaped stepped patch. The top portion of the square-shaped patch is stepped, and the lower part of the patch is arc-shaped. A square-shaped stepped patch is chosen as it is compact and gives ultra-wideband response usually. The radiating patch is connected with a microstrip feed line of 50 Ω characteristics impedance. On the bottom of the substrate, partially conducting ground plane is used to design. On the top edge of the partially conducting ground plane, a rectangular slot is cut for proper impedance matching, as shown in Fig. 1(b). Due to this additional rectangular slot provide the coupling over a broad frequency range between radiating patch and ground plane.

In Fig. 1, the golden color shows the metallic part, and the white

Fig. 3. Measured and simulated reflection coefficient of the proposed UWB antenna.

color shows the dielectric substrate. Fabricated view of the proposed antenna is given in Fig. 2. The various optimized values of the proposed antenna are given in Table 1.

Fig. 3 shows the compared return loss between measured and simulated results of square-shaped-stepped patch UWB antenna. |S11| is experimentally measured with Agilent N5230A vector network
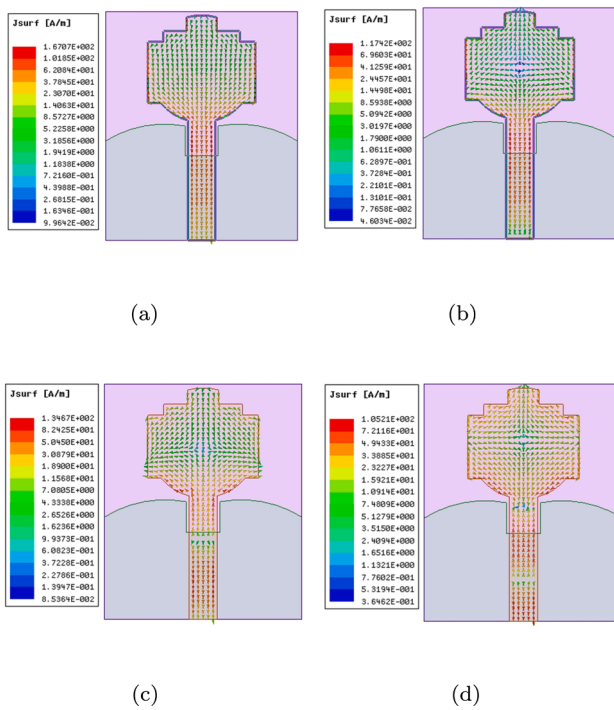
Fig. 4. Current distribution of square-shaped-stepped patch UWB antenna at frequencies (in GHz) (a) 3.5 (b) 5 (c) 8.5 and (d) 10.5.

Fig. 5. Measured and simulated Co- and Cross-polarized radiation pattern of proposed antenna at (a) 3.5 GHz (E-plane) (b) 3.5 GHz (H-plane) (c) 8.5 GHz (E-plane) (d) 8.5 GHz (H-plane) (e) 10.5 GHz (E-plane) (f) 10.5 GHz (H-plane).

analyzer. It is observed that the proposed antenna achieved substantial impedance bandwidth from 3 GHz to more than 14 GHz, considering −10 dB as a reference level. Measured and simulated results are in close agreement with each other.

Fig. 4 shows the surface current distribution of the radiating patch at different frequencies. It is observed from Fig. 4 that the upper and lower part of the patch has the maximum current density. Due to the same phase, they are combined and provide substantial bandwidth (>14 GHz). Fig. 5 shows normalized Co- and Cross-polarized far-field radiation pattern of square-shaped-stepped patch UWB antenna at different frequencies in XZ-plane (E-plane) and YZ-plane (H-plane). We can observe that an isolation of more than 30 dB in the main lobe direction in E-plane while it more than 100 dB difference between the Co– and Cross-polarized radiation pattern in H-plane which advocates that the proposed patch antenna is linearly polarized.

3. UWB bandpass filter based on inverted T-shaped multiple mode resonator (MMR)

An inverted T-shaped ultra-wideband bandpass filter based on multiple mode resonator is proposed and discussed in this section [20]. The proposed filter has a controllable notch-band characteristic. This filter is integrated with the square-shaped-stepped patch UWB antenna to make filtering antenna (or filtenna). The design process and results of the UWB bandpass filter are presented in this section. The filter bandwidth is from 3.145 to 11.01 GHz with a frequency notch band from 5.74 to 6.98 GHz. The relative dielectric constant of substrate material, substrate thickness and substrate width of the filter is the same as a square-shaped-stepped patch UWB antenna. The motive of selecting similar material is to have easy integration of this filter with the antenna to make filtenna or
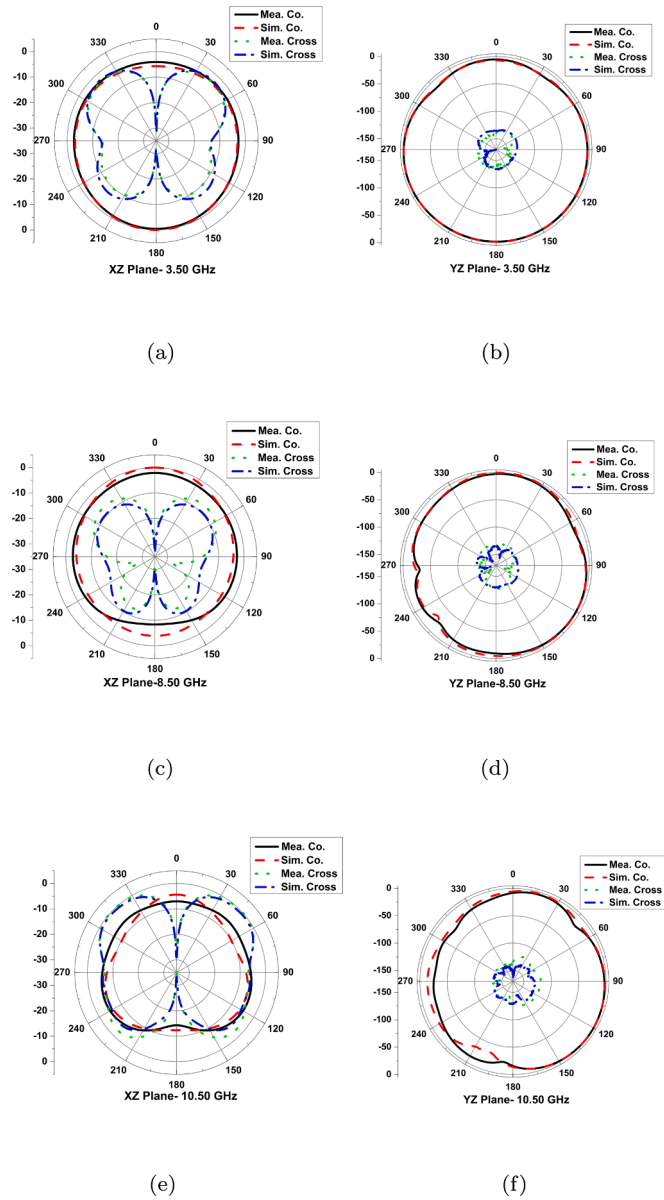
integrated filtering antenna (IFA) on a single board.

3.1. MMR structure and analysis of notch–band UWB bandpass filter

Four open stubs connected with inverted T-shaped MMR are shown in Fig. 6(a). Due to the symmetrical structure of the proposed resonator, many possible applications can find in the tunable band, harmonic suppressed BPFs, UWB, and dual-band bandpass filters. The proposed MMR has symmetry across AA' as shown in Fig. 6(a), so the odd mode and even mode analysis is used to analyze the structure. For odd-mode excitation shown in Fig. 6(b), one end of the MMR is grounded because zero voltage is obtained in the middle part of the resonator. In the case of even mode excitation shown in Fig. 6(c), one end of the
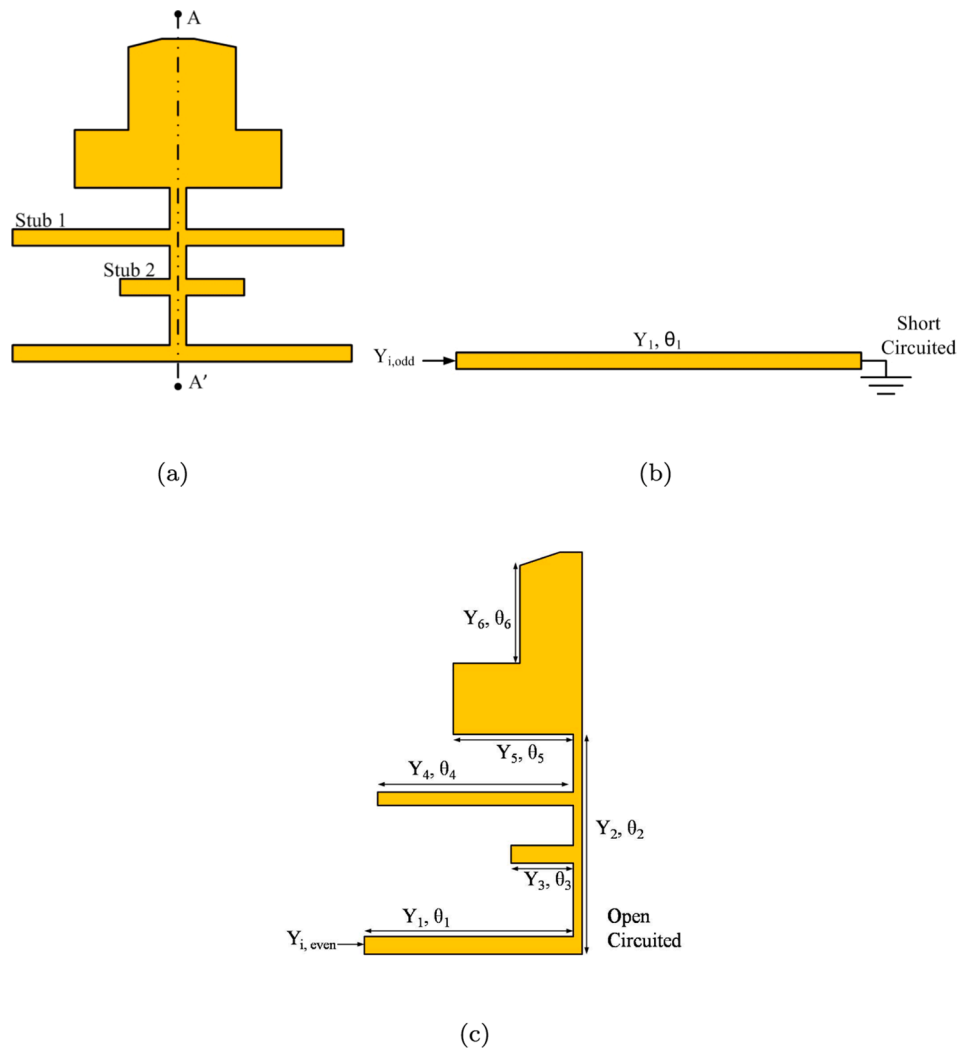
Fig. 6. (a) MMR structure in inverted T-shaped (b) Odd mode design (short-circuited), and (c) Even mode design (open-circuited).
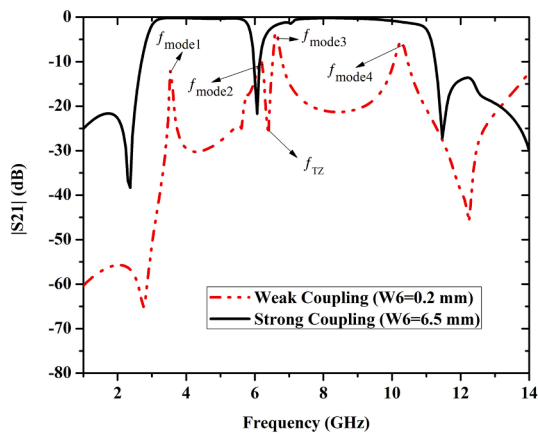
Fig. 7. Simulated |S21| of weakly coupled inverted T-shaped MMR.
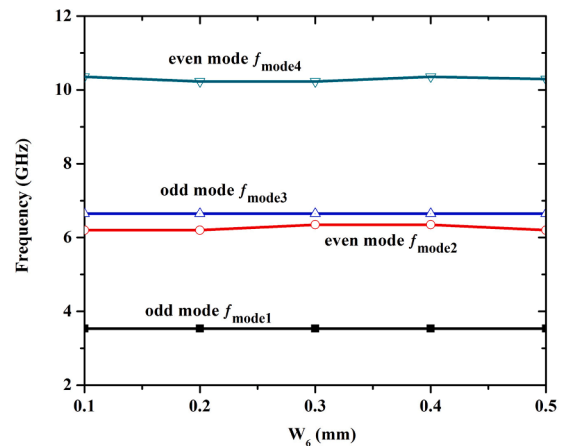
Fig. 8. Resonator characteristics with variation in W_6.

resonator is open-circuited because zero current flows through the symmetrical plane. Fig. 6 shows the approximate circuit model for odd and even mode excitation.

Y_1, Y_2, Y_3, Y_4, Y_5, Y_6 are characteristic admittances, and θ_1, θ_2, θ_3, θ_4, θ_5, θ_6 are electrical lengths of the resonator. Even mode input admittance ($Y_{i,even}$) can be analyzed as

$$Y_{i,even} = jY_1 \frac{Y_{ie1} + (\tan\theta_1 + \tan\theta_2)Y_{ie2}}{Y_{ie2}(1 - \tan(\theta_1)\tan(\theta_2) - Y_{ie1}\tan(\theta_1)} \tag{1}$$

where $Y_{ie1} = \frac{(Y_2)(Y_3)}{(Y_1)^2} + \left(\frac{Y_2}{Y_1}\right)^2 \tan\theta_5 - \frac{Y_3}{Y_1}\tan\theta_4\tan\theta_5\tan\theta_6 + \frac{Y_2}{Y_1}\tan\theta_4$ and $Y_{ie2} = \frac{Y_2}{Y_1} - \frac{Y_3}{Y_1}\tan\theta_5\tan\theta_6 - \frac{Y_2Y_3}{(Y_1)^2}(\tan\theta_4\tan\theta_6) - \left(\frac{Y_2}{Y_1}\right)^2\tan\theta_3\tan\theta_4$

(a)   (b)
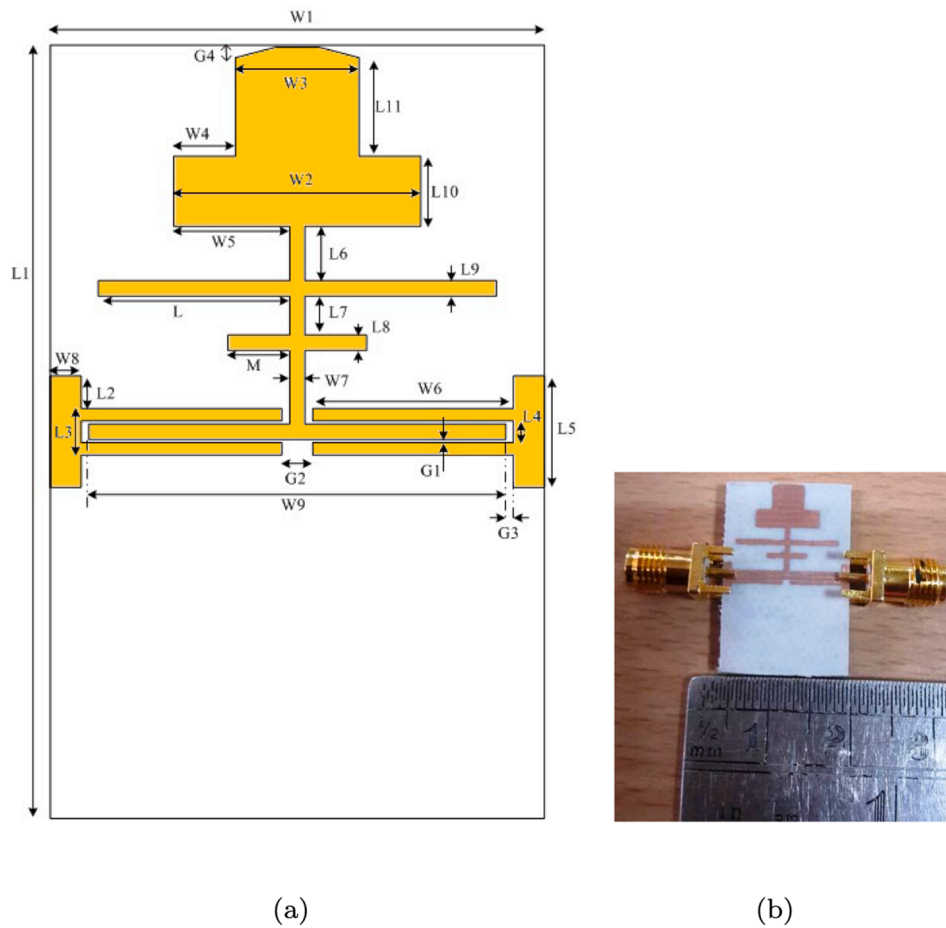
Fig. 9. (a) Layout structure of UWB filter, and (b) fabricated design of UWB filter.

Table 2
Dimensions of the UWB Filter.

Parameters	Values (mm)	Parameters	Values (mm)
L	6.2	L_1	25
L_2	1.05	L_3	1.5
L_4	0.7	L_5	3.6
L_6	1.75	L_7	1.25
L_8	0.5	L_9	0.5
L_{10}	2.26	L_{11}	3.18
W_1	16	W_2	8
W_3	4	W_4	2
W_5	3.75	W_6	6.5
W_7	0.49	W_8	1
W_9	13.5	M	2.2
G_1	0.1	G_2	1
G_3	0.25	G_4	0.4
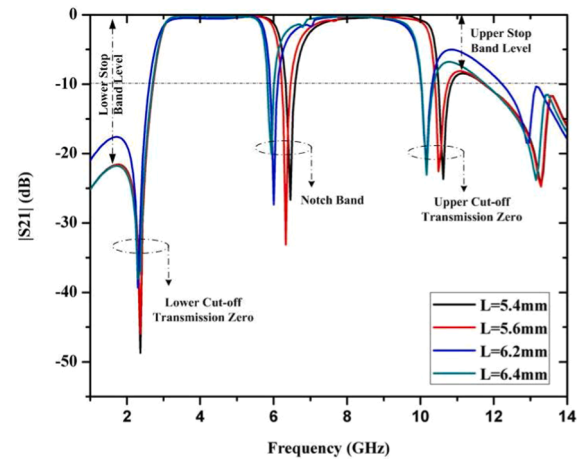
Fig. 10. Shifting the position of f_{TZu} and f_n with variation in L.

Similarly, for odd mode input admittance ($Y_{i,odd}$) can be analyzed as

$$Y_{i,odd} = -jY_1\cot\theta_1 \tag{2}$$

At resonance condition: $Y_{i,even} = 0$ and $Y_{i,odd} = 0$, so even and odd mode resonant frequencies can be calculated by following equations-

$$Y_{ie1} + (\tan\theta_1 + \tan\theta_2)Y_{ie2} = 0 \tag{3}$$

$$\cot\theta_1 = 0 \tag{4}$$

By analyzing the above equations, a controllable notch-band UWB

BPF is designed using an inverted T-shaped resonator with open stubs. Input and output interdigital coupled lines are used. The resonant frequency of the notch-band is controlled by changing the individual stub lengths. The notch-band resonant frequency (f_n) can be calculated by setting-

$$Y_{i,even} = Y_{i,odd} \tag{5}$$

$$f_n = \frac{c}{(5L + M)\sqrt{\varepsilon_{eff}}} \tag{6}$$
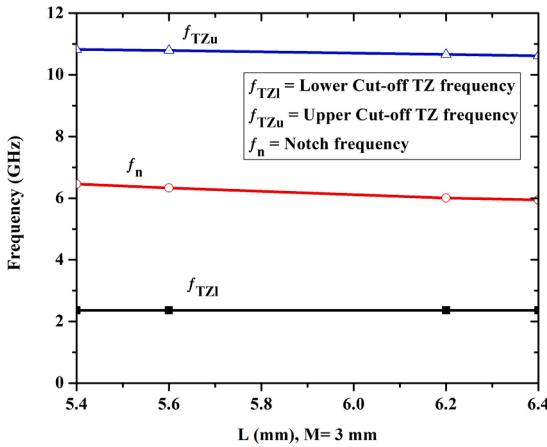
Fig. 11. Characteristics of UWB filter with a change in stub length L.
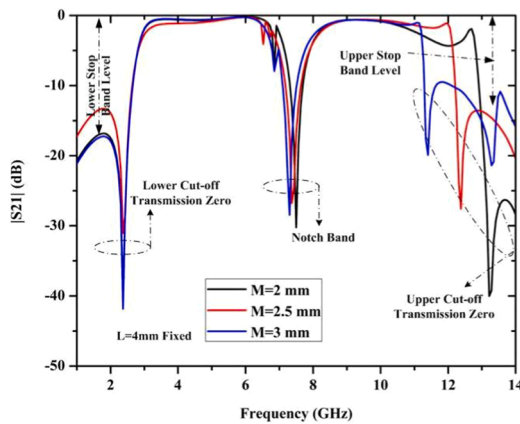
Fig. 12. Change in the position of f_{TZu} with variation in M.
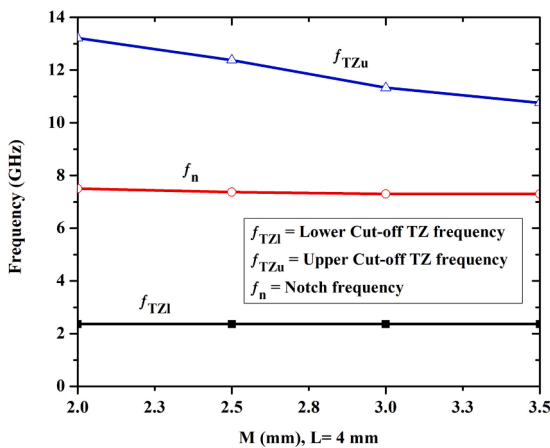
Fig. 13. Characteristics of UWB filter with a change in stub length M.

where L and M are the optimized length of stub 1 and stub 2 respectively. ε_{eff} (≈ 3.46) is the effective dielectric constant of the substrate, and c is the velocity of light in free space.

When MMR is weakly coupled, four modes are created, as shown in
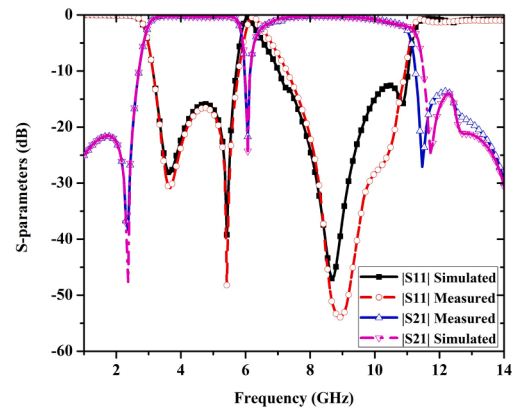
Fig. 14. Measured and simulated S-parameters of the UWB filter.

Fig. 7. Variation in length of interdigital coupled lines (W_6) generates odd mode (f_{mode1}, f_{mode3}) and even mode (f_{mode2}, f_{mode4}), as shown in Fig. 8. It is observed that with a change in W_6, dimensions mentioned in Fig. 9 (a) odd mode frequencies are almost fixed, whereas even mode frequencies are changed.

3.2. Design structure and dimensions of the proposed filter

The proposed filter is designed using Rogers4350 of the substrate material with a thickness of 1.524 mm. The value of relative dielectric constant (ε_r) of substrate material is 3.38, and the loss tangent of material is 0.004. 3D full-wave simulator ANSYS HFSSv15 [21] is used for the optimization of filter design, its response, and simulations. The optimized values of the design parameters of UWB BPF are given in Table 2. The total area covered by the UWB filter is 25×16 mm². Fig. 9 (a) and (b) shows the layout structure and fabricated view of the proposed filter, respectively.

3.3. Notch-band frequency control

Fig. 10 shows the variation in position of notch-band frequency (f_n) and upper cut-off transmission zero frequency (f_{TZu}). When increasing the value of length L (stub 1), the position of f_n and f_{TZu} are changed and shifted left towards lower frequency, whereas the position of lower cut-off transmission zero frequency (f_{TZl}) is not changed and fixed at a frequency near to 2.36 GHz. Also, the upper stop band level is slightly decreased due to the increase in the length of L. In Fig. 11 it is observed that f_n is shifted towards lower frequency and controlled by stub 1.

3.4. Control of upper Cut-off transmission zero frequency

Fig. 12 shows the variation in the position of the upper cut-off transmission zero frequency, f_{TZu} of the proposed UWB filter. When decreasing the length M of stub 2, the position of f_{TZu} is changed rapidly and shifted towards the higher frequency. Also, the upper stopband level is improved by decreasing the value of length M. However, the position of f_{TZl} has no effect and fixed at the frequency of 2.36 GHz. Fig. 13 shows that f_{TZl} and f_n are almost fixed, but substantial change is noticed in f_{TZu}. So the length of stub 2 is used to control f_{TZu} and improve the upper stopband level, which gives better band rejection at a higher cut-off frequency of UWB filter.

Fig. 14 shows the simulated and measured results of the |S11| and |S21|. Length of stub 1 and stub 2 are varied and selected their values as L = 6.2 mm and M = 2.2 mm respectively. The entire bandwidth of UWB
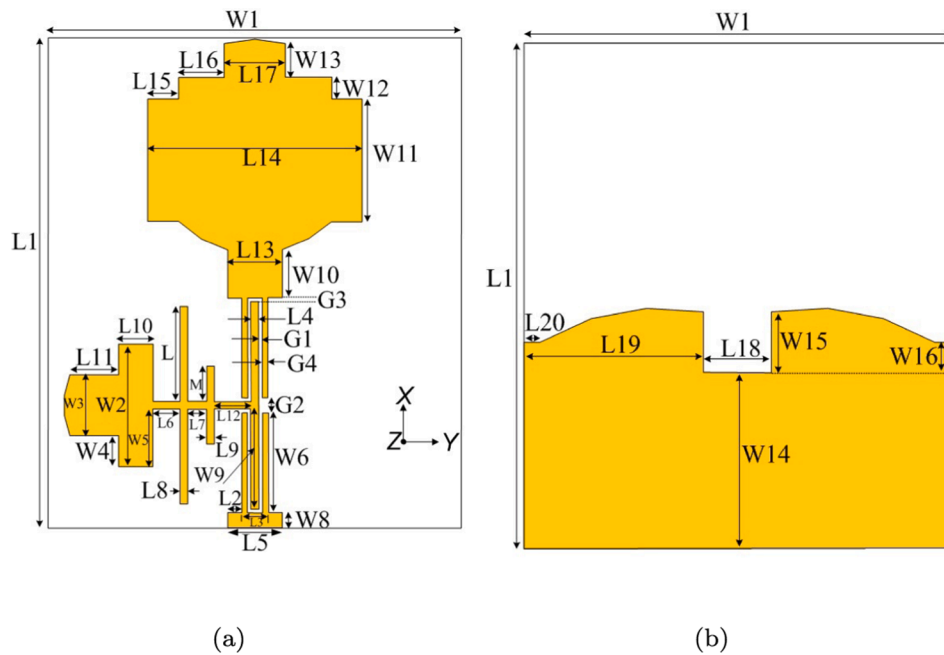
Fig. 15. Layout of the proposed notch-band IFA (a) Top view (b) Bottom view.
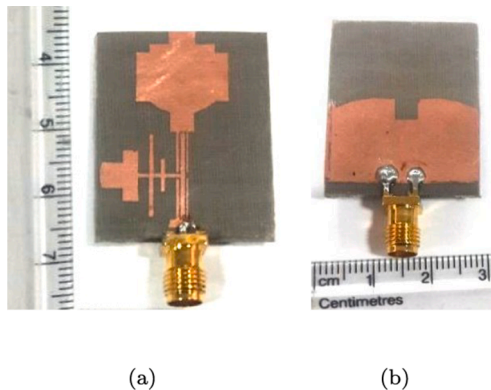
Fig. 16. Fabricated view of the proposed notch-band IFA (a) Top view (b) Bottom view.

Table 3
Parameter values of the proposed notch-band IFA.

Parameters	Values (mm)	Parameters	Values (mm)
L	6.2	L_1	32
L_2	1.05	L_3	1.5
L_4	0.5	L_5	3.6
L_6	1.75	L_7	1.25
L_8	0.5	L_9	0.5
L_{10}	2.26	L_{11}	3.18
L_{12}	2.25	L_{13}	3.6
L_{14}	14	L_{15}	2
L_{16}	3	L_{17}	4
L_{18}	4.3	L_{19}	11.35
L_{20}	1	M	2.2
W_1	27	W_2	8
W_3	4	W_4	2
W_5	3.75	W_6	6.5
W_7	0.5	W_8	1
W_9	6.5	W_{10}	3.13
W_{11}	8	W_{12}	1.4
W_{13}	2.2	W_{14}	11.14
W_{15}	3.86	W_{16}	1.93
G_1	0.1	G_2	1
G_3	0.25	G_4	0.4

is covered by the proposed filter except for a notched band at 6.1 GHz. So the frequency band from 3.139 GHz to 5.69 GHz is the first frequency band with an insertion loss of 0.67 dB and frequency band from 7.045 GHz to 11.02 GHz is the second frequency band with an insertion loss of 2.04 dB. Lower cut-off transmission zero is below −38.01 dB, and upper cut-off transmission zero is below −27 dB. Upper and lower stopband levels are 13.62 dB and 21.59 dB, respectively which, gives a better out of band rejection.

4. Integration of inverted T-shaped UWB bandpass filter with square-shaped-stepped patch UWB antenna

A square-shaped-stepped patch UWB antenna with inverted T-shaped ultra-wideband bandpass filter, called IFA or filtenna having notch-band characteristic is proposed. Fig. 15 shows the design layout of the proposed filtenna, and Fig. 16 shows the fabricated view of the filtenna. A square-shaped-stepped patch is used as a radiating element. Proposed inverted T-shaped UWB filter replaces the feed line of square-shaped-stepped patch UWB antenna and output port of the filter is directly connected to the radiating patch of the antenna. The proposed notch-band IFA is straightforward in design and compact in size. The overall

dimension of the notch-band IFA is 32 × 27 × 1.524 mm³. Optimized parameter values of the proposed notch-band IFA are shown in Table 3.

Distributions of current at frequency 8.5 GHz with angles of 0°, 90°, 180° and 270° for proposed IFA are shown in Fig. 17. It is observed that the surface current distribution from 0° to 270° are traveling anti-clockwise, claiming the Filtenna to have right-hand circular polarization (RHCP) [22]. The currents are in the same phase nearby the center and edges of the square-shaped-stepped patch, which improves the antenna gain. The high current distribution between antenna and filter shows that both are tightly coupled with each other and producing a popular UWB band. Frequencies above from 10.68 GHz are rejected by proposed filtenna confirms that the integrated BPF in proposed filtenna acts perfectly, and it covers the whole range of the UWB band, except a notched band.

The measured and simulated results of the proposed IFA are presented in Figs. 18–22. The return loss of IFA is shown in Fig. 18.

(a)

(b)

(c)

(d)
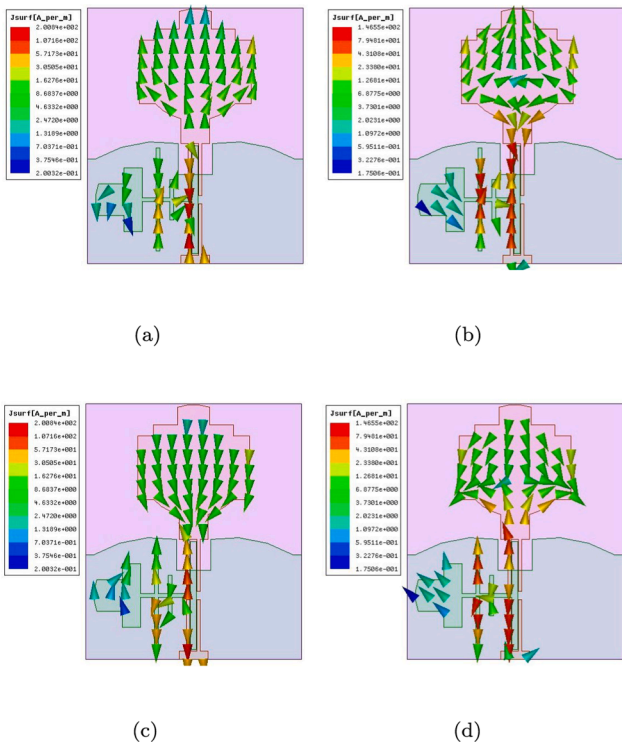
Fig. 17. Current distribution of proposed notch-band IFA at frequency 8.5 GHz with angle (a) 0° (b) 90° (c) 180° and (d) 270°.
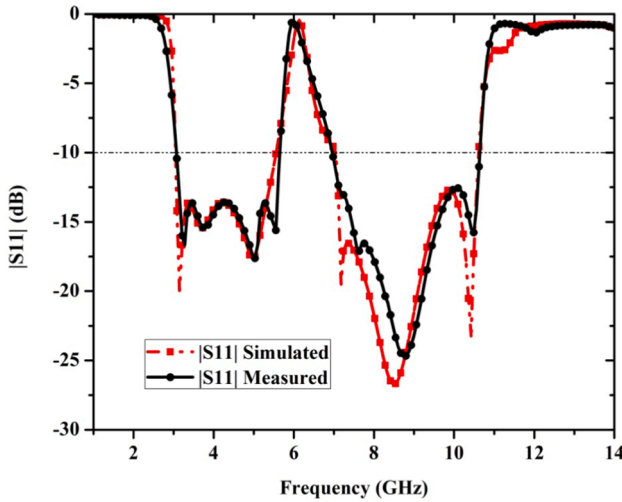
Fig. 18. Measured and simulated |S11| of proposed notch band IFA.

Proposed IFA achieved measured impedance bandwidth from 3.08 GHz to 10.55 GHz and a simulated impedance bandwidth from 3.1 GHz to 10.68 GHz. A notch-band from 5.5 GHz to 6.91 GHz in the simulated result and from 5.55 GHz to 6.98 GHz in measurements is achieved. Measured results are in good agreement with simulated results. The proposed IFA covers the entire UWB spectrum with a notch-band characteristic.

The measured results of the normalized Co- and Cross–polarization far-field radiation patterns in both *xz*-plane and *yz*-plane at frequencies 3.5, 8.5 and 9.5 GHz of the proposed IFA are given in Fig. 19. We can observe isolation of less than 3-dB in the main lobe direction as well as at different angles of elevation (*θ*) between Co- and Cross-polarization in both XZ-plane (E-Plane) and YZ-Plane (H-plane) at the given frequencies which confirms that our proposed IFA is circularly polarized.

(a)

(b)

(c)

(d)

(e)

(f)
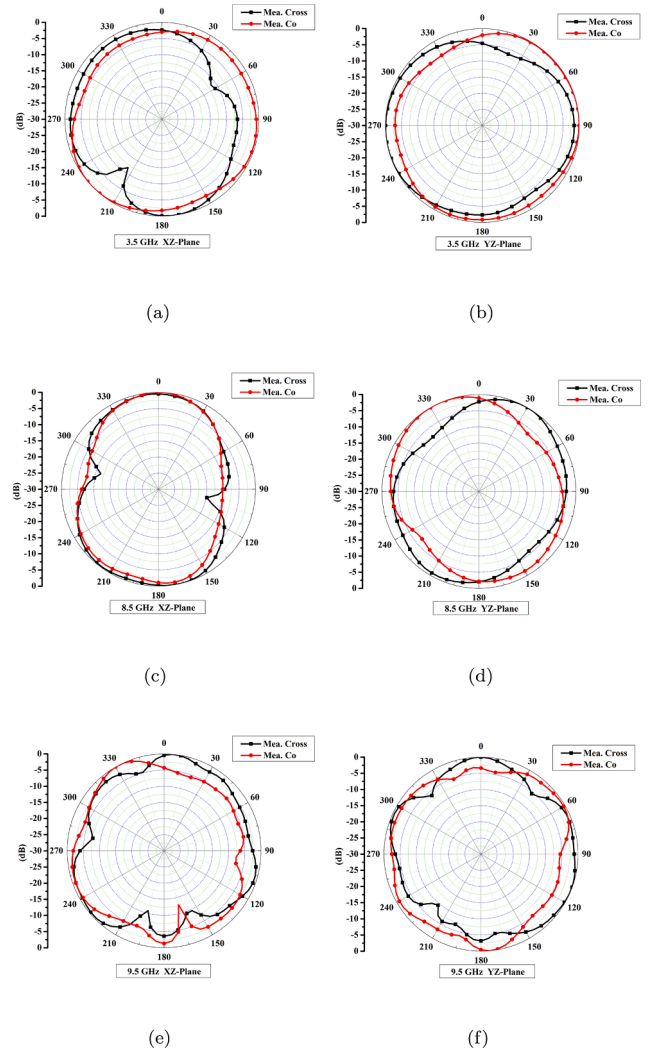
Fig. 19. Measured Co- and Cross polarization radiation pattern of IFA at frequencies (in GHz) (a) 3.5 (E-plane) (b) 3.5 (H-plane) (c) 8.5 (E-plane) (d) 8.5 (H-plane) (e) 9.5 (E-plane) and (f) 9.5 (H-plane).
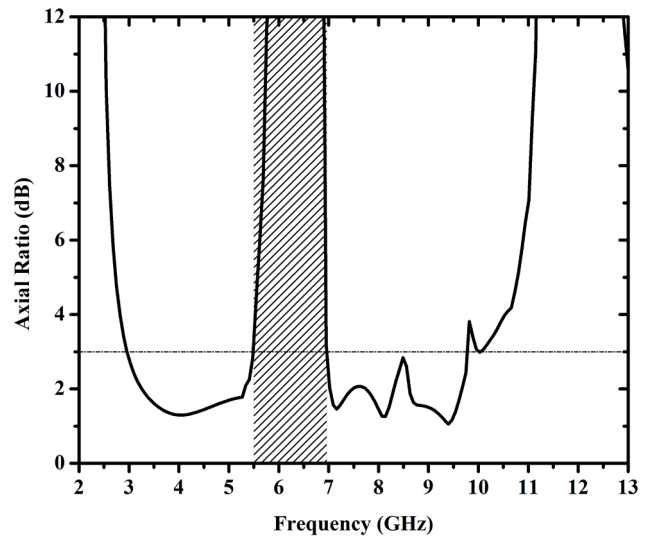
Fig. 20. Measured axial ratio of the proposed filtenna.
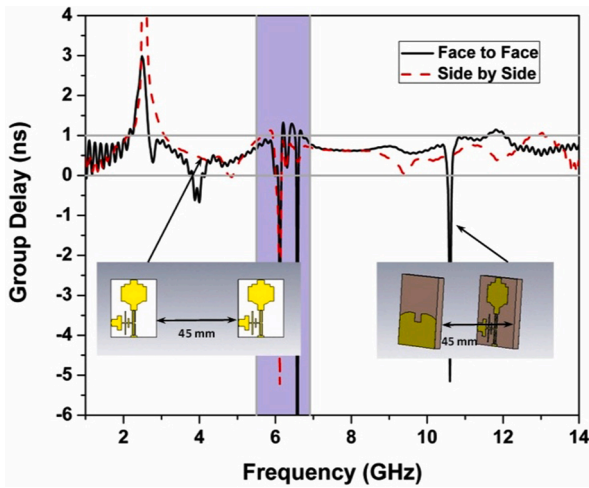
Fig. 21. Group Delay of the proposed filtenna.

The measured result of the 3-dB axial ratio (AR) is plotted in Fig. 20. It is observed that the proposed IFA has almost a UWB performance of circular polarization (except the notched band).

The time-domain analysis of the proposed filtenna is also shown in Fig. 21, where the group delay of the filtenna is calculated against the entire working frequency range. Actually, the same transmitting and receiving filtenna is kept face to face and side by side in the far-field region which is taken as 45 mm (almost two times the Fraunhofer distance) distance to measure the group delay which has come well within the desired range of ± 1 ns [23] for better performance in the operating frequency range. While on the other hand, the group delay is well beyond the desired limits, say up to −6 ns in the band-notched zone from 5.5 GHz to 6.91 GHz advocating poor performance of the filtenna as desired in the band-notched zone. Also, we can observe the sharp change in the group delay after 10.5 GHz, which is the upper cut-off frequency limit of the proposed filtenna.

The measured and simulated gain of proposed IFA and square-shaped-stepped patch UWB antenna are shown in Fig. 22. It can be seen that the measured gain of IFA is high from 3.1 to 5.5 GHz and 7–10.7 GHz. The gain of the proposed IFA varies from 3.58 to 6.8 dBi, excluding the band notched range. It is observed that the gain of a square-shaped-stepped patch UWB antenna is not bounded within the UWB range. Still, IFA achieved out of band gain reduction and notch band characteristics, which is highly desirable for UWB wireless communication system.

As we can see the comparison Table 4, we observe that proposed Filtenna with substrate area, $0.11\lambda_c^2$ is more compact than most of the Filtenna and on par with few an added advantage of interference to DSRC applications because of the wider and controllable band-notched characteristics. Also, our design has a much wider 3 dB-Axial Ratio Bandwidth (ARBW) as compared to all the Filtenna in Table 4, which are linearly polarized.

5. Conclusion

A square-shaped-stepped patch antenna for UWB band and an inverted T-shaped bandpass filter based IFA or filtenna with notched band characteristics are designed and fabricated. The proposed IFA design is straightforward and compact. Proposed filtenna covers the entire range of the UWB band, excluding a notched band. Circularly polarized radiation patterns are achieved in E-plane and H-plane. High peak gain up to 6.8 dBi (measured) is obtained. The gain of proposed IFA is reduced drastically at notch-band and out of the UWB band. Interference in ultra-wideband caused by DSRC (at 5.9 GHz) is avoided by notch-band achieved in the proposed IFA. Proposed IFA has features of
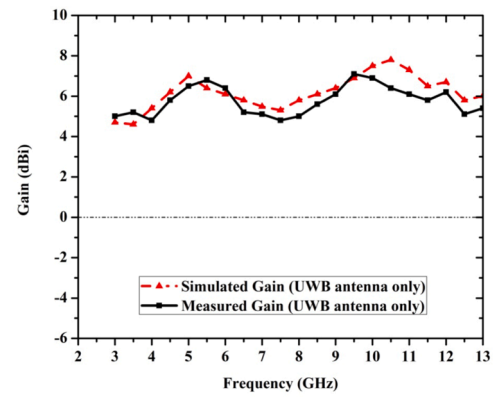
(a)
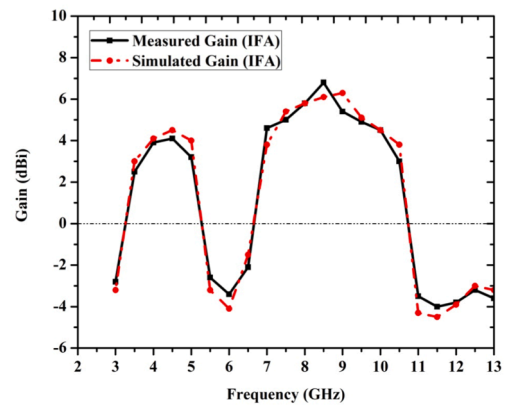
(b)
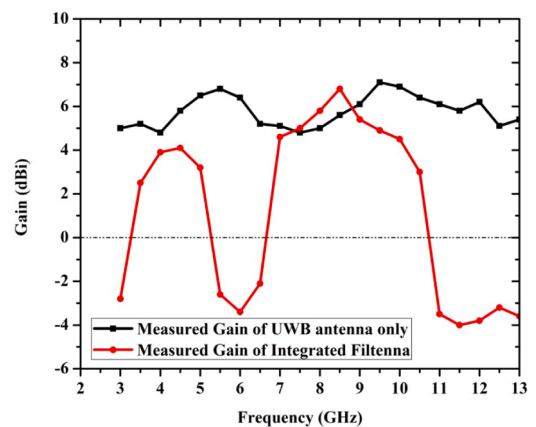
(c)

Fig. 22. Measured and Simulated gain of (a) square-shaped-stepped patch UWB antenna only, (b) Filtenna or IFA and (c)Measured compared gain between square-shaped-stepped patch UWB antenna and IFA.

Table 4
Compared with other related works.

Ref.	ε_r/ Height (mm)	Size$\lambda_c \times \lambda_c$	Peak Gain (dBi)	MSL lower /Upper (dB)	−10 dB BW (GHz)	Notched Band (GHz)	Polarization
[3]	3.38/0.813	$0.74\lambda_c \times 1.11\lambda_c$	9.05	25/25	4.42–5.42	No	Linear
[4]	2.65/2	$2.23\lambda_c \times 0.71\lambda_c$	6	30/20	4.3–4.7	No	Linear
[5]	3.38/0.508	$0.51\lambda_c \times 0.45\lambda_c$	2.5	15/15	3.82–4.49	No	Linear
[7]	3.4/0.8	$0.27\lambda_c \times 0.40\lambda_c$	4.3	NA	4.04–4.26	No	Linear
[9]	2.2/3	$1.11\lambda_c \times 1.11\lambda_c$	8.2	20/20	4.26–5.75	No	Linear
[11]	2.55/0.8	$0.58\lambda_c \times 0.58\lambda_c$	2.5	10.3/10.3	2.5–3.3	No	Linear
[12]	2.55/0.8	$0.29\lambda_c \times 0.29\lambda_c$	3.46	NA	3.1–10.6	4.5–5	Linear
[13]	3.48/1.524	$0.38\lambda_c \times 0.26\lambda_c$	2.13	NA/15	3.1–10.6	No	Linear
[14]	3.38/1.524	$0.31\lambda_c \times 0.26\lambda_c$	9.8	40/40	3.1–15	No	Linear
[16]	2.48/1.57	$0.71\lambda_c \times 0.74\lambda_c$	6.5	26/26	4.35–4.59	No	Linear
[17]	3.55/0.508	$0.29\lambda_c \times 0.28\lambda_c$	7	NA	2.9–10.75	5.01–6.19	Linear
[24]	2.2/5.2	$0.44\lambda_c \times 0.44\lambda_c$	4.7	NA	2.38–2.48	No	Circular
[25]	3.55/0.8	NA	3.3	NA	2.32–2.55	No	Circular
[26]	3.48/0.762	NA	6.4	13.9/10.8	2.35–2.57	No	Circular
This Work	3.38/1.524	$0.38\lambda_c \times 0.28\lambda_c$	6.8	21.59/13.62	3.1–10.68	5.5–6.91	Circular

MSL = Minimum Suppression Levels, NA = Not Available, λ_c= Cut-off Wavelength.

low cost, high gain, circular polarization, and notch-band characteristics, which makes it acceptable for UWB wireless communication systems.

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] Allen Ben, Brown Tony, Schwieger Katja, et al. Ultra Wide- band: Applications, Technology and Future perspectives. In: International Workshop on Convergent Technologies (IWCT); 2005.
[2] Chuang Chao-Tang, Chung Shyh-Jong. Synthesis and design of a new printed filtering antenna. IEEE Trans Antennas Propag 2011;59(3):1036–42.
[3] Hu PF, Pan YM, Zhang XY, et al. A compact filtering dielectric resonator antenna with wide bandwidth and high gain. IEEE Trans Antennas Propag 2016;64(8):3645–51.
[4] Tian Li Wu, Pan Yong Mei, Peng Fei Hu, et al. Design of a low profile and compact omnidirectional filtering patch antenna. IEEE Access 2017;5:1083–9.
[5] Zhang Yao, Xiu Yin Zhang, Yong-Mei Pan. Low-profile planar filtering dipole antenna with omnidirectional radiation pattern. IEEE Trans Antennas Propag 2018;66(3):1124–32.
[6] Chen Yu, WeiHong Zhenqi Kuai, et al. Ku-band linearly polarized omnidirectional planar filtenna. IEEE Antennas Wirel Propag Lett 2012;11:310–3.
[7] Zuo Jianhong, Chen Xinwei, Han Guorui, et al. An integrated approach to RF antenna-filter co-design. IEEE Antennas Wirel Propag Lett 2009;8:141–4.
[8] Ahmadi B, Faraji-Dana R. A miniaturised monopole antenna for ultrawide band applications with band-notch filter. IET Microw Antennas Propag 2009;3(8):1224–31.
[9] Pan YM, Hu PF, Zhang XY, et al. A low-profile high-gain and wideband filtering antenna with meta-surface. IEEE Trans Antennas Propag 2016;64(5):2010–6.
[10] Barbuto Mirko, Trotta Fabrizio, Bilotti Filiberto, et al. Horn antennas with integrated notch filters. IEEE Trans Antennas Propag 2015;63(2):781–5.
[11] Sun Guang-Hua, Wong Sai-Wai, Zhu Lei, et al. A compact printed filtering antenna with good suppression of upper harmonic band. IEEE Antennas Wirel Propag Lett 2016;15:1349–52.
[12] Wong Sai Wai, Huang Tian Gui, Mao ChunXu, et al. Planar filtering ultra-wideband (UWB) antenna with shorting pins. IEEE Trans Antennas Propag 2013;61(2):948–53.
[13] Tang Ming-Chun, Shi Ting, Ziolkowski Richard W. Planar ultrawideband antennas with improved realized gain performance. IEEE Trans Antennas Propag 2016;64(1):61–9.
[14] Ranjan Prashant, Raj Saurabh, Upadhyay Gaurav, et al. Circularly slotted flower shaped UWB filtering antenna with high peak gain. Int J Electron Commun (AEU) 2017;81:209–17.
[15] Nova OA, Bohorquez JC, Pena NM, et al. Filter- antenna module using substrate integrated waveguide cavities. IEEE Antennas Wirel Propag Lett 2011;10:59–62.
[16] Li Peng Kai, You Chang Jiang, Hong Fang Yu. Co-designed high efficiency single-layered substrate integrated waveguide filtering antenna with a controllable radiation null. IEEE Antennas Wirel Propag Lett 2018;17(2):295–8.
[17] Lee Ching-Her, Jhe-Han Wu, Hsu Chung-IG, et al. Balanced band-notched UWB filtering circular patch antenna with common mode suppression. IEEE Antennas and Wireless Propa- gation Letters 2017;16:2812–5.
[18] Ranjan Prashant, Tripathi VS. A compact triple-band microwave filter using symmetrically placed stub-loaded open-loop resonators. Int J Electron Lett 2017;6(3):364–75. https://doi.org/10.1080/21681724.2017.1378419.
[19] Prashant Ranjan. Triple-band and multiband filters using SIL and MSLOR for wireless communication system. Int J Electron Lett 2019. https://doi.org/10.1080/21681724.2019.1695150.
[20] Prashant Ranjan N, Kishore VK, Dwivedi, et al. UWB filter with controllable notch band and higher stop band transmission zero using open stub in inverted T-shaped resonator. In: IEEE Asia Pacific Microwave Conference (APMC), KUALA LUMPUR, Malaysia; 2017. p. 817–20.
[21] ANSYS High Frequency Structure Simulator (HFSS) software. https://www.ansys.com/en-in/products/electronics/ansys-hfss.
[22] Kumar A, Ansari AQ, Kanaujia BK, Kishor J. Dual circular polarization with reduced mutual coupling among two orthogonally placed CPW-fed microstrip antennas for broadband applications. Wireless Pers Commun 2019;107:759–70.
[23] Kumar A, Ansari AQ, Kanaujia BK, Kishor J. An ultra-compact two-port UWB-MIMO antenna with dual band- notched characteristics. AEU - Int J Electron Commun 2020;114.
[24] Jiang ZH, Gregory MD, Werner DH. Design and experimental investigation of a compact circularly polarized integrated filtering antenna for wearable biotelemetric devices. IEEE Trans Biomed Circuits Syst 2016;10(2):328–38.
[25] Cheng Wenbin, Li Dagang. Circularly polarised filtering monopole antenna based on miniaturised coupled filter. Electron Lett 2017;53(11):700–2.
[26] Yunlong Lu, Wang Yi, Gao Steven, et al. Circularly polarised integrated filtering antenna with polarisation reconfigurability. IET Microw Antennas Propag 2017;11(15):2247–52.

☰ IEEE Antennas and Propagation Magazine

🏠 **Home**

✏️ **Author**

💬 **Review**

Author Dashboard

1	Manuscripts with Decisions	›
5	**Manuscripts I Have Co-Authored**	›
	Start New Submission	›
	Legacy Instructions	›
	5 Most Recent E-mails	›
	English Language Editing Service	›

Manuscripts I Have Co-Authored

STATUS	ID	TITLE	CREATED	SUBMITTED
ADM: Tang-Bernas, Christina • Accept as is (09-May-2021) ✉ Contact Journal	APM-Sep-2020-0212.R1	Dielectric Resonator Antennas: Application and Development in Multiple Input Multiple Output Technology Submitting Author: Matekovits, Ladislau Cover Letter	18-Dec-2020	18-Dec-2020
ADM: Andriulli, Francesco ADM: Tang-Bernas, Christina • Major Revision (21-Nov-2020) • a revision has been submitted ✉ Contact Journal	APM-Sep-2020-0212	Dielectric Resonator Antennas: Application and Development in Multiple Input Multiple Output Technology View Submission Submitting Author: Matekovits, Ladislau Cover Letter	23-Sep-2020	23-Sep-2020

Author Notes

✔ The type has been replaced in your figures; please proof the figures carefully for any errors.

✔ The last page contains callouts that might be used in the article, depending on design and spacing considerations. Are these callouts appropriate? If not, please provide several alternatives.

✔ When the magazine is printed, you will receive courtesy copies of the issue to distribute to your coauthors. Please reply with your complete mailing address, including phone number.

✔ This article will be posted to IEEE *Xplore* once your changes are incorporated. *Please note that changes cannot be made once the article has been finalized.*

Sumer Singh Singhwal, Ladislau Matekovits, Binod Kumar Kanaujia, Jugul Kishor, Saeed Fakhte, and Amit Kumar

Dielectric Resonator Antennas

Applications and developments in multiple-input, multiple-output technology.

<AU: Please check whether the edited title and added subtitle are acceptable.>

xxxxx

This article presents a comprehensive review of multiple-input, multiple-output (MIMO) dielectric resonator antennas (DRAs) that have evolved in the past decade. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** In addition to the major challenges faced during designing an MIMO DRA, this article also discusses research gaps that must be filled in the future. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** Exploring the advantages of DRAs, numerous novel designs have been proposed in the last few years.

The major contributions of this article are 1) highlighting and comparing different MIMO DRAs on the basis of performances and applications as well as 2) elaborating the challenges, issues, and research gaps in the emerging field of MIMO DRAs. To compare the performance of MIMO DRAs, published articles concerning MIMO DRAs are divided into five segments:

- The first segment covers ultrawideband (UWB) solutions.
- Circular polarization (CP) solutions are discussed in the second segment.
- In the third segment, single-element compact MIMO DRAs are considered.

- Multiband MIMO DRAs are covered in the fourth segment.
- In the last segment, the remaining articles of MIMO DRAs are discussed. **<AU: Please check whether the preceding edited list conveys the intended meaning.>**

A comprehensive review of all of the latest trends of MIMO DRAs is presented along with seven research gaps/challenges that need to be addressed in the future.

INTRODUCTION

In the past decade, the exponential growth of communication standards has been evident around the globe. Following the principle of the demand–supply of economics, the speed of wireless data transfer and compactness of wireless devices have made a substantial impact on the present world.

The increased growth of wireless communication systems resulted in the development of the MIMO antennas. These systems, with the ability to suppress fading, provide increased coverage and connectivity, high throughput, and low latency, playing a crucial role in the development of various radio-frequency (RF) and microwave communication systems [1], [2]. The 4G and 5G wireless broadband standards depend heavily on MIMO technology.

On the other hand, DRAs have several attractive characteristics, such as no surface wave losses, high radiation efficiency, compact size, nearly constant gain, high impedance bandwidth

(IBW), and ease of excitation. Hence, they are good candidates for various applications across the spectrum, from the microwave to optical frequency bands. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

The current state of the art of DRAs in the microwave regime includes filtering DRAs [3], [4], compact DRAs with omnidirectional radiation patterns [5], wideband CP DRAs [6], [7], anisotropic DRAs [8], [9], RF energy harvesting [10], microwave image sensing [11], characteristic mode analysis of DRAs [12], frequency-tunable DRAs [13], [14], active DRAs [15], [16], and DRA arrays [17]–[19]. A substrate-integrated waveguide (SIW)-fed DRA array is shown in Figure 1 with parasitic DRAs. DRA arrays are different from MIMO DRAs in the following ways: **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

- Usually, a single feed is used in an array, and multiple feeds are used in MIMO.
- Power-divider circuitry is used in an array to feed each element, whereas the same is used in MIMO to achieve polarization diversity.
- Arrays are generally used to achieve high gain, whereas MIMO antennas are preferred to achieve diversity. **<AU: Please check whether the preceding edited list conveys the intended meaning.>**

One of the prominent features that draws attention to DRAs is their high radiation efficiency, which arises from their radiation mechanism being mostly based on displacement currents. Metallic antennas, such as patch and printed dipole antennas, suffer from high conductor losses at the millimeter-wave (mm-wave) range, where the skin effect gets more pronounced. Hence, their radiation efficiency starts to noticeably worsen as the frequency increases in the mm-waveband.

In contrast, as reported in [19]–[22], DRAs indicate higher radiation efficiency, typically above 90%, at the mm-waveband, which demonstrates their potential for applications in this band. Due to the high efficiency of DRAs in the mm-waveband, it is highly desirable to extend their application range in the terahertz regime.

This has received growing attention from researchers in the field of terahertz antennas. For example, in [23], a low-loss on-chip terahertz DRA was reported that has a radiation efficiency of 74% at 341 GHz. Also, to mitigate the size issue at terahertz frequencies, Li and Chiu [24] studied the designs of CMOS DRAs operating at their higher-order modes. At optical frequencies, the electric field penetrates into the metal and couples with the surface plasmons, which results in high dissipation losses. Due to the resilience of DRAs to losses in their ground (GND) plane, acceptable radiation efficiencies still could be obtained in the optical bands. This feature has inspired huge interest in the optics community [25]–[27].

In 2008, Ishmiya et al. [28] proposed a DRA-based MIMO antenna, which has more degrees of freedom for designing and low loss due to the absence of metal in the DRA. The advantages of MIMO technology combined with the benefits of a DRA form a new technologically advanced antenna system called the
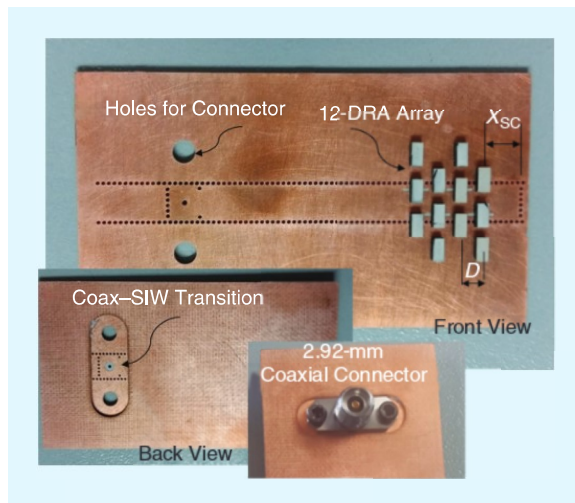
FIGURE 1. An SIW-fed DRA array [17]. *<AU: Kindly spell out X_{SC} and D so the definitions can be added to the caption.>*

DRA-based MIMO antenna. DRAs have also been operating in multimode and radiating power effectively in one direction. Also, CP can be realized in a DRA by using orthogonal modes. Exciting multimode or orthogonal modes in the DRA can be used to create multiple decoupled antenna ports, which can be considered as a basic building block of the MIMO system. **<AU: Please note that the paragraph describing the content of the article has been deleted according to magazine style.>**

CLASSIFICATION OF MIMO DRAs

In the last decade, studies on DRA-based MIMO antennas increased significantly for diverse wireless applications. To review this emerging domain of technology based on performance and characteristics, published articles concerning MIMO DRAs have been grouped into the following segments.

UWB MIMO DRAs

In line with the definition for UWB transmitters given by the U.S. Federal Communications Commission [29], a UWB antenna should have a fractional BW equal to or larger than 0.20 **<AU: Kindly provide units for the preceding value, if applicable.>** or a UWB BW equal to or larger than 500 MHz. In [30], a dual-port half-cylindrical DRA (CDRA)-based MIMO antenna was discussed; the frequency agility was achieved by the addition of reconfigurable parasitic slot loading. A p-i-n diode was used as the switching element. Two orthogonal GND planes were used for generating two modes (TE and HE) **<AU: Kindly spell out TE and HE.>** in the half-CDRA, and three MIMO configurations with three different frequencies (622, 700, and 780 MHz) were proven by varying the switching state of the p-i-n diode and the value of the capacitance.

A simple rectangular DRA (RDRA) with two ports was proposed in [31] for LTE applications. A coplanar waveguide (CPW) feed with an inductive slot and a coaxial probe feed were applied on an RDRA to excite the $TE_{\delta21}^x$ and $TE_{2\delta1}^y$ modes, respectively. A metallic strip was used to connect the
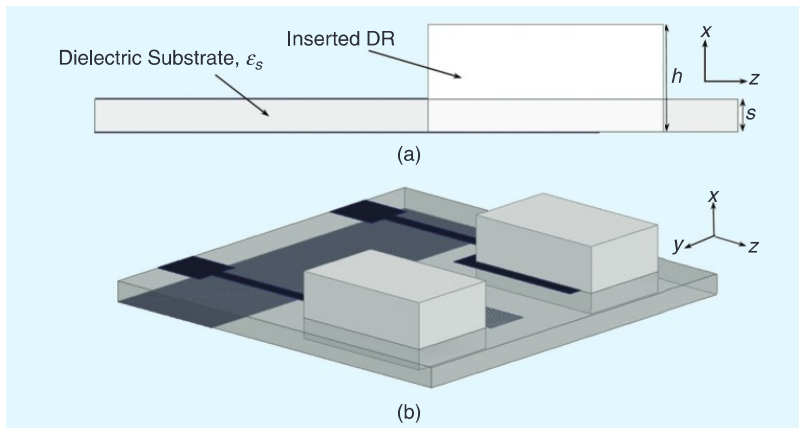
FIGURE 2. A UWB MIMO DRA with 106% fractional BW. (Source: [37].) *<AU: Kindly provide a caption for parts (a) and (b). Do you have permission from the source to use this figure?>*

DRA to the coaxial probe to make the antenna operate at 2.6 GHz. The dimensions of the metallic strip were found by parametric analysis.

In [32], four $HE_{11\delta}$ modes were excited within the DRA by slots, which were etched out of the GND plane of the FR4 *<AU: Kindly spell out FR4.>* substrate onto which the DRA was mounted. Each of these slots was excited by two 50-Ω microstrip (MS) transmission lines; 7-dB IBW was reported, which may be useful in some low-power-requirement applications.

Two F-shaped DRAs were proposed in [33] for an MIMO DRA with multiple GND planes. The use of multiple GND planes instead of the connected GND plane does not fall under the definition of MIMO antennas. The use of multiple GND planes is not a practical choice since, in a real system, the signals should have a common reference plane, i.e., a single common GND plane, so that all signal levels within the system can be interpreted properly based on that reference level (i.e., 0 V or the GND level). If separate GND planes are used, one cannot guarantee that the system will work since the assumption of having all GND planes with the same voltage level is invalid [34].

In [35], two mushroom-shaped DRAs, made up of RT/Duroid *<AU: Kindly spell out RT, if possible.>* 6010 ($\epsilon_r = 10.2$, and $\tan\delta = 0.002$), were excited by a conformal trapezoidal patch and form an MIMO antenna for WB applications. DRAs were placed orthogonal to each other to achieve polarization diversity. The peak gain was achieved from 3.34 to 7.4 dBi throughout the IBW. The mushroom-shaped DRA was formed by adding an RDRA and half-split CDRA on the top of it. This DRA resonates at three different frequencies (5.37, 6.95, and 9.58 GHz) corresponding to three modes: TE_{111}, TE_{113} and TE_{115}, respectively. Here, the matching and excitation of different modes that merge to offer a wider BW were provided by the conformal patch.

A single cross-shaped, DRA-based, dual-band MIMO antenna for GPS and LTE-band applications was proposed in [36]. A cross-shaped DRA with carved corners and a cylindrical slot in between the DRA resulted in BW enhancement. A slot on the GND plane between two feed points altered the current path

and reduced mutual coupling. To add frequency agility in the design, p-i-n diodes were used in the GND plane.

A UWB MIMO DRA with wide local area network (WLAN)-band rejection was proposed in [37], as shown in Figure 2. Two RDRAs were excited using MS feed lines, and a stub connected to the GND plane helps to achieve a wider BW with acceptable isolation throughout the BW. One more structure, in the form of an L-shaped parasitic strip connected to RDRAs on the top of the substrate, is said to claim band rejection in the WLAN band from 4.98 to 6.08 GHz by providing *strong disturb*ance of the DRA modes, leading to a strong impedance mismatch near the notch frequency. However, the realized gain varies from 1.1 to 5.3 dBi throughout the band except the band rejection, where the gain drops to –1.64 dBi. In this work, an RDRA was inserted into the Taconic RF-35 substrate with a thickness of 1.524 mm, such that it touched the GND and MS line feed above the substrate.

In [38], the authors proposed a dual-element, fractal T-shaped, DRA-based MIMO antenna for WB applications. Here, a C-shaped periodic defective GND structure (DGS) was applied to achieve isolation between ports. In [39], a triangular-shaped DRA excited by three probe feeds was proposed for switch-beam operation. It focused on the beam-steering operation in WLAN applications; the radiation pattern is tuned by switching the angular excitation port.

The authors of [40] used the conformal rectangular strip for excitation, whereas, in [35], a conformal trapezoidal-shaped strip was used to excite the mushroom-shaped DRA. However, an A-shaped DRA design was taken from equilateral-triangular DRA having dimensions corresponding to the fundamental TM_{101} mode. The DRA was made up of RT/Duroid 6010 with eight stack layers to achieve the desired height. A 20-dB isolation was realized by introducing sufficient interelement spacing between the two elements.

Two RDRAs were used in [41] to build a dual-port MIMO antenna for LTE applications. A T-type matching circuitry was used to match the antenna on the serial resonance of the first resonant mode at 800 MHz. Probe-fed RDRAs were placed diagonally for higher isolation.

A self-complementary hybrid ring DRA (RiDRA)-based MIMO antenna was proposed in [42]. Two RiDRAs, made up of alumina ceramic ($\epsilon_r = 9.8$, and $\tan\delta = 0.002$), are excited by using a stepped slot-based GND and MS line above the FR4 substrate ($\epsilon_r = 4.4$, and $\tan\delta = 0.02$). The slot behaves as a resonating structure whose frequency is associated with the dimensions of the slot, mainly its length. Here, WB IBW is achieved by merging the dual bands developed due to the stepped slots and two bands corresponding to the $HE_{11\delta}$ and $TE_{01\delta}$ modes.

A two-element MIMO antenna with dual-band capability was proposed in [43]. It included two CDRAs placed on an

annular ring printed line, which acted as a magnetic dipole and was responsible for $HEM_{11\delta}$ <AU: Kindly spell out HEM.> and $TE_{01\delta}$ in the CDRAs. A slit in the GND plane was used to reduce mutual coupling in the lower operating frequency band. To improve isolation in the upper band, two radiators were arranged orthogonally.

The authors of [44] proposed a two-CDRA-based, triple-band MIMO antenna with a Y-shaped MS line feed. To increase isolation between the ports, two metallic strips and a slot on the GND plane were used. It was dictated that the Y-shaped feed act as a short magnetic dipole to excite the $HEM_{11\delta}$ mode in the CDRA. Moreover, it was shown that the lower band response is produced due to the metallic reflector placed on the GND plane, midband response is produced due to the $HEM_{11\delta}$ mode in the CDRA, and upper band response is produced due to the Y-shaped printed line. In addition to the triple band, this antenna also exhibits CP in the upper band, which was due to the Y-shaped line's ability to produce orthogonal modes.

In [45], a triband, DRA-based MIMO antenna was proposed for WLAN/WiMax applications. Interestingly, it fulfilled the criteria for UWB given by [29] in all of the three bands, with the fractional BWs corresponding to 0.51, 0.21, and 0.26 <AU: Kindly provide units for the preceding values, if possible.> in their respective bands. It was claimed that the $HEM_{11\delta}$ and $HEM_{12\delta}$ modes were generated due to the CDRA and T-shaped printed line, which acts as an electric dipole. Additionally, a circular ring-shaped printed patch of the size of the CDRA base was found accountable for the TM_{11} mode. To enhance isolation between the ports, the concept of a slitted GND plane was used in the antenna.

A two-element, DRA-based MIMO antenna was proposed in [46]. Here, the DRA was constructed using stacking of three RDRAs. It exhibited around 46% fractional BW, but only the simulated results were discussed.

Two RiDRAs excited using a U-shaped printed line were proposed in [47] for MIMO applications in the ultrawide range of the BW (3–7 GHz). To improve isolation between the ports, a dual L-shaped DGS and rectangular DGS were carved from the GND plane.

A WB CP MIMO antenna with two RDRAs made up of Eccostock Hik bar ($\epsilon_r = 10$, and $\tan\delta = 0.002$) was presented in [48]. The diagonal placement of RDRAs gives rise to high isolation. Each RDRA was excited by an E-shaped conformal copper strip, which, in turn, was excited using the probe feed. It shows a wide axial ratio BW (ARBW) along with a wide IBW. The wide ARBW was achieved by the optimization of a conformal patch and diagonal position of both of the RDRAs. It was claimed that the degenerate mode pair $TE^{x}_{\delta13}$ and $TE^{y}_{1\delta3}$ was excited to generate the CP radiated wave.

Chen et al. [49] used an electromagnetic bandgap (EBG) surface along with a DRA to obtain a CP MIMO radiator for 5G applications. Here, a two-diagonal-edge truncated DRA was excited by a cross-ring slot to generate CP fields. In this work, for the first time, an EBG surface was used in a DRA-based MIMO antenna. Small triangular stands at the edges of the DR were utilized to hold it, avoiding any glue, which can deteriorate
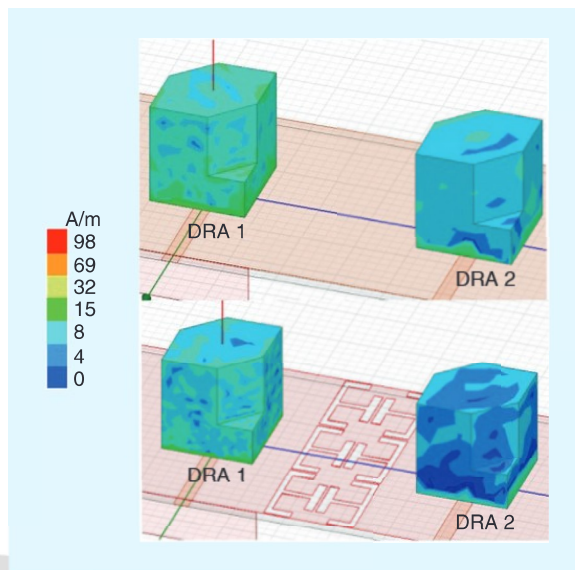
FIGURE 3. A surface current distribution comparison of an MIMO DRA with and without an EBG [49].

its performance. The optimized planar EBG structure was used to reduce mutual coupling between the two truncated DR elements. A surface current distribution comparison of an MIMO DRA with and without an EBG is shown in Figure 3.

The authors of [50] proposed a single-element, dual-port MIMO antenna for WLAN applications. The geometry of the antenna is shown in Figure 4. The shape of the DRA was obtained by merging an L-shaped DRA and CDRA at its center. This compact antenna exhibited high gain but moderate isolation between ports. Gain enhancement and IBW broadening were achieved by etching vertical copper strips on the surface of the DRA. For better impedance matching and tuning, four air gaps with optimized radii were utilized in the antenna.

In [51], a dual-band, two-element MIMO DRA was proposed with CP in the upper band. Two RiDRAs were excited using a plus-shaped aperture feed, and a DGS in the GND was used to achieve high isolation. In [52], two I-shaped DRAs were excited by circular aperture-coupled feeds, and high isolation was achieved by utilizing meander lines in between the ports.

A stacked CDRA-based, dual-band antenna was proposed in [53] for WLAN applications. Here, broadside radiation patterns were achieved by exciting $HEM_{11\delta}$ and $HEM_{12\delta}$ in two stacked CDRAs of different dimensions. Isolation between the ports was achieved by utilizing spatial diversity. In [54], the authors proposed the concept of CP adaptability in a DRA-based MIMO antenna for sub-6-GHz applications. Two CDRAs were excited using a conformal probe and L-type feed network, which resulted in CP fields. RiDRAs have also been used to show the concept of changing CP BW with varying inner shells of the DRA.

The authors of [55] reported a dual-port, CP MIMO DRA for C-band applications. A single DR element was formed by combining an RDRA and a half-CDRA, whereas two elements were joined together with no space in between them to achieve compactness. This resembles an elliptical-shaped DRA. Each

single DR element of the antenna operates with the $TM_{11\delta}$ and $TM_{21\delta}$ modes.

A two-port, compact wideband, quarter-CDRA (Q-CDRA) with CP was proposed in [56] for MIMO applications. The geometry of the antenna is shown in Figure 5. It consists of two Q-CDRs facing each other, both excited by a quasi-spiral MS transmission line to achieve CP. The image theory is applied by placing electrical and magnetic walls to reduce the size of the CDR.

A dual-port, dual-CP DRA with a wide ARBW (43.5%) was reported in [10] for electromagnetic energy harvesting. A series feeding structure that consists of four crossing slots and one MS line was used to excite the DRA. However, since the reported antenna is used for receiving energy, the performance of the antenna is not affected by the relatively low port isolation of 10 dB.

CP MIMO DRAs

CP antennas are preferred over their linearly polarized counterparts because of advantages such as the ease of reception and higher reliability of the signal. For MIMO antenna systems, CP antennas are preferred to increase the diversity performance. UWB CP MIMO DRAs [32], [44], [48], [49], [51], [54]–[56] were discussed in the preceding section.

In [57], a cross-shaped DRA with a broadside CP and vertically polarized omnidirectional radiation patterns was proposed by exciting two different modes within the DR, as shown in Figure 6. The size of the DR was reduced by using thin plates of high-permittivity material $(\epsilon_r = 50)$. A sequential rotation feed

technique was utilized to increase the symmetry of the modes and reduce coupling between the two ports.

The authors of [58] presented a simulated antenna without verification with the fabricated prototype's measured results. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** In this work, a corner truncated V-shaped DRA was excited by two MS line feeds to generate CP fields in the antenna.

Sahu et al. [59] proposed an L-shaped-based, CP MIMO DRA for WLAN applications. Two L-shaped DRAs were placed self-complementarily and excited using a coaxial probe feed arrangement. Two self-complementary, C-like-shaped DGSs were applied to achieve the desired isolation between ports. Two degenerate orthogonal modes ($TE^x_{\delta 11}$ and $TE^y_{1\delta 1}$) were excited in the DRA, using the optimized position of the coaxial probe.

Das et al. [60] proposed a dual-port, dual-CDRA-based CP MIMO antenna for WLAN applications. CP was generated by using a modified circular-shaped aperture slot. To achieve isolation between the ports, the number of DGS slits was applied in the GND plane. Due to perturbation of the circular-shaped aperture, two degenerate orthogonal modes ($HE^x_{11\delta}$ and $HE^y_{11\delta}$) were generated in the CDRA to produce CP fields. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

A dual-band MIMO DRA with an ARBW in both of the operating bands was proposed in [61]. The geometry of the antenna is shown in Figure 7. An arc-shaped line feed network
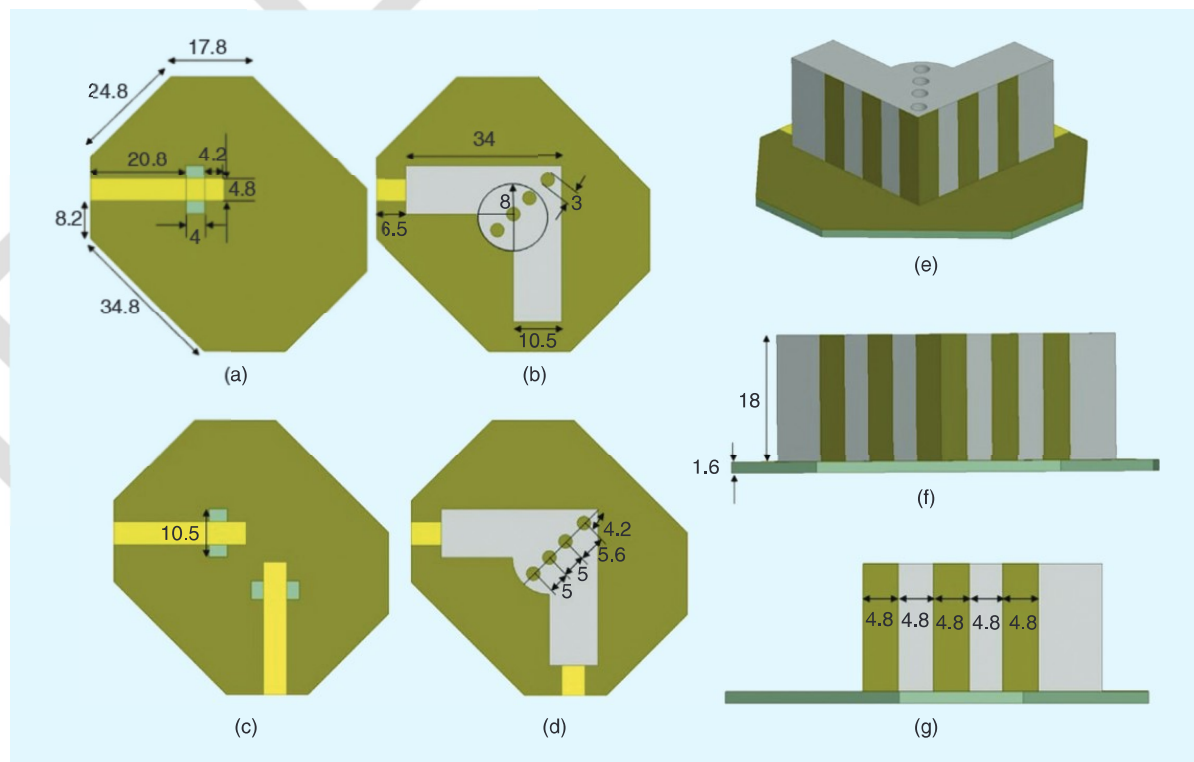
FIGURE 4. The geometry of an MIMO DRA with copper strips for gain enhancement. (Source: [50].) *<AU: Kindly provide a caption for parts (a)–(g). Do you have permission from the source to use this figure?>*

was used to achieve CP and impedance matching, whereas the slit in the GND plane was used to achieve isolation between the ports. Low correlation was achieved by polarization diversity and spatial diversity.

SINGLE-ELEMENT MIMO DRAs

The compactness of the antenna is one of the desirable properties in applications such as satellite and mobile communications. Compactness is one of the major attractions in multiport MIMO antennas. A small size can be chosen using a high-permittivity DR material, as permittivity of the DR is inversely related to its dimensions. However, a high-permittivity DR lacks the property of field leakage, which affects its radiation properties as an antenna. In view of this, a single-element MIMO DRA is one of the thrust areas that requires attention. In previous sections, some single-element MIMO DRAs have already been discussed [30]–[32], [36], [50], [57], [58]. Figure 8 shows the geometry of a dual-band cross-shaped compact MIMO DRA [36]. Some other geometries are discussed in this section.

A half-split CDRA with two orthogonal GND planes was chosen with two feeds in [62]. A mode-degeneration method derived from perturbation theory was proposed to make the TE and HE modes of the split-CDRA resonate at the same frequency. In addition, very low mutual coupling between the two ports as observed due to the presence of orthogonal GND planes. In [63], a CDRA excited by two orthogonal probe feeds for MIMO applications was discussed. Here, a simple frequency-domain channel capacity analysis method was derived for the DRA-based MIMO systems, which directly used complex S parameters for calculating the normalized correlation matrix eigenvalues.

A dual-mode DRA with magnetic- and electric-monopole-like radiation patterns was proposed in [64] by exciting the ($TE_{01\delta}$ and $TM_{01\delta}$) modes in a single annular CDR. Two groups of four radially arranged MS feeding lines with two different lengths were utilized to increase the BW of the TE mode and suppress the influence of higher-order modes.

A dual-band and dual-polarized CDRA made up of lower-permittivity material (K9 <AU: Kindly spell out K9, if possible.> glass: $\epsilon_r = 6.85$) was investigated in [65]. The fundamental mode (HEM_{111}) and higher-order mode (HEM_{113}) of the CDRA were excited by the strip- and slot-fed excitation methods. To achieve impedance matching, a pair of unequal arc-shaped slots was used for one port, whereas an impedance transformer was inserted in the MS feedline for the second port.

An RDRA was fed with two MS line feeds with a conformal patch in [66]. The top face of the DR was covered with a square copper patch. Two orthogonal modes ($TE_{\delta 11}$ and $TE_{1\delta 1}$) were excited inside the DR to achieve high isolation. The vertical metal strips

were connected with the MS line to improve impedance matching.

The design formulas for a three-port CDRA-based MIMO antenna were discussed in [67]. <AU: Please check whether the preceding edited sentence conveys the intended meaning.> The $TM_{01\delta}$ mode of the DRA was excited using an axial coaxial
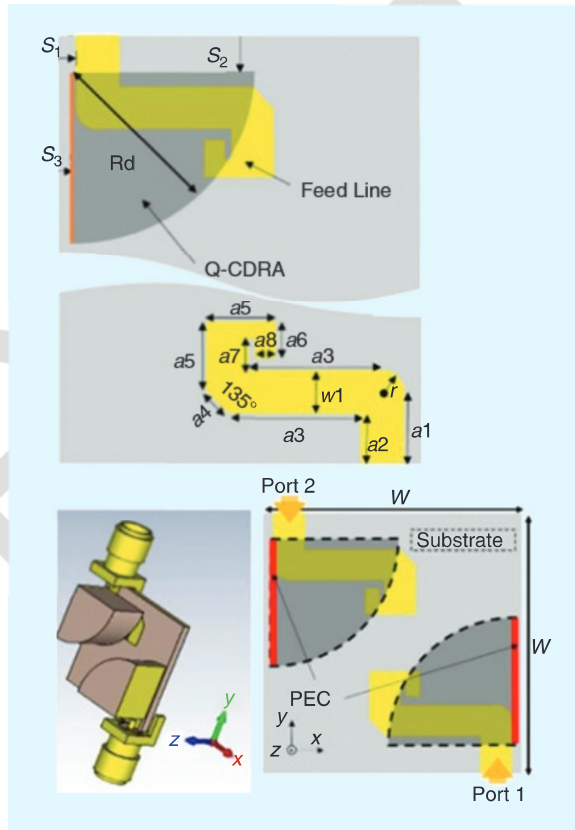
FIGURE 5. A compact wideband Q-CDRA-based MIMO antenna. (Source: [56].) *<AU: Kindly spell out Rd, S, a, w, W, r, and PEC so the definitions can be added to the caption. Do you have permission from the source to use this figure?>*
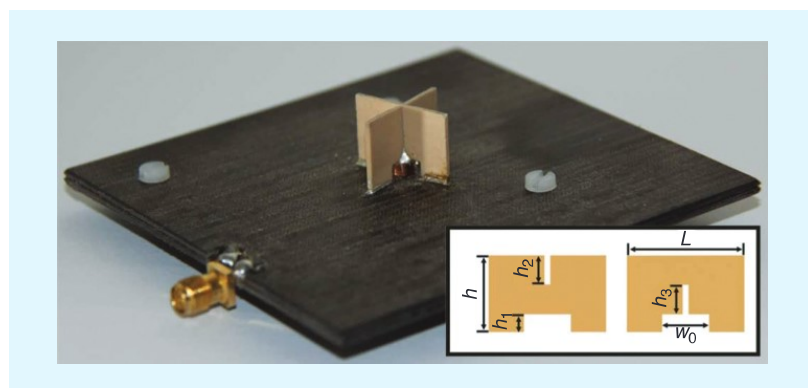
FIGURE 6. A high-permittivity ($\epsilon_r = 50$), cross-shaped antenna, where *h* indicates height, *w* represents width, and *L* is the length [57]. *<AU: Kindly check that h, w, and L are defined correctly.>*
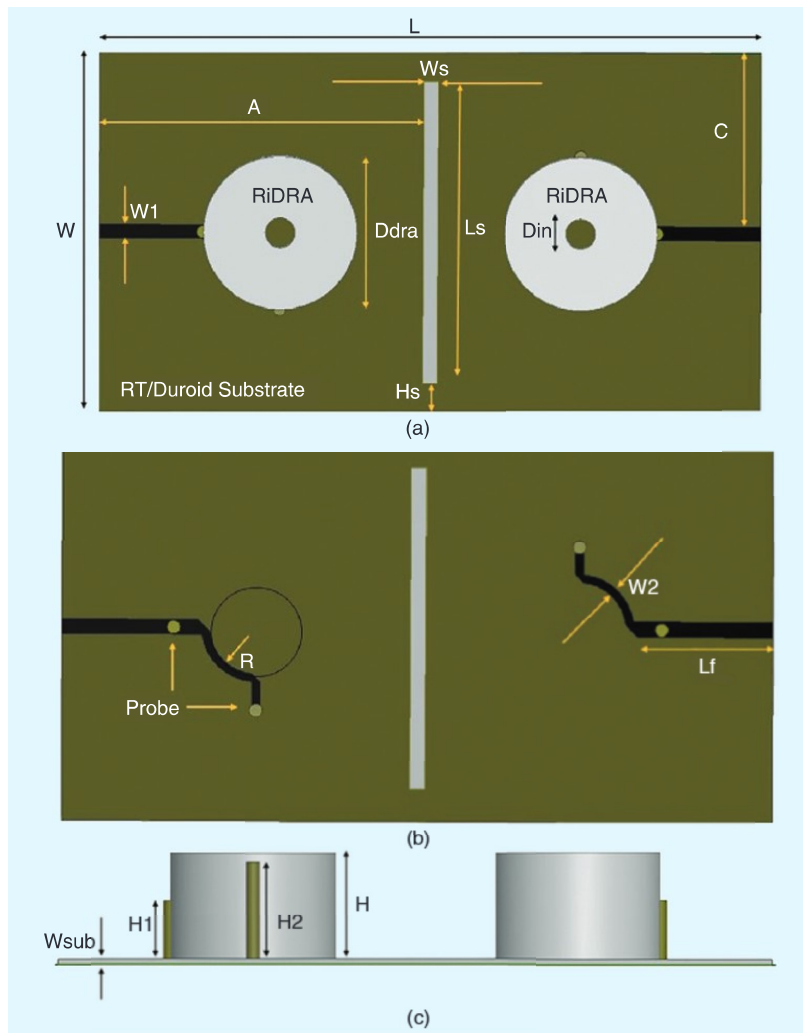
FIGURE 7. The geometry of an MIMO DRA with a CP BW in both of the operating bands. (Source: [61].) *<AU: Kindly spell out Ws, RT, A, C, Ls, Din, Hs, R, Lf, W1, W2, Wsub, H1, and H2 so the definitions can be added to the caption. Please provide captions for parts (a), (b, and (c). Do you have permission from the source to use this figure?>*
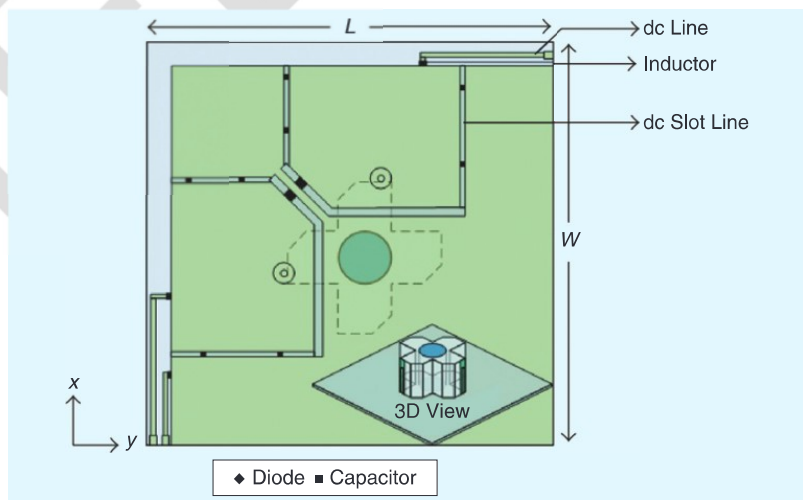
FIGURE 8. A dual-band, cross-shaped compact MIMO DRA. (Source: [36].) *<AU: Do you have permission from the source to use this figure?>*

probe, whereas each of the orthogonal $HEM^x_{12\delta+1}$ and $HEM^y_{12\delta+1}$ modes was excited by a pair of balanced slots. These balanced slots of both of the ports were fed by using two Wilkinson power dividers printed on the substrate. In [68], two annular RDRAs stacked on each other were excited by two probe feeds. Here, only simulated results were discussed, which have acceptable values for MIMO operations.

In [69], a single RDRA element was housed in a thin FR4 substrate and fed by two MS feed lines. Both the feeding lines excited the $TE^x_{\delta11}$ mode in the DR. Two slits in the GND plane were introduced to decrease the mutual coupling between the ports. In [70], an L-shaped DR element with a cylindrical air gap was excited using two slot-coupled feed lines. The cylindrical air gap was introduced to provide isolation between both of the ports. To improve impedance matching, two copper strips were etched at the back of the DR element.

A compact dual-band MIMO stacked DRA was proposed in [71] for WiMAX/WLAN applications. Stacking is used to make the design compact, and the DGS technique was applied to achieve high isolation between the antenna ports. For both of the ports, the excited modes radiate like a short magnetic dipole. Two coaxial probes of equal heights were used to excite two different modes in the DR that radiate like a short magnetic dipole.

In [72], an RDRA was excited using two feeds—the probe and aperture-coupled feeds—resulting in a dual-band antenna for LTE applications. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** However, in this work, multiple GNDs were taken for two RDRAs. In [73], a cross-shaped DRA was excited using a probe feed and an aperture-coupled feed. A rectangular aperture-coupled feed as used to excite the $TE^x_{\delta11}$ and $TE^x_{\delta13}$ modes, whereas the probe port feed was used to excite the $TE^y_{1\delta1}$ and $TE^y_{1\delta3}$ modes at 3.5 and 5.25 GHz, respectively. A metallic strip was also etched on the surface of the DRA to achieve better impedance matching.

An RDRA excited using three ports was proposed in [74] for X-band applications. Here, the method to generate three mutually decoupled modes was discussed, where the decoupling between two of the modes ($TE^x_{\delta np}$) was achieved by minimizing the spatial overlapping of field magnitudes instead of exciting mutually perpendicular field components. The third mode, $TE^y_{m\delta p}$, was excited using the probe feed, where the field components of that $TE^y_{m\delta p}$ mode were perpendicular to those of the $TE^x_{\delta np}$ mode, which was needed to excite the three mutually decoupled modes.

A differential dual-band, dual-polarization DRA was investigated and realized in [75] using the cross-shaped DR fed by two same pairs of differential schemes. The TE_{111} and TE_{113} modes were utilized to obtain two resonant frequencies for the dual-band response. In [76], a CDRA was excited using two aperture-coupled feeds. During port 1 excitation, a power divider was used to excite the DR through two narrow, annular-shaped slots in the GND plane. These slots behaved like a magnetic dipole, which resulted in $HE^y_{11\delta}$ mode generation in the DRA. During port 2 excitation, a simple MS line was used with an annular-shaped aperture in the GND.

An L-shaped DRA excited using two probe feeds was proposed in [77] for dual-band LTE applications. A DGS slot between ports was applied to increase isolation. The proposed MIMO antenna also demonstrated high throughput for different modulation schemes. In [78], an equilateral-triangular DR element was excited by two parallel MS line feeds with conformal strips on the DRA. To achieve higher isolation between feeds, the front edge of the DR element was perturbed in the form of semicylindrical geometry, as shown in Figure 9.

In [79], a WB dual-polarized metal loop dielectric resonator magnetoelectric (MLDR-ME) dipole antenna with low backward radiation and symmetrical E- and H-plane radiation patterns for a broadside array was proposed. It comprised two perpendicular metal semiloop radiators and a cross-shaped dielectric resonator. The metal semiloops served as the feeding structure to the two orthogonal modes in the dielectric resonator and the two magnetic dipole antennas, whereas each of the resonant modes in the dielectric resonator created an electric dipole. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

In [80], a differential dual-polarized filtering DRA using a four-leaf-clover-shaped DR as the radiator was proposed. Two pairs of orthogonal modes with out-of-phase field distributions in the DR are excited by T-shaped feeding strips. The T-shaped feeding strips also facilitate integration applications by offering flexible input impedance for the radiation part. After integrating a pair of high-performance filtering baluns on the back of the reflecting GND, the proposed DRA also shows high-frequency selectivity for each polarization without increasing the antenna size.
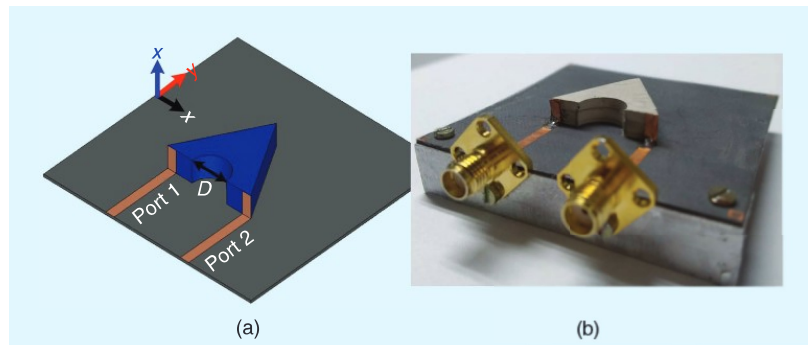
FIGURE 9. A stacked and perturbed equilateral-triangular-shaped MIMO DRA [78]. (Source: [78].) *<AU: Please provide a caption for parts (a) and (b). Do you have permission from the source to use these images?>*

MULTIBAND MIMO DRAs

Multiband antennas have always attracted researchers due to their diverse commercial applications in different bands simultaneously. Features of multiband geometries [41], [43]–[45], [51]–[53], [61], [65], [68], [70]–[73], [77] have already been discussed in previous sections, so the remaining publications are reviewed here. Figure 10 shows the geometry of a triple-band, CDRA-based MIMO antenna [45].

Eight CDRAs were excited by eight feeds in [81] to make a dual-band MIMO antenna. One group of four CDRAs covered
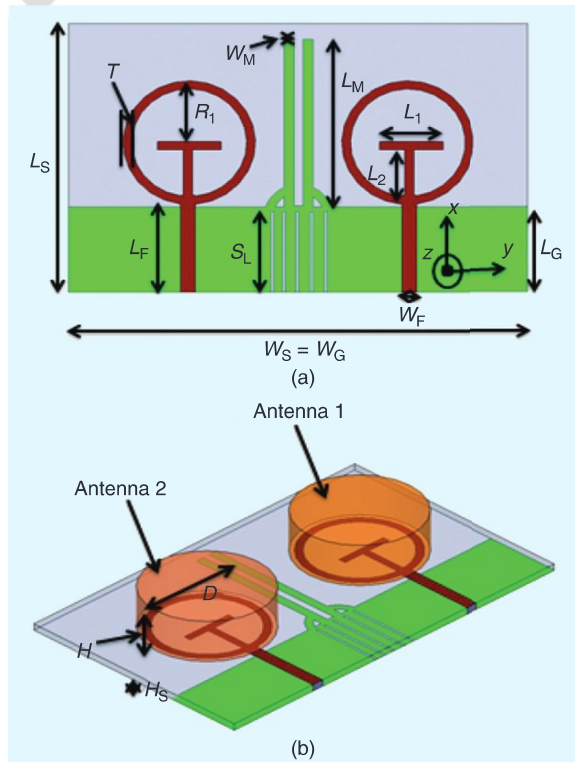
FIGURE 10. The geometry of a triple-band MIMO antenna. (Source: [45].) *<AU: Please provide a caption for parts (a) and (b). Kindly define all variables in the images so the definitions can be added to the caption. Do you have permission from the source to use this figure?>*

the 2.45-GHz band, whereas the other group of four CDRAs covered 5.8-GHz band. The isolation between the lower and upper band elements was better than 15 dB. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** At the lower and higher bands, the $HEM_{11\delta}$ and $TE_{01\delta}$ modes were excited, respectively. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** A center metallic reflector element to reduce the correlation coefficients via tilting the fields and, thus, improving the system operation was also introduced.

In [82], a four half-split CDRA-based, dual-band MIMO antenna was proposed for LTE and WLAN applications. The dual-band operation was achieved by simultaneously exciting the $TE_{01\delta}$ and $TE_{02\delta}$ modes in an aperture-coupled half-split CDRA. Four elements were arranged in a 2×2 format and also orthogonally to each other to achieve better radiation characteristics.

In [83], two RDRAs were excited using an MS line feed followed by meander lines. However, in that work, multiple GNDs were taken for two feeds, which contradicts the MIMO definition given in [34]. In [84], the authors proposed a partial reflector surface structure with varying circular patches on it to improve the isolation between CDRAs for multiband MIMO applications. The geometry of the antenna is shown in Figure 11.

In [85], an MIMO antenna was proposed with four RiDRAs excited by an aperture-coupled feed. Due to the use of polarization and space diversity, the isolation value is more than 20 dB for all of the ports. A pentagon ring-shaped aperture excited the dual modes $HE_{11\delta}$ and $HE_{12\delta}$ in the ring ceramic resonator.

OTHER MIMO DRAs

This section is dedicated to those published articles relating to MIMO DRAs that do not fall into any of the categories discussed in the previous sections. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** In [86], a dual-port, eight-CDRA-based MIMO antenna was proposed for handheld devices. There were two arrays of four CDRAs each, which were excited by an aperture-coupled feed. A passive MS line-based feed network was used to excite the array elements from two ports. In [87], four arrays with four elements each were
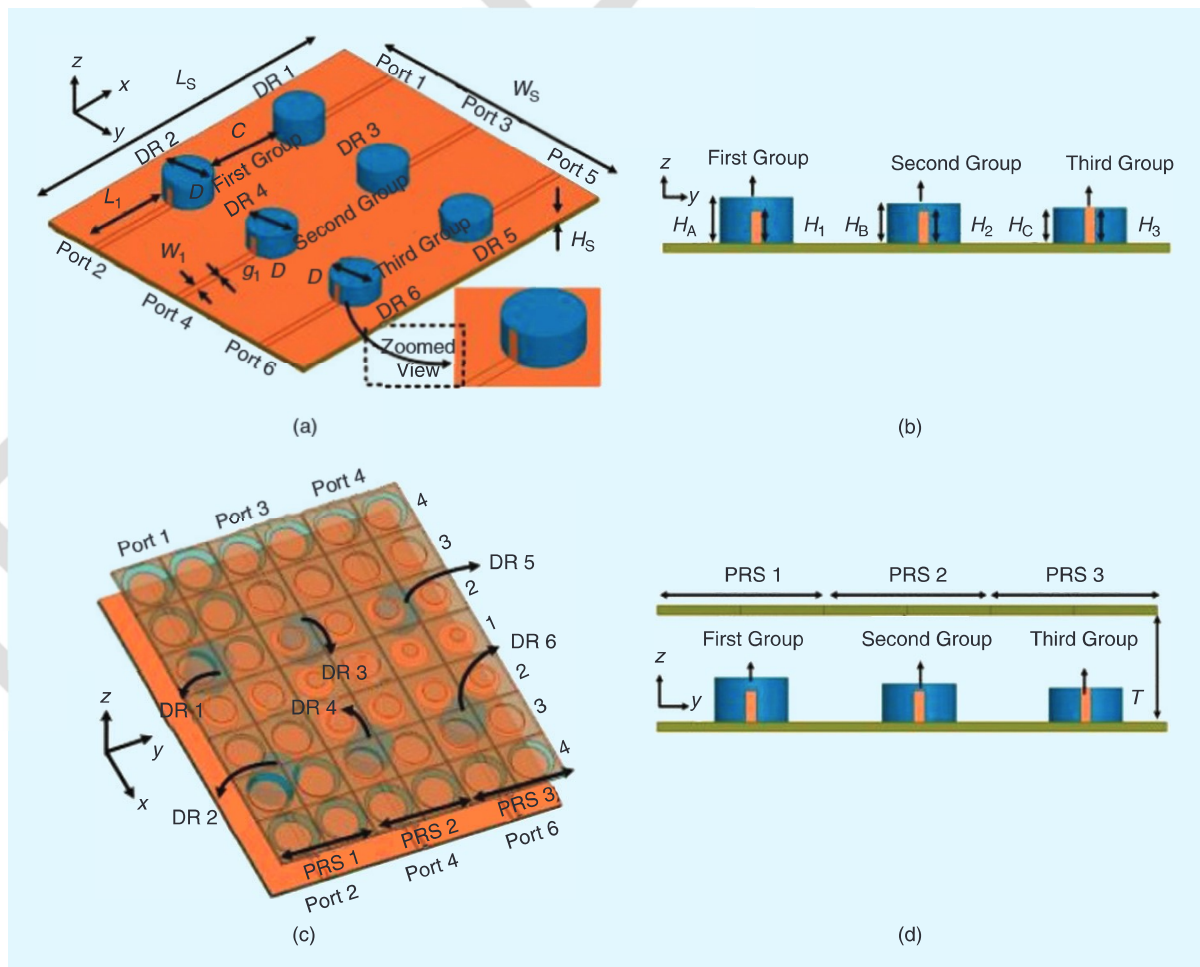
FIGURE 11. The geometry of a triple-band MIMO DRA with a partial reflector surface [84]. *<AU: Please provide a caption for parts (a)–(d). Kindly spell out PRS and define all variables in the images so the definitions can be added to the caption.>*

excited by four ports. However, only the simulated results were discussed, without fabrication of the design in [86] and [87].

A dual-port MIMO antenna was proposed in [88] with a technique to improve the isolation between two adjacent DRAs operating at 60 GHz. The improved isolation was achieved by using a metasurface shield constructed of a unique split-ring resonator that is designed to provide band-stop functionality over the operating frequency range centered at 60 GHz. Integrating the array of the SRR **<AU: Kindly spell out SRR.>** cell structure between the H-plane DRs along the E-plane results in a substantial reduction in the mutual coupling between the adjacent radiators. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

An eight-element MIMO antenna system was proposed in [89] for 4G and 5G mobile communication. It consists of two arrays with four elements that are closely spaced, which are selected to compensate for the free-space path losses that can be observed at mm-wave frequencies and provide beam-steering capability to reduce channel correlation.

A metamaterial polarization rotator wall was used in [90] to reduce mutual coupling due to spatial electromagnetic fields between two CDRAs in an MIMO system. In [91], a beam-tilting MIMO antenna loaded with the metamaterial unit cells was investigated. The authors proposed the metamaterial unit cells, which are used to create a negative refractive index medium, which is responsible for the beam-tilting mechanism. The proposed beam-tilting antenna consists of a double-feed DRA and 1×4 NRIM **<AU: Kindly spell out NRIM.>** array fixed by nylon studs over the DRA.

A three-port, two-CDRA-based MIMO antenna was discussed in [92]. Two CDRAs were excited by two probe-coupled feeds and one MS line feed with an array mode. The isolation between ports was improved by etching two C-shaped defected GND slits. Due to the coaxial probe and MS line-fed ports, the orthogonal modes $HE_{11\delta}^x$ and $HE_{11\delta}^y$, respectively, were generated on the CDRAs. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>**

Das et al. [93] proposed a compact back-to-back, two-element CDRA-based, four-port MIMO antenna. These CDRAs were glued on opposite sides of the FR4 substrate; the top two ports were fed using a CPW conformal MS line, and the bottom two ports were excited using MS line-fed conformal strip lines. Isolation between the ports was improved by generating orthogonal modes in CDRAs and the use of opposite excitation. In [94], a four-element, eight-port MIMO antenna with pattern diversity was proposed. Due to cross configuration of the four CDRAs and eight feeds, the GND could not be made common with all of the ports.

A simple decoupling method using metallic vias was investigated in [95] in a two-RDRA-based MIMO antenna. The geometry of the antenna is shown in Figure 12. It was found that, by placing the vias vertically inside the DRA elements at the appropriate positions, the mutual coupling can be substantially reduced, without increasing either the footprint or height of the antenna. At the same time, the decoupling vias slightly affect the field distributions of the excited antenna, and, therefore, the radiation patterns are not deformed and kept almost the same as the single DRA. In addition, since the DR is still the main radiator and the additional vias do not resonate in the operating band, the loss caused by the vias is desirably negligible.

A four-port MIMO antenna was proposed in [96], where antenna beams are separated spatially to achieve high isolation and low correlation. The authors used a two-phase gradient frequency structure simulator (FSS) and placed both sides of the substrate as a superstrate. Two CDRAs are positioned at the top of the substrate, and another two are placed at its bottom. With this special technique, the antenna achieves more than 20 dB of isolation and excellent envelope correlation coefficient values (less than 0.1).

A method to improve the isolation between two adjacent RDRA elements was proposed in [97] by introducing a metal strip printed on the upper surface of each RDRA. The strongest part of the coupling field moved away from the adjacent exciting slot of the DRA so that a maximum improvement of 12 dB on the isolation over 27.5–28.35 GHz was achieved.

A CDRA was split into four quad CDRAs in [98], and these four elements were excited using four MS line feeds. Exploiting image theory, a copper strip was etched on the sides of the quad CDRA elements so that each element was characterized as a full CDRA. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** In [99], an eight-element-, 16-port-based MIMO DRA was proposed. Here, four CDRAs were placed back to back on a substrate and excited using CPW and MS line feeds.

Two CDRAs fed by a printed ridge gap waveguide (PRGW) for the beam-switching antenna system based on the FSS were investigated in [100] at 28 GHz. In the MIMO antenna, each CDRA was fed by a rectangular slot etched on the top plate of
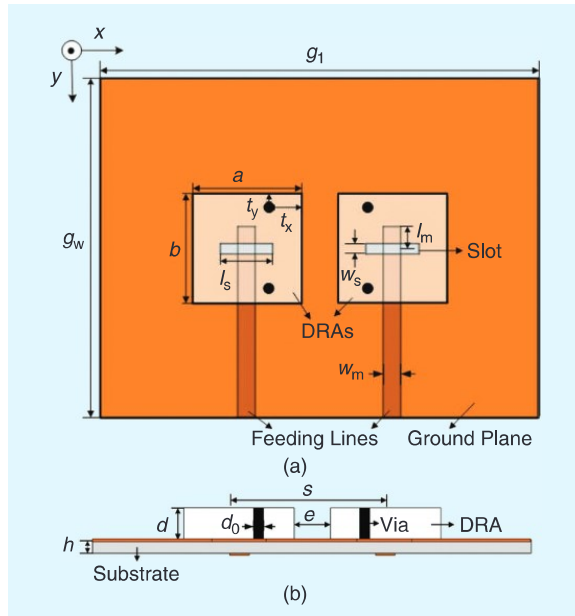
FIGURE 12. The geometry of an MIMO DRA with metallic via-based decoupling [95]. *<AU: Please provide a caption for parts (a) and (b).>*

Efforts To Overcome The Existing Impairments In Implementing Helicopter Satellite Communication System (HSCS)

Rashmi N[1*], Hamsavahini R[1], Abhinandini H R[2], Sanjana B A[2]

[1]*Department of Electronics and Communication Engineering, BMS Institute of Technology andManagement, Bangalore, Karnataka, 560064,India*

[2]*Department of Electronics and Communication Engineering, BMS Institute of Technology andManagement, Bangalore, Karnataka, 560064,India*

rashmiswamy@bmsit.in

Abstract

Communication can be broadly classified as analog and digital. These two types only deal with terrestrial networks. But, when we want communication to be achieved with very high efficiency it is very much essential to deal with satellites. Satellite Communication is becoming more advantageous day by day. Almost in all the areas, the importance of satellite communication is being recognized. The role of satellite communication in the disaster management has played an important role these days. This has gained very high popularity because of the failure or delay of the organization to detect the disaster at the right time. Satellite communication is generally useful for information gathering during the disaster [1]. Helicopters are used in the satellite communication technology and this technique is called as Helicopter Satellite Communication System (HSCS). There are few technical issues existing in this system too. Thus to overcome these issues and make it more efficient many schemes have been implemented. One such scheme is hybrid modulation with Automatic Frequency Control (AFC). [2]

1. Introduction

Communication is the most important part of our daily life. Not only owing to one's life, communication also plays a major role when it comes to avionics. Avionics is the science which deals with aircrafts and rotorcrafts. Helicopter is the most important rotorcraft which is used in Indian Army, Navy and Air force. Hence communication is very much important between the crew members of the helicopter as well as the pilots of other helicopters. Thus to facilitate such communication, Analog Communication System was introduced Analog Intercommunication System provides the direct mode of communication. The problems associated with the analog system are, deformation and the loss appearing in the signal and the proper control over the audio, that is, its volume and frequencies were not taken care of. Thus to overcome these problems, Digital Intercommunication System was introduced. Digital intercommunication system provides reliability and resistance against tightly packed voice signals. It also provides the communication of data from air to air, air to ground and ground to air using high frequency technologies such as Very High Frequency and Ultra High Frequency. The foremost aim of the Digital Intercommunication System is providing two way communication amidst the crew members (pilot and co-pilot) in the flight or helicopter (any aircraft or rotorcraft). In addition to this, the system has the provision to communicate the signals to the crew members at the ground station using a jack externally during the sustentation operations at any time necessary. This system is a chief control for all the communication equipment established on the Advanced Light

Helicopter. The Digital Intercom System connects itself with the various radio channels, warnings and other types of audio outputs. This system also provides the provision to communicate with the pilot and co-pilot of the same or different helicopters, and also with the Air Traffic Controller (ATC). [3] There were some of the problems associated with Digital Communication Systems too. Thus replacement of this old technology communication with Helicopter Satellite Communication is done in the recent years. Helicopter Satellite Communication plays a vital role in today's world because; helicopters are not just limited to one specific application. The Helicopters are used over a large area covering a wide range. They are used to acquire the information and broadcast the same in the cases of incidents such as disasters which take place over a large scale and also in the case of other emergencies. According to the present day scenario, the repeater stations are widely used to transfer the information such as video clippings and the photographs captured by the helicopter. Capturing of the videos using the helicopter is one of the best methods to obtain on scene information for relief and the management of disaster. [4.

There are few technical issues in this system. Many solutions are being provided to overcome these issues and declare the Helicopter Satellite Communication System (HSCS) as the best technique to provide communication using Helicopters.

2. Predicaments Associated With The Digital Communication System

The high requirement of bandwidth is the major problem in these systems. The proportion or the amount of information which can be transmitted per unit time depends on the bandwidth of the media usually used in the respective transmission. The analog systems have less requirement of bandwidth when compared to the digital systems. This is because, the digital signals, for example, the square pulse, delta function and many other digital signals, have many rising and falling pulses. These pulses have a lot of variation in their voltages within a very small amount of time. Thus we can say that, almost within a very few milliseconds, there will be a huge variation in its voltage. Because of this reason, these signals have very high frequency during their rise and fall. As the frequency increases, bandwidth also increases. Thus they occupy very large bandwidth. High power consumption is also a major issue in digital systems. The digital circuits consume a lot of power. This can be proved by the power dissipation in integrated circuits that is, in digital VLSI (Very Large Scale Integration). With regard to the integrated circuits, three major causes for the large power utilization are given by- utilization of the power dynamically, utilization of the power statically and utilization of the power in the short circuit. Improvement or rise in the power density and fall in the hardware measurements of the transistor are the stamp of authentication of the present day computer chips. Both of these technologies are increasing the thermal conductivity management challenge inside the chip and the package surrounding it. Both of these technologies are increasing the thermal conductivity management challenge inside the chip and the package surrounding it, as well as escalating research related progress on the materials which possess very high conductivity. The management of the thermal conductivity is largely acknowledged as an important characteristic of the computer design, due to which, the temperature significantly affects the performance of the device. Adding on to this, the lifetime of the device can be reduced extremely because of the excess thermal stresses which take place at the interfaces. [5]. Hardware complexity in the digital systems is another hurdle for the communication. The logic gates serve as an important factor in the designing of a digital circuit. These circuits put together manifest a digital system. The major problem while designing any kind of digital circuit is the hardware complexity of that circuit. When the concept of conventional gate logic is adopted in the design of the digital circuit, then the utilization of more number of components is the major challenge which it has to face. Thus these systems have more complex circuitry and require a sophisticated device making. [6]. The digital systems can be easily broken or damaged, that is, even if a small amount of digital data is unable to be recovered or if it is not interpreted exactly, the meaning of the large amount of data can change completely.

3. Role of Satellite Communication For Disaster Management

Whenever the earthbound communication resources are inadequate it means that there is a huge damage caused to the infrastructure or it can also be due to the enormous traffic. Hence we cannot depend entirely on the earthbound or the terrestrial communications for the disaster management operations. And it is also essential that for the disaster management operations the communication should be very flexible and reliable as it is the part of emergency communications. Thus everyone who desires to have such type of communications is able to experience perfect working of the communication system every time and everywhere under any circumstances. Furthermore, the broadband communications have been playing an important role in increasing the communications during the emergencies. By considering the radio links as a basis, the systems which are based on the satellites have a lot of benefits which include reliability, availability, robustness, quickness, deploy ability and accessibility from any location. The Satellite Communication System is a core network which is used for disaster management and its recovery which includes the topology of the satellite network, configuration of service and also for the technical applications. The communication provides the favorable path for managing the critical situations such as occurrence of the disaster and also helps to provide relief during these emergency situations. [7]

4. The Satellite Communication System Employing Helicopters

5. How Exactly The System Looks Like?

The continuous prevention of the radio waves being reached to the destination in the Helicopter and Satellite Communication is mainly due to the blades of the Helicopter. These helicopter blades intervene or interrupt the helicopter directly by communicating with the satellite in order to check whether the antenna is situated right under the blade of the helicopter. The time break is about 6.93 milliseconds and the amount of interception taking place is for 39.1 milliseconds. By picturing the case of the helicopter transmission, the system which usually transmits at the exact time of the blade has to be designed and this is considered to be an alternative to solve this complicated issue. [4]

6. The Designing Requirements For Helicopter Satellite Communication System

While designing the system based on Satellite Communication and Helicopter, the most important conditions to be followed were; 1. The system which has to be designed must be very small in size; 2. The system to be designed should not be bulky and heavy, instead light weight is preferred; 3. A very effective link must be established between a helicopter and the satellite for the communication to take place, and it should be irrespective of the helicopter direction; 4. The power due to transmission should not harm the pilot, co-pilot and other crew members; 5. The transmitting power should be maintained such that it should not prevent the radio waves reaching the destination due to the blockages in the rotor blade; 6. The video capturing objectives must be based on the positioning function. [4]

7. Technical Issues Ceasing To Implement The System

- Avoidance of the shadowing due to rotor blades;
- Satellite tracking;
- Polarization tracking;
- Avoidance of interference with other satellites [1]

8. Helicopter Videoing

The technique of capturing the video using helicopter is the most popular method to obtain the information for emergency communications such as disaster management. A new system has been developed which captures the information related to the disaster and immediately accounts for the transmission of this information. Such a system is called as Ku band utilization in the satellite communication with helicopter. As we have already discussed about the major issue in such type of communications, that is, interruption due to rotor blades of the helicopter, there are techniques proposed to overcome these drawbacks. These techniques are called as time diversity transmission technique which is applicable in the forward link, and the other technique namely blade synchronized transmission technique which can be applied at the return link. The estimation of the position is also implemented for on scene videoing pictures. [8]

9. Techniques To Improve The Efficiency of Hscs

In this system, there exists two major issues and they have to addressed properly; 1. Rotor blades of the helicopter is the major part which is leading to the blockage of the received signal; 2. The Doppler Shift in the system which accounts for the fluctuation of the carrier frequency. New techniques have been proposed to overcome these two drawbacks. There were many other techniques which were implemented to overcome these issues; but they were inefficient. The continuous blockage of the signal received can be cancelled by using a hybrid modulation scheme with Automatic Frequency Control (AFC). This technique is better when compared to the use of former diversity technique based on time, as this increases the amount of data transmitted per unit time; whereas AFC transmits the original amount of data per unit time. [2]. Most of the communications in the helicopter are handled by a radio which operates in very high frequency range and it has the limits such as its coverage area and the range at which it can operate. As the amount of area occupied by the relay station which is established on the ground is only around 40km (radius), there is a huge requirement for many numbers of such stations to be established. If we consider an idea of establishing relay stations, it doesn't work out in the areas covered completely with mountains or sea. Thus to overcome all these geographical limitations, the satellite itself is made as a relay station in this system. Whenever the helicopter is outside the terrestrial frequency range, the pilot operations must be undertaken with a very high efficiency because even it includes the satellite communication with the Air Traffic Controller. [2]   This system which is based on Satellite and Helicopter comprises of the two important stations; one at the ground and one at the helicopter. Communication between the two stations is considered to be bidirectionaland it can be achieved using a transponder. The transmission link is established between these two stations, that is forward and reverse link. The location of the Antenna should be such that it should not interfere with the rotor blades of the helicopter. It would be considered the best way if it is situated at the tail of the helicopter. Some of the experiments have proven that, as the antenna is located at a large distance from the main rotor blades, then the amount of interruption of signal with the blades would be avoided. [2]

   Establishing a system which communicates using the satellite technology with a helicopter is a tedious task, because the rotor blades of the helicopter is the major obstruction in such systems, as this leads to the Doppler shift. Thus because of these major impairments, many applications such as Global Positioning System, degradation in the performance of bit error rate and the discontinuity in the received signal over the recovered carrier. [2]

10. Results And Discussions

The transmission from the helicopter can automatically stop when 1. The tracking error angle becomes large 2. The receiving level becomes below the limited value 3. The posture of the helicopter inclines beyond the anticipation 4. The onboard equipment fails. The final flight test included the following ventures.

- Verification of the satellite tracking performance
- Measurement of the communication quality

- Operation check of transmission protection
- Picture transmission and position fixing. [1]

The parameters of simulation in order to calculate BER is given below:

Figure1. Diagram showcasing how exactly the communication takes place by adopting the Principles of satellite communication in the Helicopter

Figure 2 Phase symbol before and after AFC

Figure 3 BER behavior for different schemes

Table   **I**   **.**

Helicopter	
Rotor revolutions	$\omega = 6.45$ rev/s (387 rev/min)
Blades	$M = 4$
Maximum speed	$V = 68.4\ m/s$ (246km/hour)
Maximum acceleration	$A = 3.09\ m/s2$ (11.1km/hour/s)
Modem	
Modulation	BPSK with time diversity
Time diversity factor	$L = 2$
Information data rate	$R_b = 3600\ bit/s$
Carrier frequency	$f_C = 12.5\ GHz$
Demodulation	Coherent detection
Forward error correction	
Code	Convolutional code with
Viterbi decoding	
Rate	$R = 1/2$
Constraint length	$K = 7$

11. CONCLUSION

The difficulties in analog as well as digital communication systems led to the development of new technology in communication era which is called as Helicopter Satellite Communication System (HSCS). Thus the helicopter satellite system was developed successfully and it was also demonstrated for the first time. The key technologies of the HSCS for disaster control operations were experimentally demonstrated. The developed HSCS system provides a real time information collection and transmission directly via a satellite. Research and Development aiming at small-size, light-weight and wide-band HSCS is proceeding. Continuous research and development collaborated with disaster management bodies will be carried out. [1] This paper concludes that the Helicopter Satellite Communication system is one of the best ways to realize communication. A novel modulation scheme based on hybrid techniques for the accurate cancellation of the periodic or continuous blockage in Helicopter Satellite Communication System is implemented instead of conventional time diversity scheme. The proposed AFC scheme incorporates adaptive modulation in an efficient way. The traditional time diversity scheme increases the transmission bit rate to reduce errors, but in the novel AFC scheme, similar type of signal is made to transmit just by changing the number of time slots. [2]

REFERENCES

[1] WataruChujo, "Helicopter Satellite Communication System for Disaster Control Operations", Kashima Space Research Centre, National Institute of Information and Communications Technology, 893-1, Hirai, Kashima, Ibaraki 314-8501, Japan.

[2] Emad. A. Elaty, Sami A El-Doli, Mohammed Abd-Elnaby, Fathi E Abd El-Samie, "An Adaptive Modulation Scheme in Helicoper-Satellite Communication", The Mediterranean Journal of Electronics and Communications, Vol.9, No.3, September 2013.

[3] DankanGowda, MahaveerPenna, SatyasrikanthPalle, SaiPriya K A and Chaithra K N, "Digital Intercommunication System in Advanced Light Helicopter", IEEE, Advances in Computational Sciences and Technology, ISSN 0973-6107 Volume 10, Number 5, pp. 699-708, Research India Publications, 2017

[4] Masaki Satoh et al, Huan-Bang Li, Yoshiyuki Fujino, HiromtsuWakana, Amane Miura, Yutaka Ozaki, Hiroyuki Satou, Eiji Watanabe, Manabu Sawa, "Helicopter-Satellite Communication System Developed for Transmission of Disaster and Emergency Information", Communication Research Laboratory, American Institute of Aeronautics and Astronautics, Inc., AIAA-2003-2319, 2003.

[5] Patrick K. Schelling, Li Shi and Kenneth E. Goodson, "Managing heat for Electronics", Materialstoday, June 2005.

[6] Muhammad Khalid, ShubhankarMajumdar and Mohammad JawaidSiddiqui, "Reduction of Hardware Complexity of Digital Circuits by Threshold Logic Gates Using RTDs", Information and Communication Technology for Intelligent Systems, Smart Innovation, Systems and Technologies 106.

[7] Yang-Min Lee, Bon-lun Ku and Do-SeabAhn, "A Satellite Core Network System for Emergency Management and Disaster Recovery", ICT Standardization program of MKE (The Ministry of Knowledge Economy) [20l0-Pl-02, Standardization on PPDR communication for next generation], IEEE, 2010.

[8] Huan-Bang Li, Masaki Sato, Amane Miura, Shinichi Taira, and HiromitsuWakana, "Ku-band Helicopter Satellite Communications for on scene disaster information transmission", National Lnstitute of Information and Communications Technology 893- 1 Hirai, Kashima-shi, Ibarah 3 14-8501 Japan, 0-7803-8523-3/04/$20.0, IEEE, 2004.

Noise Identification and Removal Technique for Precise Lung Cancer Cell Detection using Active Contour Model

Shilpa Hiremath[1], Dr. A Shobha Rani[2]

BMS Institute of Technology and Management, Bangalore

Abstract

Image processing is a domain where a lot of research is required, as each other field is associated with image processing directly or indirectly. In our work we considered the medical field, where segmentation of an image is required to extract the information from the given image and illustrate anatomical structures based on the feature of the image known prior to us or based on the expert knowledge. The input image that we consider is the noisy medical image which is applied to various filtering techniques to remove the noise present in an image, after recognizing the nature of noise present in an image. Edge detection technique is applied for the filtered image to obtain the better results. Finally edge segmentation is done to the output of the edge detected output, to identify the region of tumor.

Keywords: Active Contour, Edge detection, Frost Filter, Image noise, Image Segmentation, Region Growing.

1. INTRODUCTION

Medical imaging application plays an important role in the anatomical structures. Understanding the data from the medical image is a very tedious task due to the numerous features that an image has, which leads to the complication of segmentation hence image segmentation of a medical image is an inspiring work that a researcher can take up. Information content of an image can be better analyzed and read by the user by performing the segmentation of an image. Currently lot of work is going on in the medical image processing area namely in the computer tomography image, X-ray images, Magnetic Resonance Images and Position Emission Tomography images etc. [1].

By performing the image segmentation technique on a medical image, we can obtain the regions of interest for visualization which is very necessary for the doctor to take any decision about the patient like for adjuvant therapy, planning for a surgery, model for teaching, prosthetic design, and various others [22] which can be seen in figure 1.

Figure 1: Need for Medical Image Segmentation

Segmentation of an image usually divides the entire image into its constituent components [35]. Image segmentation helps in extracting the selected object in each image hence Segmentation is usually followed by image representation and description and identifying the object in an image.

Before the part of segmentation methods, we apply image to filtering. In-order to retain all the fine details and edges of an image and to remove the noise we apply filtering technique. These techniques differ in their straightforward procedures applied to remove the noise or suppress the noise. Few techniques identify the noise and then apply the filtering technique and few others directly apply the filtering technique without considering the noise parameter.

In our paper we are considering the noise factor, firstly we are recognizing the nature of noise, which is present in an image, the image we have considered here is a medical image and then a suitable smoothening technique is applied which is used to remove the noise from an image and then segmentation of this image is done which helps in medical image diagnosis.

2. LITERATURE SURVEY

Lot of work have been carried out in the field of medical image analysis [35]. Here we will provide brief survey on the work being carried out in this field by various researchers by considering thresholding and region growing value.

(a) Thresholding:

AbuBaker, A. A., Qahwaji, R. S., Aqel, M. J., & Saleh, M. H. [23] proposed two image segmentation technique upon considering threshold values by scanning each row in an image and making a small modification to this method to find the threshold value by averaging.

Ng, H. P., Huang, S., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. [24] considered the threshold value by studying the gradient magnitude image also by considering the histogram of the gradient magnitude map the automated threshold value is obtained.

Kotropoulos, C., & Pitas, I. [25] have well thought out the use of support vector machine to segment the image and find the region of interest (ROI).

Khare, A., & Tiwary, U. S. [26] proposed a method to denoise the medical image by considering the threshold value and proposed a method, where the threshold value can adapt to change in the value based on the mean, median and standard deviation of wavelet coefficients.

(b) Region Growing:

Haider, W., Sharif, M., & Raza, M. [28] worked with breast and lung cancer where they used edge detection and segmentation as morphology. The main aim of this work is to detect the tumor region.

Siddique, I., Bajwa, I. S., Naveed, M. S., & Choudhary, M. A. [29] researchers worked on the brain images; they have considered the concept of region growing technique based on initial seed point concern.

Oghli, M. G., Fallahi, A., & Pooyan, M. [30] have considered the same factor as in [29] but proposed an automatic approach for region growing. The seed point is selected based on co-occurrence matrix.

Deng, W., Xiao, W., Deng, H., & Liu, J. [31] have studied the variations happening near the edges and the gradient value. The method turns out to be very efficient near the edges as it helps in identifying the tumor region.

Poonguzhali, S., & Ravindran, G. [32] have thought of the ultrasound images and developed a segmentation technique which is more efficient for speckle noise.

The researchers in [33] Guan, H., Li, D. Y., Lin, J. L., & Wang, T. F. have developed a method comprising of two methods namely region growing and region merging for ultrasound images.

Pan, Z., & Lu, J. [34] used Baye's concept for segmenting a medical image, the researchers developed a multistage approach which adjusted the parameter for region growing and improves the efficiency.

3. METHODOLOGY

Considering the original noisy medical image, the very first step is to apply the filtering technique depending upon the type of noise present in an image. In our work we have considered three types of noise i.e., impulsive noise, Gaussian noise, and multiplicative noise. If the noise present is impulsive then we use median filter to remove the impulsive noise, if the noise present is Gaussian then use Gaussian filter else, we apply frost filter to remove the multiplicative noise. The next step is to detect the edges of an image by performing the edge detection technique and for this edge detected output we

perform image segmentation using snake method. Figure 2 shows the flow chart of the steps performed to identify the lung cancer region in an image.

Figure 2: Flow chart of computed tomography imaging segmentation and analysis of lung cancer

4. NOISE IDENTIFICATION

Any unwanted information which is present in the image can be considered as image noise [27]. Image noise makes its presence in an image because of various reasons like atmospheric effects while capturing the image, it can be because of the sensor which is used to capture the image or even during the transmission of an image, noise can be added. Irrespective of the way the noise is being added, it is affecting the image and changing the pixel value or say intensity value. Hence identifying those pixel values which are changed because of this noise is very crucial in image processing.

There are two methods which helps in analyzing the histogram of the image in a step-by-step way namely: The Peak detection method and valley extraction method. Once the analysis of histogram of the medical image is done then this is used in extracting the valleys which are used as threshold value for segmentation process. A distinctive histogram analysis is usually done in three stages:

Stage 1: Identify the leading modes of the histogram.
Stage 2: The valleys among different modes is recognized.
Stage 3: The identified threshold is applied to the medical image for segmentation.

For identification of the noise present in an image first we need to study the behavior of the noise i.e., the variation of the gray level values of the local homogeneous regions which are affecting the image. Depending upon the image size the pixel values considered for the local homogeneous region varies, approximately it must have at least 128 pixels.

Here in our work, we have considered two step which are used to find the noise type which is affecting the medical image.

Step 1: If impulsive noise is present in an image being considered then apply filtering technique to remove it, else if.

Step 2: Additive noise is present apply filtering technique to remove additive noise or else apply filtering technique to remove the multiplicative noise.

(a) Detection of Impulsive Noise:

An input image is said to be affected by impulsive noise if the ratio of the dynamic mean of a homogeneous region to that of the dynamic max of a homogeneous region is greater than the threshold value.

(b)**Detection of Additive or Multiplicative Noise:**

An input image is said to be affected by additive noise if the gray level value of the homogeneous region is closer to the value of zero. Else it can be considered as a multiplicative noise is present in an image.

4. IDEOLOGY OF MEDICAL IMAGE SMOOTHING

The process of removing the noise from an image to retain the original image by applying a suitable filtering technique is said to be as image smoothing. The noise can appear in an image because of various reasons, and it is because of numerous non-idealities which occur in an image process. We usually say noise is present if the intensity value or say the pixel value of an image is changed usually having the values near or equal to the minimum or maximum of the allowable dynamic luminance range.

Almost all applications which deal with images has to be noise free in order for getting the correct results and having a good performance. Hence, this acts as a primary step for any image processing task to be carry forwarded. However, this is a tedious task since removal of noisy pixel should not alter the original pixel of an image which may represent the edge of an image or texture of an image etc. [3].

The image under consideration is denoised in two stage, in first stage if the noise present in a medical image is identified as impulsive noise then a suitable filter is applied usually it is a median filter which is very good for removal of impulsive noise. And, in second stage if the noise present in a medical image is additive then a filter suitable for removal of additive noise is applied i.e Gaussian filter or if the noise present is multiplicative, the image is then applied to Frost filter to remove it [5].

(a) **Impulsive Noise Smoothing:**

For better image restoration we use median filtering which is basically a non-linear filter. This median filter replaces the center pixel by the new pixel value by considering rank order information and then considering the median value of the pixel within the window. This technique of filtering using median value uses a two-dimensional window which is applied to every pixel of the input image.

$$y_{median} = \begin{cases} median\ \{x_k\}; & if\ |median\ \{x_k\} - x| > threshold \\ x & ;\quad otherwise \end{cases}$$

(b) **Additive Noise Smoothing:**

Gaussian filtering technique is used to remove the additive noise present in an image. This linear filtering technique removes the Gaussian noise present in an image by performing convolution method over the window. Digital images are usually represented in the form of a discrete value; hence the Gaussian window will also have a discrete approximation of the two-dimensional Gaussian function. The Gaussian expression can be represented as below:

$$g(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma}} \cdot e^{-\frac{x^2}{2\sigma^2}}$$

(c) **Multiplicative Noise Smoothing:**

Frost filtering technique is used to remove the multiplicative noise present in an image. The window is moved over the entire image to identify and remove the multiplicative noise. This filtering technique considers the window and substitutes the center pixel with a weighted sum of the pixel values. As the distance of the center pixel value increases the value of the weighting factors decrease [5].

5. IDEOLOGY OF MEDICAL IMAGE EDGE DETECTION

Edge detection is defined as the process of detecting the edges in an image. Edges are usually having a sharp transition when compared to the other regions of an image which means that the pixel value of an edge varies very significantly from its neighboring pixel. The Edge detection technique helps in identifying the boundaries, sharp transitions, corners of an image and remove the smooth transitions which assists in understanding the image. Further processing of an image like image enhancement, registration, segmentation, or identification of an object in an image can be performed based on the results obtained from edge detection. Multidimensional filtering method can be used to perform edge detection.

M-Dimensional filtering method:

A real image which has intrinsic geometrical constructions in particular the object boundaries or edges or corners of an object in an image are more highlighted for the visual information. Hence, performing a filtering technique to visualize these vital information of edges plays a very important step.

A (discrete domain) filter can be characterized in two forms either by its impulse response which is denoted by h[n] or by transfer function denoted by H(z) = Z{h[n]}. This filtering technique performs the convolution operation on the input image and the impulse response and produces the resultant output. In the resultant output, each sample has the inner product with a translated copy of the flipped input signal and complex conjugated filter impulse response. That is, the output signal of the filtering operation is a linear combination of the impulse response h and its translated copies [7].

6. IDEOLOGY OF MEDICAL IMAGE SEGMENTATION

Image segmentation is defined as partitioning of an image into various sections. This is domain dependent where it considers the disjoint regions present in an image which are visually different, homogeneous, and meaningful in their pixel level or texture or some other characteristics and considers it to be as one group. The main objective is to streamline the depiction of an image into a more meaningful image and easy to analyze it.

(a) Region growing method:

In region growing method the very first step is to select some start point popularly known as seed point. Seed point can be selected based on application or based on user criteria like considering some grey scale region and one pixel in that range as a seed point, color criteria or grey level texture etc. The region can grow starting from these seed points and further can extend to the adjacent points based on the criteria being selected. This process is continued till all the allocated pixels are compared using the difference between the intensity value of the pixel and the regions mean value. This factor is used to identify the similarity in the regions. If the difference value is small, then the threshold value under consideration then it is considered as the belonging to the same region else to the other region. This iterative process stops if the resultant value of the difference is larger than the threshold value being considered.

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5

0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1
3	2	1	0	1	1	0	2

Before region growing

Threshold < 3, image size = 8X8,

red indicates seed point

0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1
3	2	1	0	1	1	0	2

After region growing

Threshold < 3, image size = 8X8,

 red indicates seed point (i.e pixel value = 6);

white region indicates one segment and gray indicates the other segment

Figure 3: Region Growing Method

(c) **Split & merge image segmentation:**

In this technique, a threshold value is considered and based on this threshold we can decide which all regions in an image can be merged and considered as one segment or object. The entire image is split into blocks and then difference between the maximum and minimum pixel value in each block is calculated. This difference value is checked with the threshold value and if this value is below the threshold value then it is considered as the one region else further the block is spilt into smaller block and the process is repeated until no blocks can further be split. Threshold value plays a very important role in this technique hence choosing the right value of threshold is very significant [8]. In the below figure the image is of size 8X8, considering the entire image as one segment the maximum value is 7 and minimum value is 0 in this segment. The threshold value chosen is less than or equal to 3. Since the difference value is equal to 7 which is greater than threshold value hence the image is divided into four segments indicated by white region, green, yellow and gray region. Now, considering the grey region the max value is 7 and min value is 0 which gives the result of 7 which is more than the threshold value this leads to further splitting of the grey region which are indicated by different colored numbers and the process is repeated. Figure b shows the resultant image pixel values obtained by performing the split and merge segmentation.

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5
0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5
0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1

3	2	1	0	1	1	0	2		3	2	1	0	1	1	0	2	
Threshold<=3																	

Figure 4: Split and Merge Image Segmentation

(C) Active contour (Snake) method:

Snake method is usually implemented after the edge detection technique, where all the local regions of the object are identified in the edge detection method later which snake method is implemented which reduces the energy function. This method is used in various applications where identification of an object in an image is required, there is a need to identify the shape of the object, or there is a need for stereo matching etc. This method is said to be as active method as this reduces the energy required and has dynamic behavior. This method leads to an arbitrary shape which is close to the object contour as much as possible. Basically, we can consider this energy functions as two components: potential component and internal deformation energy.

a) If the contour which is aligned to the image edge is small, then its potential energy.

b) If the contour which is smooth is small, then its internal deformation energy.

The termination functional can be implemented with a gradient direction calculus in a slightly smoothed version of the image.

An Active contour can be represented as v(s) = (x(s), y(s)) and its energy functional is represented as follows:

$$E = \int_0^1 E_{int}(v(s))ds + \int_0^1 E_{image}(v(s))ds + \int_0^1 E_{ext}(v(s))ds$$

Where E_{int} represents the internal energy of the spline due to bending, E_{image} represents the image forces and E_{ext} represents the external constraint forces, *a(s) and B(s) are* the controlling spline energy values. Hence, the internal spline energy expression can be given as:

$$E_{int} = \frac{(a(s)|v_s(s)|^2 + B(s)|v_{ss}(s)|^2)}{2}$$

The total energy of an image is given by the sum of weighted values of the functions of all the three energies.

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term}$$

The advantageous factor of this model is that without extra processing, the object shape in an image is being identified in the piece-wise linear manner at the time of convergence.

7. RESULTS AND DISCUSSION

The objects in an image which is of our interest and the boundaries of an image namely lines and curves can be identified by the image processing technique called image segmentation. The proposed algorithm is implemented using MATLAB platform and the test results after performing various operations are as shown in figure 5.

a	b
c	d

e	f
g	h
i	j

Figure 5: (a) Original image, (b) Image with impulse noise, (c) Image with multiplicative noise, (d) Image with additive noise, (e) Filtered additive noise, (f) Filtered multiplicative noise, (g) Filtered output, (h) M-Dimensional filtered output, (I) Region growing segmentation, (j) Active counter (Snake) segmentation.

8. CONCLUSION

In this paper we have successfully identified the type of the noise and filtered it by using suitable noise removal filters. Multidimensional filter provides a powerful computational tool including natural images, medical diagnostic images. The specialty of merge-split algorithm which is to find the variation among the maximum and minimum values of a pixel within the window is like the edge detection technique. In images such as an image which has a smooth texture or an image which has less gradient, the edges play a very important role since this is the only region in the image where there is a large difference in the pixel values as compared to other regions in an image. Hence, merge and split technique is very helpful in such kind of images. This technique splits the blocks into individual pixels in which larger merged blocks appear in the interiors. This technique plays an important role and is carried out in the beginning of the image segmentation which leads to quicker region growing. From the initial contour, the object boundaries can be automatically located with the help of active contour. The most advantageous benefit of active contours is its ability to give a piece-wise linear description of the object shape at the time of convergence, without extra processing.

REFERENCES

[1] Pujar, J. H., Gurjal, P. S., Shambhavi, D. S., & Kunnur, K. S. (2010). Medical image segmentation based on vigorous smoothing and edge detection ideology. International Journal of Electrical and Computer Engineering, 5(2), 121-127.

[2] Bindu, C. H. (2009). An improved medical image segmentation algorithm using Otsu method. International Journal of Recent Trends in Engineering, 2(3), 88.

[3] Lagendijk, R. L., & Biemond, J. (2009). Basic methods for image restoration and identification. In The essential guide to image processing (pp. 323-348). Academic Press.

[4] Pujar, J. H., & Kunnur, K. S. (2010). A NOVEL APPROACH FOR IMAGE RESTORATION VIA NEAREST NEIGHBOUR METHOD. Journal of Theoretical & Applied Information Technology, 14.

[5] Frost, V. S., Stiles, J. A., Shanmugan, K. S., & Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Transactions on pattern analysis and machine intelligence, (2), 157-166.

[6] Liu, K. (2006). An Implementation of the Median Filter and Its Effectiveness on Different Kinds of Images. Thomas Jefferson High School for Science and Technology Computer Systems Lab, 2007.

[7] Do, M. N., & Lu, Y. M. (2011). Multidimensional filter banks and multiscale geometric representations. Foundation and Trends in Signal Processing, 5(3), 157-264.

[8] Kamdi, S., & Krishna, R. K. (2012). Image segmentation and region growing algorithm. International Journal of Computer Technology and Electronics Engineering (IJCTEE), 2(1), 103-107.

[9] Aly, A. A., Deris, S. B., & Zaki, N. (2011). A novel image segmentation enhancement technique based on active contour and topological alignments. arXiv preprint arXiv:1106.0371.

[10] Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. International journal of computer vision, 1(4), 321-331.

[11] Pujar, J. H., & Shambhavi, D. S. (2010, March). A novel digital algorithm for Sobel edge detection. In International Conference on Business Administration and Information Processing (pp. 91-95). Springer, Berlin, Heidelberg.

[12] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. IEEE Transactions on pattern analysis and machine intelligence, 22(8), 888-905.

[13] Liu, K. (2006). An Implementation of the Median Filter and Its Effectiveness on Different Kinds of Images. Thomas Jefferson High School for Science and Technology Computer Systems Lab, 2007.

[14] Nafi'iyah, N., Fatichah, C., Astuti, E. R., & Herumurti, D. (2020, December). The Use of Pre and Post Processing to Enhance Mandible Segmentation using Active Contours on Dental Panoramic Radiography Images. In 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (pp. 661-666). IEEE.

[15] Hsu, W. Y., Lu, C. C., & Hsu, Y. Y. (2020). Improving segmentation accuracy of CT kidney cancer images using adaptive active contour model. Medicine, 99(47).

[16] Wu, X., Tan, G., Li, K., Li, S., Wen, H., Zhu, X., & Cai, W. (2020). Deep Parametric Active Contour Model for Neurofibromatosis Segmentation. Future Generation Computer Systems, 112, 58-66.

[17] Inayat, A., Jawaid, M., & Talpur, S. (2020). Optimizing Active Contour Evolution for Robust Object Segmentation in Medical Imagery.

[18] Malligeswari, N., & Kavya, G. (2021). A Versatile Segmentation Approach For Diagnosis of Lung Cancer.

[19] Voronin, V., Zhdanova, M., Semenishchev, E., Zelensky, A., & Agaian, S. (2020). An active contour model for medical image segmentation using a quaternion framework. Electronic Imaging, 2020(10), 62-1.

[20] Malathi, M., Kesavan, S., & Praveen, K. (2021). *Segmentation of Spine Tumour Using K-Means and Active Contour and Feature Extraction Using GLCM.* In AI Innovation in Medical Imaging Diagnostics (pp. 194-207). IGI Global.

[21] Ma, J., Wang, D., Wang, X. P., & Yang, X. (2020). *A fast algorithm for geodesic active contours with applications to medical image segmentation.* arXiv preprint arXiv:2007.00525.

[22] Masood, S., Sharif, M., Masood, A., Yasmin, M., & Raza, M. (2015). *A survey on medical image segmentation.* Current Medical Imaging, 11(1), 3-14.

[23] AbuBaker, A. A., Qahwaji, R. S., Aqel, M. J., & Saleh, M. H. (2006, January). *Average row thresholding method for mammogram segmentation.* In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (pp. 3288-3291). IEEE.

[24] Ng, H. P., Huang, S., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2008, August). *Medical image segmentation using watershed segmentation with texture-based region merging.* In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 4039-4042). IEEE.

[25] Kotropoulos, C., & Pitas, I. (2003). *Segmentation of ultrasonic images using support vector machines.* Pattern Recognition Letters, 24(4-5), 715-727.

[26] Khare, A., & Tiwary, U. S. (2005). *Soft-thresholding for denoising of medical images—a multiresolution approach.* International Journal of Wavelets, Multiresolution and Information Processing, 3(04), 477-496.

[27] Hiremath, S., & Rani, A. S. (2008). *A Concise Report on Image Types, Image File Format and Noise Model for Image Preprocessing.*

[28] Haider, W., Sharif, M., & Raza, M. (2011). *Achieving accuracy in early stage tumor identification systems based on image segmentation and 3D structure analysis.* Computer Engineering and Intelligent Systems, 2(6), 96-102.

[29] Siddique, I., Bajwa, I. S., Naveed, M. S., & Choudhary, M. A. (2006, December). *Automatic functional brain MR image segmentation using region growing and seed pixel.* In International Conference on Information & Communications Technology, December (Vol. 6).

[30] Oghli, M. G., Fallahi, A., & Pooyan, M. (2010, May). *Automatic region growing method using GSmap and spatial information on ultrasound images.* In 2010 18th Iranian Conference on Electrical Engineering (pp. 35-38). IEEE.

[31] Deng, W., Xiao, W., Deng, H., & Liu, J. (2010, October). *MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve.* In 2010 3rd International Conference on Biomedical Engineering and Informatics (Vol. 1, pp. 393-396). IEEE.

[32] Poonguzhali, S., & Ravindran, G. (2006, December). *A complete automatic region growing method for segmentation of masses on ultrasound images.* In 2006 International Conference on Biomedical and Pharmaceutical Engineering (pp. 88-92). IEEE.

[33] Guan, H., Li, D. Y., Lin, J. L., & Wang, T. F. (2007, May). Segmentation of ultrasound medical image using a hybrid method. In 2007 IEEE/ICME International Conference on Complex Medical Engineering (pp. 644-647). IEEE.

[34] Pan, Z., & Lu, J. (2007). A Bayes-based region-growing algorithm for medical image segmentation. Computing in science & Engineering, 9(4), 32-38.

[35] 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 (pp. 1-22). IGI Global.

Digital Design of Optimized DCTQ Processor for Image Compression

Suryakanth B | B M Srinvasulu | Hussain Ghantiwala | Aviral Singh | Mallikarjun J Pathanshetty

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BMS INSTITUTE OF TECHNOLOGY

EMAIL:1BY17EC034@bmsit.in

ABSTRACT: In this era of Internet of Things, wherein every 'thing' is integrated within the existing internet architecture, it becomes quite necessary that embedded computing systems process quickly, occupy less area and consume low power. This would enable them to work quickly with real time data and have a large shelf life. As such there is a need for development of optimized algorithms and their efficient implementation in hardware. This paper presents a novel architecture for obtaining, DCTQ coefficients suitable for FPGA Implementation. The design is highly parallel and pipelined so as to exploit the massive parallelism of FPGAs and occupies considerably less area with a very high processing speed.

INTRODUCTION

Digital images have been widely used today in various applications and it is growing rapidly. Video and television transmission is becoming digital. More and more digital image sequences are used in various multimedia applications. Image compression addresses the problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage transmission requirements. Image compression basically stems from two fundamental concepts. They are

1. Redundancies reduction: It aims at removing duplication from the signal source (image/video).
2. Irrelevancy reduction: It erases that part of the signal which will be not noticed by the signal receiver, namely the Human Visual System.

The redundancy in data representation is reduced such a way that it can be subsequently reinserted to recover the original data, which is called decompression of the data. Data compression can be understood as a method that takes an input data, D and generates a shorter representation of the data (D) with a smaller number of bits compared to that of D. The reverse process is

called decompression, which takes the compressed data c (D) and generates or reconstructs the data, D. Most often the compression (coding) and decompression (decoding) systems together are called a "CODEC". The

objective of image compression is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. Image compression can be lossy or lossless. Lossless compression is sometimes preferred for artificial images such as technical drawings, icons or comics. This is because lossy compression methods, especially when used at low bit rates, introduce compression artifacts. With the development of imaging multimedia applications (image archiving, network image transmission, document imaging, digital photography, medical imaging, remote sensing etc.) image compression requires higher performance as well as new features. Compression is an active research area in which great effort has been made to deliver new standard by providing features inexistent in previous standards, but also by providing higher efficiency for features that exist in other.

LITERATURE SURVEY

A good algorithm for compression is the hybridization of discrete cosine transform on the discrete wavelet transform coefficients. This algorithm performs much better in terms of peak-signal to noise ratio with a higher compression ratio compared to standalone DCT and DWT algorithms. This scheme is intended to use as the image/video compressor engine in medical imaging and video applications, such as telemedicine and wireless

capsule endoscopy.

A good compression algorithm of higher compression ratio based on SPIHT (Set Partition in Hierarchical Tree). This algorithm shows a higher peak signal to noise ratio (PSNR) with same compression ratio.

A good algorithm which outperforms all other previous Algorithm is the Discrete Cosine Transform which attempts to decorrelate the image and after de correlation each transform coefficient can be encoded independently without losing compression efficiency. Like other transforms, the Discrete Cosine Transform (DCT) attempts to de correlate the image data. After de correlation each transform coefficient can be encoded independently without losing compression efficiency. Discrete Cosine Transform (DCT) has emerged as the de-facto standard for image transformation in most visual systems. DCT has been widely deployed by modern video coding standards, for example, MPEG, JVT etc. This article introduces the DCT, elaborates its important attributes and analyzes its performance using information theoretic measures. Discrete Cosine Transform is widely used in such problems of digital signal processing as feature extraction, filtering, and especially, in image compression. DCT has very good energy compaction properties and near optimal performance which is closest to that of the Karhunen-Loeve transform (KLT). The JPEG and MPEG coding algorithms, based on DCT, have been established as industry standards for image compression.

In Discrete Cosine Transform, computation complexity is quite large. For reducing the computational complexity several fast DCT algorithms has developed such as, DCT based on the Residue Number System (RNS). This algorithm uses the combination of Fast Cosine Transform (FCT) algorithm and the matrix-vector multiplication (MVM). RNS based FCT MVM implementation provides a throughput improvement over the equivalent binary system up to 72%, while its advantage over the binary distributed arithmetic implementation is up to 128%. A custom architecture for Discrete Cosine Transform (DCT) using No-Instruction Set Computer (NISC) technology that is developed for fast processor customization. Using several software transformations and hardware customization, performance achievement is 10 times more, 2 times power reduction, 12.8 times energy reduction, and 3 times area reduction compared to an already optimized soft core MIPS implementation. In DCT computation, complexity increases mainly for

huge amount of multiplication. So, idea arises that, if it is multiplication less, then computation can be faster. A good technique is that which implements direct 2-D DCT approach with regular butterfly structure.

This architecture employs eight 1-D DCT processors and four post-addition stages to calculate two-dimensional DCT coefficients. Each 1-D DCT processor is designed using Algebraic Integer Encoding architecture which requires no multiplier, therefore the entire 2-D DCT design is multiplier less. Critical path is shortened by employing five pipeline stages on each compression and decompression system. A good architecture is one which uses the 2-D DCT separability property, such that the whole architecture is divided into two 1-D DCT calculations by using a transpose buffer. A new alternative of DCT is Integer Cosine Transform (ICT) used for image processing. Parallel-pipeline architecture of an 8 x 8 ICT processor for image compression whose characteristics are high throughput, low latency, reduced internal storage and 100% efficiency in all computational elements. A good architecture of an efficient implementation of a two-dimensional DCT/IDCT transforms processor via a serial-parallel systolic array that does not require transposition. A fast two-dimensional discrete cosine transform optimized for software implementation on a RISC microprocessor with an integer multiplier accumulator (MAC) unit. The number of processor cycles as well as computational error is less than that of the row-column approaches based on fast 1D DCTs.

A new processor that computes the Discrete Cosine Transform (DCT) which overcomes the major drawbacks of the original implementation resulting in a design with considerably less area consumption and higher speed. To achieve these results, a novel architecture, based on the CORDIC Circular Rotation Algorithm, is introduced; it reduces the required area by more than 60% compared to the use of standard CORDIC architectures. Furthermore, bit serial arithmetic is used, resulting in a very compact design. In order to get maximum throughput ,the processor is fully pipelined, achieving a Performance efficient even for signals as fast as HDTV.

PROPOSED METHODOLOGY

The conventional memories that we use either have single address and single data output for ROM or separate read and write addresses and data bus for a dual port RAM. In real time systems such as video processing systems, these conventional memories may not be of help since the typical applications often demand access to two memory locations simultaneously, write word-wise and read column-wise, etc. These application-specific requirements, arising mainly due to the need for efficient implementation of computationally intensive algorithms.

• The ROM stores the cosine terms 2 × C instead of C in order to improve accuracy.

• Two-stage pipelining for 2 × C matrix of cosine terms to keep pace with dual RAM used in the DCTQ design.

• ROM size is 8 × 64 bits. Two locations, each of size 64 bits, can be accessed and output to the data bus, 'dout1' and 'dout2' simultaneously using the two addresses, 'addr1' and 'addr2' respectively.

This ROM, with single address and single data output, is used to store inverse quantization values. In quantization process, we need to divide the DCT coefficients by the corresponding quantization values. However, division can also be implemented as multiplication if we take the inverse of the quantization values. The dual RAM consists of two RAMs, each of which stores the image information. This information will be written from a host computer such as a PC into one of the RAMs through peripheral connect interface (PCI) bus. Initially, one of the double memory buffers, RAM 1,is filled and once it is full, the image information is written to the second RAM. While the second memory, RAM 2, is being written into, the RAM 1 will be read concurrently to process the DCTQ coefficients. If RAM 1 is in read only mode, then RAM 2 is automatically configured to the write only mode and vice-versa. The RAM is written row-wise and read column-wise. This is due to the complexity of the DCT algorithm (While reading, the RAM is accessed column-wise, since in the computation of C*X of the DCT algorithm we need to multiply a row of C matrix with the column of X (image input) matrix as explained).

Serial and Parallel adders are chosen depending on Application type and chip area. Multipliers are also Pipelined to increase the computation speed.The two-stage matrix multiplication can be implemented by parallel architecture, where in eight partial products,

which are the row vectors of C*X generated in the first stage, are fed to the second stage. Subsequently, multiplying row vector of CX by the CT matrix generate seight DCT coefficients. Quantized outputs can be obtained by dividing each of the 64 DCT coefficients by the corresponding quantization values in the Single address ROM.

Algorithm for Parallel Matrix Multiplication for DCTQ

DCT is an orthogonal transform consisting of a set of vectors that are sampled cosine functions. 2D-DCT of a block of size 8 × 8 pixels of an image is

$$DCT = \begin{bmatrix} c_{00} & c_{01} & \dots & c_{07} \\ c_{10} & c_{11} & & c_{17} \\ \vdots & \vdots & & \vdots \\ c_{70} & c_{71} & & c_{77} \end{bmatrix} \begin{bmatrix} X_{00} & X_{01} & \dots & X_{07} \\ X_{10} & X_{11} & \dots & X_{17} \\ \vdots & \vdots & & \vdots \\ X_{70} & X_{71} & \dots & X_{77} \end{bmatrix} \begin{bmatrix} c_{00} & c_{10} & \dots & c_{70} \\ c_{01} & c_{11} & \dots & c_{71} \\ \vdots & \vdots & & \vdots \\ c_{07} & c_{17} & \dots & c_{77} \end{bmatrix}$$

defined as

$$DCT(u,v) = \frac{1}{4} c(u) c(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x, y) \left[\cos \frac{(2x+1) u\pi}{16} \right] \left[\cos \frac{(2y+1) v\pi}{16} \right]$$

where f (x, y) is the pixel intensity and
$$c(u) = c(v) = 1/\sqrt{2} \text{ for } u = v = 0 \text{ and}$$
$$= 1 \text{ for } u, v = 1 \text{ to } 7.$$

The DCT can be expressed conveniently in a matrix form: DCT = C X C^T .where X is the input image matrix, C the cosine coefficient matrix, and C^T, its transpose with constants (1/2)c(u) and (1/2)c(v) absorbed in C and C^T matrices respectively. For a clearer understanding, the DCT may be expressed in an expanded form:

$$p_{jk} = \sum_{i=0}^{7} c_{ji}\, x_{ik}$$

$$= \begin{bmatrix} \sum_{i=0}^{7} p_{0i}c_{0i} & \sum_{i=0}^{7} p_{0i}c_{1i} & \cdots & \sum_{i=0}^{7} p_{0i}c_{7i} \\ \sum_{i=0}^{7} p_{1i}c_{0i} & \sum_{i=0}^{7} p_{1i}c_{1i} & \cdots & \sum_{i=0}^{7} p_{1i}c_{7i} \\ \vdots & \vdots & & \vdots \\ \sum_{i=0}^{7} p_{7i}c_{0i} & \sum_{i=0}^{7} p_{7i}c_{1i} & & \sum_{i=0}^{7} p_{7i}c_{7i} \end{bmatrix}$$

$$= \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{07} \\ p_{10} & p_{11} & \cdots & p_{17} \\ \vdots & \vdots & & \vdots \\ p_{70} & p_{71} & \cdots & p_{77} \end{bmatrix} \begin{bmatrix} c_{00} & c_{10} & \cdots & c_{70} \\ c_{01} & c_{11} & \cdots & c_{71} \\ \vdots & \vdots & & \vdots \\ c_{07} & c_{17} & \cdots & c_{77} \end{bmatrix}$$

Where

The two-stage matrix multiplication shown above can be implemented by parallel architecture, wherein eight partial products, which are the row vectors of CX generated in the first stage, are fed to the second stage. Subsequently, multiplying row vector of CX by the C^T matrix generates eight DCT coefficients, corresponding to a row of C X C^T. While computing the $(i+1)^{th}$ partial products of CX, the ith row DCT coefficients can also be computed simultaneously since the ith partial products of CX are already available. Application of DCT on an 8 × 8 pixel block, thus, generates 64 coefficients in a raster scan order.

Quantized outputs can be obtained by dividing each of the 64 DCT coefficients by the corresponding quantization table values given in the standards [27] as per the expression:

DCTQ (u,v) = DCT (u,v) / q (u,v) ; u, v = 0 to 7

These stages can be pipelined in such a way that one DCTQ output can be generated every clock cycle.

BLOCK DIAGRAM

Raw image data → Dual Ram (2^6 *8 bit) (Provides simultaneous Reading and Writing) → Arithmetic operation required to obtain DCT Coefficient → RAM (To write the o/p data)

ADVANTAGES

The main advantages of this system are reductions in storage hardware, data transmission time, and communication bandwidth. This can result in significant cost savings. Compressed files require significantly less storage capacity than uncompressed files, meaning a significant decrease in expenses for storage. A compressed file also requires less time for transfer while consuming less network bandwidth. This can also help with costs, and also increases productivity.

CONCLUSION

In this paper we have designed a system whose inputs are pixel values of an image and the output is the computed DCTQ of those pixels. The system computes DCTQ of 64-pixel values and outputs each value at the end of each clock cycle. The core of the project is to design and verify the most efficient Arithmetic circuits which we have obtained by using pipeline concept. Pipelining increases the throughput as well as chip area. As the cost of silicon is pretty less and it is abundantly available on earth our focus in mainly on increasing the overall systems performance.

REFERENCES

[1] Kapil Ram Gavali, Poonam Kadam, Dwarkadas J "A High Throughput Architecture of DCTQ Processor suitable for FPGA Implementation utilizing Fox Algorithm".

[2] Yogesh M. Jain, Aviraj R. Jadhav and Harish V. Dixit, Akshay S. Hindole and Jithin R. Vadakoott, Devendra S. Bilaye "A Novel VLSI Design of DCTQ Processor for FPGAImplementation".

[3] Digital VLSI Systems Design by Dr. S. Ramachandran.Saurabh Chaudhury, Debashish Dash" Design and Implementation of a DCTQ Processor for Low Area, Power and High Performance"

[4] K.R. Rao and P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press, New York, NY, 1990.

[5] B.G. Lee, A new algorithm to compute the discrete cosine transform, IEEE Trans. Acoust., Speech and Signal Proc., ASSP-32, pp. 1243–1245, 1984.

[6] B.G. Lee, Input and output index mappings for a prime-factor-decomposed computation of discrete cosine transform, IEEE Trans. Acoust., Speech and Signal Proc., ASSP-37, pp. 237–244, 1989.

[7] A.N. Skodras, and A.G. Constantinides, Efficient input-reordering algorithms for fast DCT, Electron. Lett., 27, pp. 1973–1975, 1991.

[8] P. Lee and F.Y. Huang, an efficient prime-factor algorithm for the discrete cosine transform and its hardware implementations, IEEE Trans. Signal Process., 42, pp. 1996– 2005, 1994.

[9] C.L. Wang and C.Y. Chen, High throughput VLSI architectures for the 1-D and 2-D discrete cosine transforms, IEEE Trans. Circuits Syst. Video Technol., 5, pp. 31–40, 1995

[10] Yung-Pin Lee, Thou-Ho Chen, Liang-Gee Chen, Mei-Juan Chen and Chung-Wei Ku, A cost-effective architecture for 8 × 8 2-D DCT/IDCT using direct method, IEEE Trans. Circuits Syst. Video Technol., 7, 1997.

[11] Yukihiro ARAI, Takeshi AGUI and Masayuki NAKAJIMA, A fast DCT-SQ scheme For images, Trans. IEICE, E71, pp. 1095–1097, 1997.

[11] Yi-Shin Tung, Chia-Chiang Ho and Ja-Lung Wu, MMX-based DCT and MC Algorithms for real-time pure software MPEG decoding, IEEE Computer Society Circuits and Systems, Signal Processing, 1, Florence, Italy, pp. 357–362, 1999.

[12] H.S. Hou, A fast recursive algorithm for computing the discrete cosine transform, IEEE Trans. Acoustics, Speech: Signal Proc., ASSP-35, pp. 1455–1461, 1987.

[13] C. Loeffler, A. Ligtenberg and G.S. Moschytz, Practical fast 1-D DCT algorithms with 11 multiplications, Proceedings of IEEE ICASSP, 2, pp. 988–991, 1989.

[14] N.I. Cho and S.U. Lee, DCT algorithms and VLSI implementations, IEEE Trans.Acoust., Speech and Signal Process., ASSP-38, pp. 121–127, 1990.

[15] N.I. Cho and S.U. Lee, Fast algorithm and implementation of 2-D discrete cosine transform, IEEE Trans. Circuits Syst., 38, pp. 297–305, 1991.

[16] Y.P. Lee, T.H. Chen, L.G. Chen, M.J. Chen and C.W. Ku, A cost-effective architecture for 8 × 8 2D-DCT/IDCT using direct method, IEEE Trans. Circuits Syst. Video Technol., 7, pp. 459–467, 1997.

[17] M. Yoshida, H. Ohtomo and I. Kuroda, A new generation 16-bit general purpose programmable DSP and its video rate application, IEEE Workshop on VLSI Signal Processing, pp. 93–101, 1993.

[18] I.Kuroda, Processor architecture driven algorithm optimization for fast 2-D DCT, IEEE Workshop on VLSI Signal Processing, VIII, pp. 481–490, 1995.

PLACEMENT MANAGEMENT SYSTEM

Akash P. Galagali
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: akashgalagali4@gmail.com

Chethan
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: chethanjkulkarni@gmail.com

Mrs. Mamatha K.R
(Assistant professor)
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: mamathakr@bmsit.in

Debanshu Biswas
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: debanshubiswas111@gmail.com

Siddesh S
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: 1by17ec164@bmsit.in

Abstract— As we are moving forward in life, we are facing challenges everyday and till date we are trying to have a better solution to it. With the increasing number of human population, the crave for jobs and recruitment is also increasing in heavy demand day by day and this won't stop ever. Rather it would get worse. So, looking into this factor the placement offices are facing problems to conduct the recruitment process smoothly. They are looking forward to some automated and online process so that the officers can take a break and cut down their stress of work duty. In the view of this we bring you "Placement Management System."

Keywords— ◎ Django ◎ HTML ◎ CSS ◎ JS

◎ GIT ◎ Docker ◎ Bash Scripting ◎ Nginx

I. INTRODUCTION

The project "PLACEMENT MANAGEMENT SYSTEM" is a web-based application developed for the placement department of the college in order to provide the details of its students in a database for the companies to their process of recruitment. The main purpose of the System is to avoid manual process because the manual work makes the process slow and other problems such as inconsistency & ambiguity on operations. In order to avoid this web-based placement managed system is proposed, where the student information in the college with regard to placement is managed efficiently. It intends to help fast in fast access procedures in placement related activities and ensures to maintain the details of the student. Students logging should be able to upload their personal and educational information. The key feature of this project is that it is onetime registration enabled. The placement cell calls the companies to select their students for jobs via the campus interview. The placement cell allows the companies to view the student resumes in selective manner. They can filter the student profile as per their requirement. The job details of the placed students will be provided by the administrator. The administrator plays an important role in our project. Our project provides the facility of maintaining the details of the students and gets the requested list of candidates for the company who would like to recruit the students based on given criteria.

Figure 1: Student Registration Process

Figure 2: Working at the Admin Panel

II.   PROBLEM ANALYSIS

Nowadays campus placements are conducted in all colleges. Various software and other sector companies are conducting campus selections for selecting merit candidates. When campus selections are conducted the students should provide their curriculum vitae to the concern officer for attending the campus interviews. This routine process is maintained manually, like maintenance of their resumes in papers. This can be automated by designing software.

The old manual system was suffering from many drawbacks. The PMS System is to maintain the data and make the process smooth, maintaining the data and retrieving information is very tedious and lengthy. there used to be lots of difficulties in associating any particular transaction with particular context. It is very difficult to find errors while entering the records. Once the records were entered it was very difficult to update records.

Main reason behind it is a lot of information to be maintained and managed. For this reason we provide features in PMS that can potentially automate the whole system and reduce the errors and manual work.

In Colleges, Training and Placement Officers(TPO) have to manage the student profile and data for training and placement process manually. Placement officers have to collect the information from companies and notify the students who are clearing the company criteria and this

manual process is tedious and erroneous. Any modification in the details of a student or updates required in the profile of the student has to be searched and get done manually.

III.   PROPOSED SYSTEM

Proposed system is inexpensive and no manual workload is needed. Everything, however smaller the details for the registration process are, they are going to get into the system through this Online Placement Management System.

The software used in this regards are as listed below:
- Django
- HTML
- CSS
- JS
- GIT
- Docker
- Bash Scripting
- Nginx

All these softwares are easily available in the market. If need be any college can develop it for the better conduction of placement process.

Below given are the basic functionalities of the softwares used in this project:

- Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design.

- Hypertext Markup Language (HTML) is the standard markup language for documents designed to be displayed in a web browser.

- Cascading Style Sheets(CSS) is designed to enable the separation of presentation and content, including layout, colors, and fonts.

- JavaScript(JS) conforms to the ECMAScript specification and is one of the core technologies of the World Wide Web.

- GIT is a distributed version-control system for tracking changes in any set of files, originally designed for coordinating work among programmers cooperating on source code during software development.

- Docker is a set of platform as a service (PaaS) products that use OS-level virtualization to deliver software in packages called containers.

- Bash Scripting helps in executing a shell command, running multiple commands together, customizing administrative tasks, performing task automation etc.

- Nginx is used as a web server designed for maximum performance and stability.

IV.   FLOW OF THE PROCESS

In this project we will be using Django and our main framework backing up most of the authentication, communication with the frontend and the database, and effectively handling traffic, sending emails, sending SMS and much more. Front end will be handled by HTML, CSS and JS. JavaScript will be handling all the robust frontend actions like making XHR requests, visibility of elements, interactions with the users, etc. MySQL will be the database holding all the information about the students and PTO.

EXISTING SYSTEM	PROPOSED SYSTEM
1. Manual SGPA CGPA calculation	1. Automatic CGPA SGPA calculation
2. Manual shortlisting	2. Can generate Shortlist with Criteria
3.Notifying students individually	3. Notifying students on one click
4. No option for students to update Resume	4. Can Update resume at any time
5. Collect Feedback Manually	5. Can write Feedback on site
6. Manual Reminding about Training and Placement Fees	6. Automatic Reminding about Training and Placement Fees

Figure 4: Allover Statistical Display of Registrations

Figure 5: Notification box

Applications: -

1. It can be used in colleges to ease the way of the placement campus drive.
2. On a large scale it will help in the recruitment process more efficiently.

Advantages: -

1. Easy to Install
2. No Manual workload
3. Easy to use
4. More accuracy

V.   RESULTS

This paper depicts how the system will work and how smooth the conduction process is going to take place from the very beginning of registration till the recruitment of the individuals and how stress-free the process is this for the placement officers.

VI.   CONCLUSION

According to various surveys done, this turns out really well for both students as well as the officers and the companies' HRs too. We hope to look forward to implementation of this system not only in our college but in all the colleges of this region as well as the country. This way the process of recruitment is going to be hassle-free.The main advantage of this paper is that it suggests the solution to a very existing

problem at a very low cost. The accuracy of the system is no doubt of high rate for good possible outputs.

VII.　ACKNOWLEDGMENT

VIII.　REFERENCES

- https://www.slideshare.net/
- https://acadpubl.eu/
- http://ijcsit.com/
- http://data.conferenceworld.in/

IMPLEMENTATION OF 3-DIMENSIONAL DISCRETE WAVELET TRANSFORMS (3D-DWT) USING LIFTING BASED ALGORITHM

Hamsavahini R[1]*,
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Mamatha K R[2]
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Abstract— **In this paper we implemented 3D discrete wavelet Transform(DWT) using lifting based algorithm .The lifting based DWT architecture has the advantage of low power computational complexities transforming the signals with extension and regular data flow. We also provide a survey on these architectures for both 1-D and 2-D DWT. 3-D DWT uses a cascade combination of three 1-D wavelet transforms.**

KEYWORDS: DWT, LIFTING, IMAGE COMPRESSION

I.INTRODUCTION

Wavelet transforms is one of the most modern areas of mathematics. In the last few decades, there has been an enormous increase in the applications of wavelets, in which it is used as an analytical tool in various areas of technical research such as electronics, computers and telecommunication systems that require efficient speed, resolution & real time memory and optimization with maximum hardware utilization. A wavelet, as the name suggests, is a little piece of a wave. They are used in representing data and other functions that satisfy certain mathematical requirements. The fundamental idea behind wavelets is to analyze according to scale. Wavelet algorithms process data at different scales or resolution. Wavelet transforms offer a wide variety of useful features in contrast to other transforms such as Fourier or cosine transform. Fourier transform is a powerful tool for analyzing the components of stationary signal, but it failed for analyzing non-stationary signal, whereas wavelet transforms allows the components of a non-stationary signal to be analyzed by using a set of damped oscillatory functions known as wavelet basis. Wavelet transforms in its continuous form is represented as CWT. A discrete and fast implementation of CWT is known as standard DWT. It decomposes the signal into different sub-bands with both time and frequency and facilitates to arrive at high compression ratio.

Recent advances in implementation of image coding and progressive image transmission through DWT demand both a large number of computations and large storage features that are not desirable for either high speed or low power applications. A methodology for implementation of the above complex applications have been proposed known as lifting based DWT that often has far fewer computations than DWT. In lifting scheme, the signal is divided like a zipper and a series of convolution-accumulate operations across the divided signals is applied, since this technique applies to each of this individual divided signals. Finally, brief summaries are given in below sections to conclude the paper.

II.WAVELETS

2.1. Wavelet Definition: A 'wavelet' is a small wave which has its energy concentrated in time. It has an oscillating wavelike characteristic and is a suitable tool for transient, non-stationary or time-varying phenomena.

2.2 Wavelet Characteristics: The difference between wave (sinusoids) and wavelet is shown in figure 1.1. Waves are smooth, predictable and everlasting, whereas wavelets are of limited duration, irregular and may be asymmetric. Waves are used as deterministic basis functions in Fourier analysis for the expansion of functions (signals), which are time-invariant, or stationary. The important characteristic of wavelets is that they can serve as deterministic or non-deterministic basis for generation and analysis of the most natural signals to provide better time-frequency representation, which is not possible with waves using conventional Fourier analysis

(a)   (b)

Fig.2. Representation of a (a) wave (b) wavelet

2.3 Wavelet Analysis: The wavelet analysis procedure is to adopt a wavelet prototype function, called an 'analyzing wavelet' or 'mother wavelet'. Temporal analysis is performed with a contracted, high frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low frequency version of the same wavelet. Mathematical formulation of signal expansion using wavelets gives Wavelet Transform (WT) pair, which is analogous to the Fourier Transform (FT) pair. Discrete-time and discrete-parameter version of WT is termed as Discrete Wavelet Transform (DWT).

2.4. Advantages of Wavelets:
*(a).*Wavelet analysis is an exciting new method for solving difficult problems in mathematics, physics, and engineering, with modern applications as diverse as wave propagation, data compression, signal processing, image processing, pattern recognition, computer graphics, the detection of aircraft and submarines and other medical image technology.
*(b).*Wavelets allow complex information such as music, speech, images and patterns to be decomposed into elementary forms at different positions and scales and subsequently reconstructed with high precision.
*(c).*Wavelets are a powerful statistical tool which can be used for a wide range of *applications*, namely: Signal processing, Data compression, Fingerprint verification, Blood-pressure, heart-rate and ECG analyses, DNA analysis, and protein analysis.

III.DISCRETE WAVELET TRANSFORM

3.1. Wavelet Transforms: Wavelet transform has gained widespread acceptance in speech, image and video processing, and in signal processing in general because of its attractive characteristics to represent non-stationary signals in both frequency and time domains. The wavelet transform is computed separately for different segments of the time-domain signal at different frequencies. *Multi-resolution analysis*: analyzes the signal at different frequencies giving different resolutions. Multi-resolution analysis is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Good for signal having high frequency components for short durations and low frequency components for long duration, e.g. Images and video frames.
3.2. Discrete Wavelet Transforms: It is a wavelet transform for which the wavelets are discretely sampled. DWT of a signal x(n) is calculated by passing it through a series of filters. First the samples are passed through a low-pass filters with impulse response g(n) resulting in a convolution of the two.

$$\mathbf{Y[n]=(x*g)[n]=} \sum_{k=-\infty}^{\infty} \boldsymbol{x[k]g[n-k]} \quad \mathbf{(1)}$$

The signal is also decomposed simultaneously using a high-pass filter h(n). The outputs giving the detail coefficients(from the high pass filter)and approximate coefficients(from low pass filter) as shown in fig.3.These 2 filters are related to each other and they are known as a quadrature mirror filters. The frequency of the signal have now been removed, half the samples can be discarded according to Nyquist'srule. The filter output's are then sub sampled by 2(g-high pass and h-low pass).

$$\mathbf{Y_{low}[n]=}\sum_{k=-\infty}^{\infty} \boldsymbol{x[k]g[2n-k]} \qquad \mathbf{(2)}$$

$$\mathbf{Y_{high}[n]=}\sum_{k=-\infty}^{\infty} \boldsymbol{x[k]h[2n-k]} \qquad \mathbf{(3)}$$

This decomposition has halved the time resolution since only half of each filter output characterizes the signal.Each output has the frequency band of the input so the frequency resolution has been doubled.

Fig.3. One-Dimensional DWT

IV.TWO-DIMENSIONAL DWT

One-Dimensional DWT can be easily extended to two dimensions which can be used for the transformation of two dimensional images. A two dimensional digital image which can be represented by a 2-D array x[n1,n2] with n1 rows and n2 columns, where n1 and n2 are positive integers.First, a one-dimensional DWT is performed on rows to get low frequency 'L' and high frequency 'H' components of the image.The fundamental block diagram of a 2-D DWT is as shown in fig 4a.

Fig.4 Two-Dimensional DWT

At each scale, an image f(x,y) is decomposed into an approximation image a_j of a low-pass band, and three detail images d_j^x, d_j^y, d_j^{xy} corresponding to a horizontal high-pass

band d_j^x, a vertical high-pass band d_j^y, and a diagonal high-pass band d_j^{xy} as given below:

$$a_j(x,y) = \left\langle f(x,y), \phi_j(x)\phi_j(y) \right\rangle$$
$$d_j^x(x,y) = \left\langle f(x,y), \psi_j(x)\phi_j(y) \right\rangle$$
$$d_j^y(x,y) = \left\langle f(x,y), \phi_j(x)\psi_j(y) \right\rangle$$
$$d_j^{xy}(x,y) = \left\langle f(x,y), \psi_j(x)\psi_j(y) \right\rangle$$

Compare with single dimension case

$$Wf(b,s) = \int f(t) \frac{1}{\sqrt{s}} \psi\left(\frac{t-b}{s}\right) dt = f * \overline{\psi}_s(b)$$

$$= \;\; <f(t),\ \psi(t)> \tag{6}$$

Due to down-sampling of the coefficients, the size of the image at lower level is half of that of the higher level. In two-dimensional image analysis, the three wavelets are given by:

$$\psi^1(x,y) = \varphi(x)\psi(y)$$
$$\psi^2(x,y) = \psi(x)\varphi(y)$$
$$\psi^3(x,y) = \psi(x)\psi(y) \tag{7}$$

Extract image details at different scales and orientations. At each scale, we end up with three "detail" images:(a) Low-pass filtered in the x-direction and high-pass filtered in the y direction, $\psi^1(x,y) = \varphi(x)\psi(y)$, yielding detail D^x. (b)Low-pass filtered in the y direction and high-pass filtered in the x direction, $\psi^2(x,y) = \psi(x)\varphi(y)$ yielding detail D^y, and, (c)finally, high-pass filtered in both x and y directions yielding detail D^{xy}

I_2	D_2^x	D_1^x
D_2^y	D_2^{xy}	
D_1^y		D_1^{xy}

Fig.5 Components of decomposed image, three details and one approximation at each scale.

So, we have three orientations for details: Horizontal, Vertical and Diagonal. To cover the entire image using one dimensional wavelets, we consider image composed of rows or columns which are one dimensional in rows or columns. We apply wavelet transform for each row (i.e. keeping x

constant but changing y) starting from the top row. Then apply wavelet transform on the results of row operations but now move column-wise starting from the left column where we use wavelet or scaling function depending on whether we want LL, LH, HL or HH signal components. Since we are using one dimensional wavelet transform at each scan of x or y direction, we have only two filters low pass 'L' and high pass 'H' where we use them in sequence in x and y directions. To obtain low frequency band in x and y direction, we need to use to low pass filters as shown in Fig.4c Similarly we use LH, HL and HH for other frequency bands.

Fig.6. Different levels of decomposition of an image

V.3-DIMENSIONAL DWT

The 3D DWT can be considered as a combination of three 1D DWT in the x, y and z directions as shown in the fig. 5. The preliminary work in the DWT processor design is to build 1D DWT modules, which are composed of high-pass and low-pass filters that perform a convolution of filter coefficients and input pixels. After a one-level of 3D discrete wavelet transform, the volume of image is decomposed into HHH,HHL,HLH,LHL,LHH,LLH and LLL signals as shown in fig.5.The process results in 8-data stream. The approximate signal resulting in scalar operations only goes to the next octave of the 3D transform. It has roughly 90 percent of the total energy. Meanwhile, the seven other streams contain the detail signals. Note that though conceptual drawing of the 3D DWT for one octave has 7 filter phases, this doesn't mean that the process needs 7 physical pairs. For example, a folded architecture maps multiple filters onto one filter pair.

Fig.7. 3-Dimensional DWT

VI.LIFTING BASED DWT

The development of the lifting scheme was proposed for the construction of bi-orthogonal wavelets. The main feature of the lifting scheme is that all constructions are derived in the spatial domain. It does not require complex mathematical calculations that are required in traditional methods. Lifting scheme is simplest and efficient algorithm to calculate wavelet transforms. It does not depend on Fourier transforms. Lifting scheme is used to generate second - generation wavelets, which are not necessarily translation and dilation of one particular function.

The memory modules are divided into multiple banks to accommodate high computational bandwidth requirements .a number of architectures have been proposed for calculation of the convolution based DWT. The architectures are mostly folded and can be broadly classified into series (where the inputs are supplied to the filters in a serial manner) and parallel architectures (where the inputs are supplied to the filters in a parallel manner). The basic principle of lifting scheme is to factorize the poly-phase matrix of a wavelet filter into a sequence of alternating upper and lower triangular matrices and diagonal matrix. This leads to the wavelet implementation by means of a banded-matrix multiplications.

Constructing wavelets using lifting scheme consists of three steps: (a)Split step: where the data is split-up into odd and even sets. (b)Predict step: in which odd set is predicted from even set. Predict phase ensures polynomial cancellation in high pass. (c)Update phase: that will update even set using wavelet coefficient to calculate scaling function. Update stage ensures preservation of moments in low pass. Block diagram of forward Lifting scheme is as shown in fig.6a.

6.1. Working of lifting scheme:

The basic idea behind lifting scheme is that, it tries to use the correlation in the data to remove redundancy. First split the data into 2 sets(split phase) i.e., odd sample and even sample. Because of the assumed smoothness of the data, we predict that odd samples have a value that is closely related to their neighboring even samples. We use N even samples to predict the value of a neighboring odd value (predict phase).With a good prediction method, the chance is high that the original odd sample is in the same range as its prediction. We calculate the difference between the odd sample and its prediction and replace the odd sample with this difference.

6.2. For Image Compression:
In each lifting stage, the predicting or updating operations are carried out in the direction of image edges and textures in a local window, and are not necessarily horizontal or vertical. This adaptation can significantly reduce the signal energy of high-pass sub-bands. High angular resolution in prediction is achieved by the use of fractional pixels in prediction and update operations. The fractional pixels can be calculated by any existing interpolation method. In order to guarantee perfect reconstruction, the predicted and updated samples are always in integer pixel positions.

Fig.8. Block diagram of forward Lifting scheme

Fig.9. Architecture of DWT based on lifting scheme

6.3.Reasons for the choice of Lifting scheme

We have used lifting scheme of wavelet transform for the digital speech compression, because lifting scheme is having following advantages over conventional wavelet transform technique. (1)It allows a faster implementation of the wavelet transform. It requires half number of computations as compare to traditional convolution based discrete wavelet transform. This is very attractive for real time low power applications. (2)The lifting scheme allows a fully in-place calculation of the wavelet transform. In other words, no auxiliary memory is needed and the original signal can be replaced with its wavelet transform. (3)Lifting scheme allows us to implement reversible integer wavelet transforms. In conventional scheme it involves floating point operations, which introduces rounding errors due to floating point arithmetic. While in case of lifting scheme perfect reconstruction is possible for loss-less compression.

VII. RESULTS

Consider an color RGB image of size 225x225 and convert it into a gray scale image which is then resized into 256x256 as shown in fig 10.

Fig 10. Resized Input Image

We extract the required wavelet co-efficients and combine it with corresponding filter coefficients. Divide the image coefficients into odd and even indexed values. Convolute filter coefficients with image coefficients taking two adjacent pixel values at a time. After performing predict and update step column wise first and row wise next we get 128x128 pixel image which is the output of 1D DWT as shown in fig b.

fig 11. Output Of 1D DWT

The output of 1D DWT is given as an input to the cascaded filters of 2D DWT where the image is compressed further that leads to 64x64 pixel image as shown in fig c.

Fig 12. Output of 2D DWT

 The output of 2D DWT is given as an input to the cascaded filters of 3D DWT. The final image is the compressed image(32x32) of the original input image due to normalization/down sampling which is shown in fig d.

Fig 13. Output of 3D DWT

VII.COMPARISIONS

At first, the comparision between different stages of DWT implemented in this paper using lifting based algorithm is tabulated below:

Levels of dwt	Extent of relative compression
1st level DWT	2:1
2nd level DWT	4:1
3rd level DWT	8:1

Next, the time difference in computation speed between normal dwt and discrete wavelet transforms implemented using lifting based algorithm is shown below in fig 14 and fig 15.

As we have already seen the advantages of using lifting based algorithm that it takes less time to compute than normal DWT. As we see in fig 14. that normal dwt takes a total time of 1.108s to process an image. Where as lifting based algorithm shown in fig 15. Takes the total time of 1.044s to process same image ,there is a time difference of 0.064s for just one image but in real world we deal with millions of images at a time where time constraint is more effective and hence lifting based occupies high priority.

Fig 14. Time simulation results of normal DWT

Fig 15. Time simulation results of lifting based DWT

CONCLUSION

In this paper, we implemented the lifting based implementations of 1-dimensional, 2-dimensional and 3-dimensional Discrete Wavelet Transform. We briefly described the principles behind the lifting scheme in-order to better understand the different implementation styles and structures.

The result of this implementation is the compressed image whose relative compression is 8 times the input image as shown in simulation figures.

REFERENCES

[1] Senthilkumar.M ,Uma .S " High speed 3d DWT VLSI architecture for image processing using lifting based wavelet transform".International journal of communication and computer technologies. Vol 01-no.36.

[2] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecturefor Image Processing".IMACST: vol 3.

[3] Kishore Andra, Chaitali Chakrabarti, member IEEE, and Tinku Acharya, sentor member IEEE" A VLSIfor lifting based forward and inverse wavelet transform".IEEE transactions on signal processing. Vol 50,no. 4.

[4] Malay Ranjan Tripathy, Kapil Sachdeva and Rachid Talhi "3D Discrete wavelet transform VLSI architecture for Image processing" progress in Electromagnetics research symposium proceedings Moscow, Russia.

[5]Patrick Longa " An Optimized architecture for 2D discrete wavelet transform on FPGA's usingdistributed arithmetics 2006.

[6] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," IEEE Signal Processing, vol. 8, issue: 4,pp. 14-38, Oct. 1991.

[7] ISO/IEC. International Standard, 15444-1: 2000(E), JPEG2000 Image Coding System – Part I Core coding system.

[8] C. Chakrabarti, M. Vishwanath, and R. M. Owens, "Architectures for wavelet transforms: A survey," J. VLSI Signal Process., vol. 14, pp.171–192, 1996.

[9] W. Sweldens, "The lifting scheme: A construction of second-generation wavelets," SIAM J. Mathematical Analysis, vol. 29, no.2, pp. 511–46, 1997.

[10] P. Chen, "VLSI implementation for one-dimensional multilevel lifting-based wavelet transform," IEEE Trans. on Computers, vol.53, no.4, pp.386-398, April 2004.

[11] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecture for Image Processing".IMACST: vol 3.

[12] N.D. Zervas, G.P. Anagnostopoulos, V. Spiliotopoulos,Y. Andreopoulos, and C.E. Goutis, "Evaluation of Design Alternatives for the 2-D Discrete Wavelet Transform," IEEE Trans. Circuits and Syst. Video Tech., Vol. 11, No. 12, pp. 1246-1262, December 2001.

[13] C. Chrysafis and A. Ortega, "Line-Based, Reduced Memory, Wavelet Image Compression," IEEE Trans.Circuits and Syst. Video Tech., Vol. 9, No. 3, pp.378-389, March 2000.

Gateway Based Hybrid Hierarchical Routing Scheme with Time Synchronization in WSN

Raghunandan G. H[1] and A Shobha Rani[2]
[1]*Department of Electronics and Telecommunication Engineering*
[2]*Department of Electronics and Communication Engineering*
[1,2]*BMS Institute of Technology and Management, Bengaluru, India*

ABSTRACT

Wireless Sensor Network are spatially distributed sensors intended to monitor different physiological conditions. Sensing and communicating data from one place to another consumes more energy, therefore the management of sensor energy is very important factor. Energy utilization, synchronization and lifetime of the network is the main criteria in WSN. More energy is lost by sensors which are far from the base station.The cluster head is deployed to collect and relay information from nodes to the base station or gateway nodes to resolve this problem. To decrease energy consumption, gateway nodes are deployed between the cluster head and the base station. In this paper, a hybrid approach is used to increase the overall efficiency of the network in WSNs with time synchronization which increases the throughput of the network. The efficiency in terms of network lifetime, residual energy, data packets, throughput of the network has been improved as shown in simulation results. The performance of WSN of the proposed scheme is compared to other classical routing scheme and proposed algorithm has proved its merit.

KEY WORDS: BASE STATION (BS), CLUSTER HEAD (CH), GATEWAY NODE (GN), ROUTING PROTOCOLS, WIRELESS SENSOR NETWORK (WSN).

INTRODUCTION

Presently there have been development in Micro-Electro-Mechanical Systems (MEMS) in tandem with major developments in digital signal processing (DSP) which has. led to growth of micro-sensors. Previously few industries use wired sensors, implementation provides deployment of sensor nodes more viable than before. Previously, there has been study regarding applications of WSN such as environmental monitoring, agricultural field, military surveillance and home automation (Pavithra et al., 2019).

Deep research on routing protocols efficiency was done considering the power constraints in WSNs and there was vast deployment of nodes on large-scale , to ensure reliable and real-time data transmission. Recently there has been exposure in the field of WSNs and their applications because they are easy to deploy and are of low cost, have flexibility. a WSN have distinctive set of resource curtailment like finite on-board battery power, limited processing ability and limited communication bandwidth. Since sensors are battery-powered, energy efficiency is of vital importance in WSNs. Algorithms are used to solve the problem of power constraint without altering the standard. Local collaboration among sensors, suppression, data compression, redundant data ,avoidance of direct transmission to far distant sensors are of the major factors that influence algorithm designers to device unique distributed, scalable and energy efficient solution for Wireless Sensor Networks.

.In common, the sensor nodes measure environmental conditions,. The sensor node extract some useful information by processing the raw sensor signals . The output of this processed signal is transmitted the through direct communication or multi-hop communication with access point across other sensor nodes. In some situations, repeaters (RPs) are used for multi-hops, to support sensors installed outside the radio range. One of the components of the WSN is the base stations which has more energy, computational ,communication resources. Forwarding of data from wireless sensor network on to a server is done by the BS which acts as a gateway. Energy is a limited resource of WSN , and it determines the lifetime of WSNs. The computation subsystem has less energy consumption when compared to the communication subsystem has. The energy required for transmitting one bit may consume as same as executing a few thousands instructions. Hence, communication must be traded for computation. In different environments, together with remote and hostile regions, where ad-hoc communications are a key element WSNs is to be deployed in large numbers .For this reason, algorithms and protocols should concentrate on the subsequent issues:

- Lifetime maximization: Sensor nodes should be energy efficient and consumption of energy of the device should be less due to the limited energy resources. The radio power supply when not in use should shut off to conserve power of the node
- fault tolerance and Robustness .
- Self-configuration.

Figure 1 shows the arrangement of WSN.

Figure 1: Schematic of WSN

Hierarchical routing algorithms are more efficient algorithm. In this, CHs are responsible for collecting the data and send it to GN and then the GN sends the information to the BS. If the members of the cluster are far, then there is a wastage of energy. To solve this problem, the Centrality approach is used. In this method,

the distance between one node to every other node is calculated using the Pythagoras theorem. The node which is equidistance to all other nodes along with high energy will be selected as the cluster head. This will reduce energy consumption. Error during communication can be reduced by using time synchronization in communication. This will reduce the packet collision during data transmission. In this paper, a new synchronization method considering a new broadcast sequence is used. This determines the order of the nodes that transmit timing information, this Eliminates collision. Section II presents a summary of the existing systems. The framework of proposed system is explained in Section III . The system is evaluated in Section IV based on simulation and implementation. Section V, addresses the future scope and demands. The paper concludes with section VI.

Existing Systems: A wireless sensor network consists of distributed sensor nodes. Sensor nodes include of sensing unit based on application, processing levels. In the existing system, many limitations or difficulties cause less efficiency of the system. It includes Energy limitations storage or environmental limitations, communication constraints, and other limitations. Transfer of information to nodes needs more energy. Hence sensor nodes are affected by energy limitations. Sensor nodes are small devices and hence their capacity to store the data is less. Nodes of the network may get affected by many environmental conditions physical obstacles, unpredictable errors, and also communication interferences.

Communication constraints are limitation of bandwidth, frequent routing changes, channel error rates, and also unreliable communication. LEACH (Salah et al., 2016) is the main protocol under hierarchical routing protocols which includes set-up phase, and steady-state phase. In the first phase some nodes are selected at random probability as cluster heads (CHs) and nodes are ordered into clusters. In the second phase, the data is transmitted to the BS. CHS has to lose more energy when compared to other normal nodes. A drawback of this system is cluster head consumes more energy and sensor nodes die faster.

In (Djenouri et al., 2012) they proposed a protocol that mainly deals with the technique of the cluster head (CH) selection that makes balanced energy consumption among the sensors, and which leads to enhancement of lifetime of the sensor network. In (Sang et al., 2012) for transmission of data to the Base Station majority of the energy of sensor nodes is used. Thus, there is a fast depletion of energy. Here agglomerative to limit the energy utilization of cluster heads a portable base station is utilized along with cluster approach. But the movement of the Base Station is not always feasible. In (Ruqiang et al., 2014) this it discusses up several technical challenges and many application possibilities that occur when the sensor networks interconnect several nodes when wide networks are established. These wireless sensor networks communicate using multiple -hop wireless communications systems.

To (Raghunandan et al., 2011) guarantee reliable multi-hop communication and to maintain the routes in the network routing protocols for WSN are used. It provides us with an idea of routing protocols for WSN and compares their relative strengths and limitations to provide better energy efficiency or increase wireless. In (Raghunandan et al., 2012) mainly concentrates on synchronization schemes where a multichip extension is used which use a final local estimates, without any forwarding of synchronization signals. In the existing system, there are many routing challenges and design issues. In designing routing protocols one should consider the uniform distribution of sensor nodes; otherwise, we should go for clustering. Production costs also should be less and the quality of service should be good.

Proposed System: In a hierarchical approach, some sensors are classified as lower energy nodes and some sensors are grouped as higher energy nodes. The higher energy nodes become cluster heads (CHs) and lower energy nodes become the normal sensors that perform the ensuing operation only.

CH is decided with a different probability (D Djenouri et al., 2012) by following equation.

$$T(n)= [p/(1-p)] \times (r \bmod p-1) \qquad (1)$$

Figure 2: Concept of proposed techniques

But the cluster heads deal with the collection of information from sensors, data aggregation of these data and transmission of processed data to the next level. Figure 2 shows the concept of the proposed technique. The main aims of this protocol are to reduce power consumption, data aggregation, and time synchronization. The main aim of the proposed technique is to increase energy efficiency, improving the time synchronization in communication between each node and the receiver, and enhance the network lifetime of sensor nodes. This paper is mainly motivated to overcome the problem of higher energy consumption of sensor nodes and extending the lifetime of wireless sensor networks to provide good transmission and data sense. Since Sensor nodes away

from base stations hence they consume more energy during transmission of data to a base station. Hence gateway nodes are introduced. Gateway nodes receive the data from sensor nodes and then send them to the base station. Time synchronization is an elementary part of any network-oriented organization and system. In our proposed system grouping of sensor nodes is done called clustering. The energy consumed by the transmitter for transmission of message to a distance is given by,

$$E_T(k,r) = \begin{cases} k\left(E_{TX} + E_{fs} * r^2\right) & \text{if } r < r_o, \\ k\left(E_{TX} + E_{mp} * r^4\right) & \text{if } r \geq r_o. \end{cases} \qquad (2)$$

The energy dissipated by a receiver to receive bit message is given by

$$E_R(k) = k * E_{RX}, \qquad (3)$$

Threshold distance is given by,

$$r_o = \sqrt{\frac{E'_{fs}}{E_{mp}}}. \qquad (4)$$

Clustering is done to construct the appropriate topology of the network. The clustering-based network reduces the cost of the routing algorithm and the flooding broadcast. In this cluster-based routing approach, sensor nodes that are in a particular radio range are grouped which forms a cluster. Each group or cluster has one cluster head (CH). This CH collects all the data from sensor nodes in a cluster and it performs data fusion and sends it to the nearby gateway node. CH is selected based on the centrality approach. In this method distance of nodes between each other is determined; the node which is equidistant from the other node and the node having more energy is selected as CH. After CH selection aggregation of data is performed by removing redundant data. Data aggregation is the process of removing redundant data during transmission. This will increases the lifetime of the network.

In the proposed system base station (BS) is fixed. The same algorithm can also be applied to a movable base station. During the data communication phase, the data packets are transmitted from CH to gateway node and from gateway node to BS. This process is called multi-hop communication. The proposed technique uses the receiver-to-receiver concept implemented by the reference broadcast Synchronization which minimizes the time-critical path when compared to the sender-to-receiver method.

Implementation: The proposed system is simulated using MATLAB software. We consider 100 sq.m area and 100 nodes are distributed randomly. MATLAB Simulation is done for 4500 rounds. The proposed algorithm is compared with DR-LEACH. Fig.3 Initial network topology for LEACH. The figure 3 shows deployment of sensor and clustering of the network area in DR-LEACH.

Figure 3: Initial network topology for LEACH

Figure 4: Network framework for Proposed Technique

Figure 5: Comparision of packets sent of Proposed System vs. DR-LEACH in multiples of 10^4

Figure 6: Comparison of number of dead sensors of Proposed System vs DR-LEACH

Figure 7: Comparison of Residual energy of results of Proposed System vs DR-LEACH

Figure 8: Comparison of Network Lifetime of Proposed System vs DR-LEACH

The connectivity between individuals to CH, CH to the gateway node, and GN to Base station are shown in figure 4 of proposed system. Figure 4 shows network framework of Proposed Technique. Figure 5 packets sent of Proposed System vs. DR-LEACH. Figure 6 shows the Comparison of number of dead sensors of Proposed System vs DR-LEACH. Figure 7 Comparison of Residual energy of results of Proposed System vs DR-LEACH. The implementation results in a hybrid model comprising of energy conservation, data aggregation, and time synchronization. The results in the graph show improvements in the lifetime of the network along with time synchronization. With respect to simulation results numbers of dead nodes are less on comparison with rounds of DR-LEACH protocol. This shows that the lifetime of network using proposed system is enhanced compared to DR-LEACH which leads to maximum transmission of data. Hence the WSN using the proposed technique works with more powerfully on comparison to the DR-LEACH.

CONCLUSION

Energy consumption by sensors is a major parameter for network lifetime in WSN. In hierarchical clustering methodology, the cluster head selection plays an important role. When the complete network is considered, then the optimization of energy consumption can be done by the changing of the cluster head based on priorities at the cluster level. In this proposed system, we establish a connection from sensor nodes to CH, CH to GN, and from the gateway node to the base station. The energy utilization can be decreased by properly designing the CH selection mechanism.

The selection of cluster head proposed in this paper is a new technique discussed concerning distance from the base station and centrality approach at the cluster level. Along with that time synchronization is achieved using reference broadcast synchronization. The implemented protocol shows an improvement in the lifetime network, and the life of nodes will be extended to the maximum period. Here we are comparing the proposed system with the DR-LEACH protocol and it proves that the proposed system is more efficient than DR-LEACH.

REFERENCES

Abhishek , Sumedha, (2014) ,Minimization of Avg Energy Consumption to Prolong Lifetime of WSN, IEEE Global Conference on WCN.

Akkaya, Younis,(2005) ,A Survey of Routing Protocols in WSN, Journal of Ad Hoc Network ,Elsevier, vol. 3, pp. 325-349.

Amit , R Singh, A Nandini(2016) WSN – Challenges and Possibilities, IJCA, volume 140-No.4

D Djenouri(2012), R4 Time Synchronization in WSN, IEEE signal processing letters, Vol. 19, No. 4.

Salah, A. Boulouz(2016) ,Energy Efficient Clustering based on LEACH, International Conference on Engineering and MIS.

Moumita, Paramita , Sarmistha (2014),Analysis of EE WSN Routing Schemes, International Conference of EAIT.

Pavithra, G.S. and Babu, N.V., 2019. Energy efficient hierarchical clustering using HACOPSO in wireless sensor networks. International Journal of Innovative Technology and Exploring Engineering, 8(12).

Raghunandan, Shobha Rani, Nanditha (2014), Hierarchical clustering based routing Technique for overall efficiency of WSN, In Proceeding of the IEEE 2017 International Conference on RTECIT.

Raghunandan, Shobha.Rani, Nanditha, et al.,(2017),Hierarchical agglomerative clustering based routing algorithm for overall efficiency of WSN, In Proceeding of IEEE ICICT -2017.

Raghunandan, Lakshmi(2011) ,Secure Routing protocols for WSN, International Conference on FCS proceedings.

Raghunandan, Sagar (2012) ,A Novel Approach to increase overall Efficiency in WSN ,2012 IEEE ICCEET proceedings, pp 699-703.

Raghunandan, Lakshmi.(2011) ,A Comparative Analysis of Routing Techniques for WSN, IEEE NCOIET-2011, IEEE Conference Proceedings.

Ruqiang, Sun,Yuning(2013) ,EAS Node Design with Its Application in WSN, IEEE Transactions on instrumentation and measurement, Vol. 62.

Sang, Nguyen (2012) ,Distance Based Thresholds for CH selection in WSN, IEEE Communications Letters ,Vol 16 , Issue 9 .

Ying, Chin , Ting (2008)" AER Protocol for mobile sink in WSN," IEEE Conference on Media Computing, conference proceeding, pp. 44-49.

Yunhe, Zhihua,Q Zhang(2016) "Efficient Load Balance Data Aggregation methods for WSN based on compressive network coding ", IEEE International Conference on EICT.

the PRGW. Additionally, two FSS layers, each one including a 3×3 FSS unit cell, working in the reflection mode, were fixed in the middle of a double-feed DRA by inserting them inside the rectangular cuts, which were etched in the two vertical walls on two sides of the antenna to tilt the antenna beam at the Ka band. **<AU: Please check whether the preceding edited sentence conveys the intended meaning.>** In-depth statistical and comparative analyses are given in the supplementary material (available at https://doi.org/10.1109/MAP.2021.3089981) for interested readers. References [101]–[106] are cited in the supplementary material.

FUTURE CHALLENGES

The MIMO technology uses multiple radiators at the transmitter and receiver ends to increase the data rates for a given BW and power requirement. To realize this, one can use antenna elements like a dipole, patch, DRA, and so on. It is argued that, once a single-element DRA is designed, the multiple elements can be arranged in the proper configuration to achieve diversity (spatial, polarization, and angle).

However, there are challenges that one can encounter in designing compact MIMO systems with DRAs having closely spaced elements or a single element with multiple ports for next-generation 5G communication systems and beyond. Major challenges in DRA-based MIMO antennas can be classified as follows:

- single-element compact CP MIMO DRAs
- multiband CP MIMO DRAs
- an isolation mechanism, such as the DGS, meander line, FSS, metamaterial, insertion of vias, and others **<AU: Please check whether the preceding edited bullet point conveys the intended meaning.>**
- feeding mechanisms to excite the desired modes in nonconventional shapes of DRAs in an MIMO configuration
- a detailed theoretical mode analysis in multielement, non-conventionally shaped DRA-based MIMO antennas
- a low-dielectric-constant material of the DRA element
- the effect of casing on a DRA-based MIMO antenna.

CONCLUSIONS

In this dissemination, a comprehensive review of recently published MIMO DRAs was presented. A detailed comparison of MIMO DRAs was done on the basis of various performance parameters and radiation characteristics. The aim of the review was the identification of major research gaps/challenges for future investigators.

The results of this study indicate that there is immense scope in the emerging domain of DRA-based MIMO antennas. Due to various advantages of DRA and MIMO technology, the challenges revealed by this work will prove a milestone in the future. The future challenges to be studied by researchers include fabricating and testing lower-dielectric-constant-material MIMO DRAs, analyzing the modes of unconventionally shaped MIMO DRAs, and the design of compact single DRA-based MIMO antennas with high isolation between multiple ports or hybrid feeds to generate the desired modes in unconventionally shaped DRA-based MIMO antennas. **<AU: Please**

check whether the preceding edited sentence conveys the intended meaning.>

AUTHOR INFORMATION

Sumer Singh Singhwal (sumersinghwal@gmail.com) is a postdoctoral researcher in the Department of Electronics and Telecommunications, Politecnico di Torino, Torino, 10129, Italy. He received his Ph.D. degree in electronics and communication engineering in 2020 from Uttarakhand Technical University Dehradun. His research interests include dielectric resonator antennas for different wireless applications. **<AU: Please note that bios have been trimmed in accordance with magazine style.>**

Ladislau Matekovits (ladislau.matekovits@polito.it) is an associate professor in Instituto di Elettronica e di Ingegneria dell'Informazione e delle Telecomunicazioni, National Research Council of Italy, Torino, 10129, Italy, and he obtained the full professor qualification (Italy) in 2017. He received his Ph.D. degree from Politecnico di Torino in 1995. He has published more than 375 papers (95 journal articles). He is a Senior Member of IEEE. **<AU: Please note that bios have been trimmed in accordance with magazine style.>**

Binod Kumar Kanaujia (bkkanaujia@ieee.org) is a professor in the School of Computational and Integrative Sciences, Jawaharlal Nehru University, New Delhi, 110067, India. He has successfully executed nine research projects sponsored by several agencies of the Government of India. He is a Senior Member of IEEE. **<AU: Please note that bios have been trimmed in accordance with magazine style.>**

Jugul Kishor (er.jugulkishor@gmail.com) is an associate professor in the Department of Electronics and Communication Engineering, JIMS-EMTC **<AU: Kindly spell out JIMS-EMTC.>** Greater Noida, India. **<AU: Please provide the postal code for the current affiliation.>** He received his Ph.D. degree from IIT-ISM **<AU: Kindly spell out IIT-ISM.>** Dhanbad, Jharkhand, India, in 2017. He has published more than 50 papers with various reputed international journals and conferences. **<AU: Please note that bios have been trimmed in accordance with magazine style.>**

Saeed Fakhte (fakhte@qut.ac.ir) is an assistant professor with the Department of Electrical and Computer Engineering, Qom University of Technology, Qom, 1519-37195, Iran. His research interests include various areas of microwave and antenna engineering, such as dielectric resonator and microstrip antennas. **<AU: Please note that bios have been trimmed in accordance with magazine style.>**

Amit Kumar (amit.kumarc210@gmail.com) is with the BMS **<AU: Kindly spell out BMS.>** Institute of Technology and Management, Bengaluru, Karnataka, 560064, India. He has published more than 15 research articles in reputed journals and completed

one project of Rs. 17 Lacs funded by MHRD, **<AU: Kindly spell out Rs. 17 Lacs and MHRD.>** Government of India.

REFERENCES

[1] A. S. Konanur, K. Gosalia, S. H. Krishnamurthy et al., "Increasing wireless channel capacity through MIMO systems employing co-located antennas," *IEEE Trans. Microw. Theory Techn.*, vol. 53, no. 6, pp. 1837–1844, 2005. doi: 10.1109/TMTT.2005.848105. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[2] B. Allen, W. Q. Malik, P. J. Smith et al., "Demystifying MIMO," *IET Commun. Eng.*, vol. 4, no. 6, pp. 38–42, 2006. doi: 10.1049/ce:20060607. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[3] P. F. Hu, Y. M. Pan, X. Y. Zhang, and B. J. Hu, "A compact quasi-isotropic dielectric resonator antenna with filtering response," *IEEE Trans. Antennas Propag.*, vol. 67, no. 2, pp. 1294–1299, Feb. 2019. doi: 10.1109/TAP.2018.2883611.

[4] Y. Liu, K. W. Leung, J. Ren, and Y. Sun, "Linearly and circularly polarized filtering dielectric resonator antennas," *IEEE Trans. Antennas Propag.*, vol. 67, no. 6, pp. 3629–3640, June 2019. doi: 10.1109/TAP.2019.2902670.

[5] N. Yang and K. W. Leung, "Size reduction of omnidirectional cylindrical dielectric resonator antenna using a magnetic aperture source," *IEEE Trans. Antennas Propag.*, vol. 68, no. 4, pp. 3248–3253, Apr. 2020. doi: 10.1109/TAP.2019.2944532.

[6] M. Yang, Y. Pan, Y. Sun, and K. Leung, "Wideband circularly polarized substrate-integrated embedded dielectric resonator antenna for millimeter-wave applications," *IEEE Trans. Antennas Propag.*, vol. 68, no. 2, pp. 1145–1150, Feb. 2020. doi: 10.1109/TAP.2019.2938629.

[7] W. Yang, W. Sun, H. Tang, and J. Chen, "Design of a circularly polarized dielectric resonator antenna with wide bandwidth and low axial ratio values," *IEEE Trans. Antennas Propag.*, vol. 67, no. 3, pp. 1963–1968, Mar. 2019. doi: 10.1109/TAP.2019.2891219.

[8] S. Fakhte, H. Oraizi, and L. Matekovits, "High gain rectangular dielectric resonator antenna using uniaxial material at fundamental mode," *IEEE Trans. Antennas Propag.*, vol. 65, no. 1, pp. 342–347, Jan. 2017. doi: 10.1109/TAP.2016.2627520.

[9] S. Fakhte, H. Oraizi, L. Matekovits, and G. G. Dassano, "Cylindrical anisotropic dielectric resonator antenna with improved gain," *IEEE Trans. Antennas Propag.*, vol. 65, no. 3, pp. 1404–1409, Mar. 2017. doi: 10.1109/TAP.2016.2647689.

[10] S. Liu, D. Yang, Y. Chen, S. Huang, and Y. Xiang, "Broadband dual circularly polarized dielectric resonator antenna for ambient electromagnetic energy harvesting," *IEEE Trans. Antennas Propag.*, vol. 68, no. 6, pp. 4961–4966, June 2020. doi: 10.1109/TAP.2020.2968768.

[11] S. S. Singhwal, B. K. Kanaujia, A. Singh, and J. Kishor, "Novel circularly polarized dielectric resonator antenna for microwave image sensing application," *Microw. Opt. Technol. Lett.*, vol. 61, no. 7, pp. 1821–1827, 2019. doi: 10.1002/mop.31830.

[12] Q. Wu, "Characteristic mode assisted design of dielectric resonator antennas with feedings," *IEEE Trans. Antennas Propag.*, vol. 67, no. 8, pp. 5294–5304, Aug. 2019. doi: 10.1109/TAP.2019.2916763.

[13] K. W. Leung and K. K. So, "Frequency-tunable designs of the linearly and circularly polarized dielectric resonator antennas using a parasitic slot," in *IEEE Trans. Antennas Propag.*, vol. 53, no. 1, pp. 572–578, Jan. 2005. doi: 10.1109/TAP.2004.838762.

[14] C. X. Hao, B. Li, K. W. Leung, and X. Q. Sheng, "Frequency tunable differentially fed rectangular dielectric resonator antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 884–887, 2011. doi: 10.1109/LAWP.2011.2165929. **<AU: Please provide the issue number or month.>**

[15] T. Chiu, O. Pajona, J. Kyllonen, L. Desclos, L. Huitema, and T. Monediere, "Active steering dielectric resonator antenna for automotive," in *Proc. Int. Workshop on Antenna Technol.: Small Antennas, Innov. Struct. Appl. (iWAT)*, Athens, Greece, 2017, pp. 343–346. doi: 10.1109/IWAT.2017.7915396.

[16] S. Guo, L. Wu, K. W. Leung, and J. Mao, "Active integrated dielectric resonator antenna-in-package design," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 11, pp. 2414–2418, Nov. 2019. doi: 10.1109/LAWP.2019.2937673.

[17] W. M. Abdel-Wahab, M. Abdallah, J. Anderson, Y. Wang, H. Al-Saedi, and S. Safavi-Naeini, "SIW-integrated parasitic DRA array: Analysis, design, and measurement," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 1, pp. 69–73, Jan. 2019. doi: 10.1109/LAWP.2018.2880926.

[18] R. Movahedinia, A. Sebak, M. R. Chaharmir, M. Ranjbar Nikkhah, and A. A. Kishk, "X-band circularly polarized electronically steerable parasitic array radiator of DRA," *IEEE Trans. Antennas Propag.*, vol. 66, no. 2, pp. 721–728, Feb. 2018. doi: 10.1109/TAP.2017.2780895.

[19] Z. Chen et al., "Millimeter-wave rectangular dielectric resonator antenna array with enlarged DRA dimensions, wideband capability, and high-gain performance," *IEEE Trans. Antennas Propag.*, vol. 68, no. 4, pp. 3271–3276, Apr. 2020. doi: 10.1109/TAP.2019.2950101. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[20] J. Kowalewski, J. Eisenbeis, A. Jauch, J. Mayer, M. Kretschmann, and T. Zwick, "A mmW broadband dual-polarized dielectric resonator antenna based on hybrid modes," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 7, pp. 1068–1072, July 2020. doi: 10.1109/LAWP.2020.2988516.

[21] Y. Al-Alem and A. A. Kishk, "Wideband millimeter-wave dielectric resonator antenna with gain enhancement," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 12, pp. 2711–2715, Dec. 2019. doi: 10.1109/LAWP.2019.2949947.

[22] E. Baldazzi et al., "A high-gain dielectric resonator antenna with plastic-based conical horn for millimeter-wave applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 19, no. 6, pp. 949–953, 2020. doi: 10.1109/LAWP.2020.2984565. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[23] C. Li and T. Chiu, "340-GHz low-cost and high-gain on-chip higher order mode dielectric resonator antenna for THz applications," *IEEE Trans. Terahertz Sci. Tech.*, vol. 7, no. 3, pp. 284–294, May 2017. doi: 10.1109/TTHZ.2017.2670234.

[24] C. Li and T. Chiu, "Single flip-chip packaged dielectric resonator antenna for CMOS terahertz antenna array gain enhancement," *IEEE Access*, vol. 7, pp. 7737–7746, 2019. doi: 10.1109/ACCESS.2018.2890678. **<AU: Please provide the issue number or month.>**

[25] C. Fumeaux et al., "Terahertz and optical dielectric resonator antennas: potential and challenges for efficient designs," in *Proc. 10th European Conf. Antennas Propag. (EuCAP)*, Davos, 2016. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the page range.>**

[26] S. Bidault, M. Mivelle, and N. Bonod, "Dielectric nanoantennas to manipulate solid-state light emission," *J. Appl. Phys.*, vol. 126, no. 9, p. 094104, 2019. doi: 10.1063/1.5108641.

[27] C. Zou, W. Withayachumnankul, M. Bhaskaran, S. Sriram, and C. Fumeaux, "Dielectric resonator nanoantennas: A review of the theoretical background, design examples, prospects, and challenges," *IEEE Antennas Propag. Mag.*, vol. 59, no. 6, pp. 30–42, Dec. 2017. doi: 10.1109/MAP.2017.2752638.

[28] K. Ishimiya, J. Langbacka, Z. Ying et al.," "A compact MIMO DRA antenna," in *Proc. IEEE Int. Workshop on Antenna Technol.: Small Antennas Novel Metamater.*, Chiba, Japan, 2008, pp. 286–289. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[29] "Rules & regulations for Title 47: Part 15, section-15.503," US Federal Communications Commission (FCC), Washington, D.C., Oct. 2003. [Online]. Available: http://www.fcc.gov/oet/info/rules **<AU: Please confirm the title and the organization location.>**

[30] J. B. Yan and J. T. Bernhard, "Implementation of a frequency agile MIMO dielectric resonator antenna," *IEEE Trans. Antennas Propag.*, vol. 61, no. 7, pp. 3434–3441, 2013. doi: 10.1109/TAP.2013.2255092.

[31] S. F. Roslan, M. R. Kamarudin, M. Khalily et al., "An MIMO rectangular dielectric resonator antenna for 4G applications," *IEEE Antennas Wirel. Propag. Lett.*, vol. 13, pp. 321–324, 2014. doi: 10.1109/LAWP.2014.2305696. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the issue number or month.>**

[32] J. C. Johnstone, S. K. Podilchak, M. Clénet et al., "A compact cylindrical dielectric resonator antenna for MIMO applications," in *Proc. IEEE Antennas Propag. Soc. Int. Symp. (APSURSI)*, Memphis, TN, 2014, pp. 1938–1939. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[33] S. F. Roslan, M. R. Kamarudin, M. Khalily et al., "An MIMO F-shaped dielectric resonator antenna for 4G applications," *Microwave Opt. Technol. Lett.*, vol. 57, no. 12, pp. 2931–2936, 2015. doi: 10.1002/mop.29473. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[34] M. S. Sharawi, "Current misuses and future prospects for printed multiple-input, multiple-output antenna systems," *IEEE Antennas Propag. Mag. [Wireless Corner]*, vol. 59, no. 2, pp. 162–170, 2017. doi: 10.1109/MAP.2017.2658346.

[35] A. Sharma and A. Biswas, "Wideband multiple-input-multiple-output dielectric resonator antenna," *IET Microw., Antennas Propag.*, vol. 11, no. 4, pp. 496–502, 2016. doi: 10.1049/iet-map.2016.0515.

[36] S. Aqeel, M. H. Jamaluddin, A. A. Khan et al., "A dual-band multiple input multiple output frequency agile antenna for GPSL1/Wi-Fi/WLAN2400/LTE applications," *Int. J. Antennas Propag.*, vol. 2016, 2016. doi: 10.1155/2016/9419183. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[37] M. Abedian, S. K. A. Rahim, C. Fumeaux et al., "Compact ultrawideband MIMO dielectric resonator antennas with WLAN band rejection," *IET Microw., Antennas Propag.*, vol. 11, no. 11, pp. 1524–1529, 2017. doi: 10.1049/iet-map.2016.0299. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[38] K. Trivedi and D. Pujara, "Mutual coupling reduction in wideband tree shaped fractal dielectric resonator antenna array using defected ground structure for MIMO applications," *Microw. Opt. Technol. Lett.*, vol. 59, no. 11, pp. 2735–2742, 2017. doi: 10.1002/mop.30810.

[39] A. Behloul, I. Messaoudene, T. A. Denidni, and A. Benghalia, "Three-port triangular dielectric resonator antenna with switching beam forming operation," *Microw. Opt. Technol. Lett.*, vol. 59, no. 4, pp. 955–958, Apr. 2017. doi: 10.1002/mop.30442.

[40] A. Sharma, A. Sarkar, and A. Biswas, "A-shaped wideband dielectric resonator antenna for wireless communication systems and its MIMO implementation," *Int. J. RF Microw. Comput.-Aided Eng.*, 2018. **<AU: Please provide the volume, issue number or month, and page range.>**

[41] T. L. Chiu, L. Huitema, O. Pajona, and T. Monediere, "Compact and multiband MIMO dielectric resonator antenna for automotive LTE communications," *Int. J. Antennas Propag.*, vol. 2018, 2018. doi: 10.1155/2018/8231081.

[42] G. Das, A. Sharma, and R. K. Gangwar, "Wideband self-complementary hybrid ring dielectric resonator antenna for MIMO applications," *IET Microw., Antennas Propag.*, vol. 12, 2018. **<AU: Please provide the volume, issue number or month, and page range.>**

[43] G. Das, A. Sharma, and R. K. Gangwar, "Dielectric resonator-based two-element MIMO antenna system with dual band characteristics," *IET Microw., Antennas Propag.*, vol. 12, no. 5, 2018. **<AU: Please provide the page range.>**

[44] N. K. Sahu, G. Das, and R. K. Gangwar, "Dual polarized triple-band dielectric resonator based hybrid MIMO antenna for WLAN/WiMAX applications," *Microw. Opt. Technol. Lett.*, vol. 60, no. 4, pp. 1033–1041, 2018. doi: 10.1002/mop.31102.

[45] A. Sharma, G. Das, and R. K. Gangwar, "Design and analysis of tri-band dual-port dielectric resonator-based hybrid antenna for WLAN/WiMAX applications," *IET Microw., Antennas Propag.*, vol. 12, no. 6, pp. 986–992, 2018. doi: 10.1049/iet-map.2017.0822.

[46] T. Kumari, G. Das, and R. K. Gangwar, "Wideband multi-segment dielectric resonator antenna for MIMO applications," in *Proc. Int. Conf. Microw. Photonics (ICMAP 2018)*, 2018. **<AU: Please provide the page range.>**

[47] T. Kumari, G. Das, A. Sharma, and R. K. Gangwar, "Design approach for dual element hybrid MIMO antenna arrangement for wideband applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2018. **<AU: Please provide the page range.>**

[48] J. Iqbal, U. Illahi, M. I. Sulaiman et al., "Mutual coupling reduction using hybrid technique in wideband circularly polarized MIMO antenna for WiMAX applications," *IEEE Access*, vol. 7, pp. 40,951–40,958, 2019. doi: 10.1109/ACCESS.2019.2908001. **<AU: Please provide the issue number or month.>**

[49] H. N. Chen, J. Song, and J. Park, "A compact circularly polarized MIMO dielectric resonator antenna over electromagnetic band-gap surface for 5G applications," *IEEE Access*, vol. 7, pp. 140,889–140,898, Sept. 2019. doi: 10.1109/ACCESS.2019.2943880.

[50] S. S. Singhwal, B. K. Kanaujia, A. Singh, and J. Kishor, "Dual-port MIMO dielectric resonator antenna for WLAN applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[51] G. Bharti, D. Kumar, A. K. Gautam, and A. Sharma, "Two-port ring-shaped dielectric resonator-based diversity radiator with dual-band and dual-polarized features," *Microw. Opt. Technol. Lett*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[52] A. K. Biswas and U. Chakraborty, "Complementary meander-line-inspired dielectric resonator multiple-input-multiple-output antenna for dual-band applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[53] T. Kumari, G. Das, A. Sharma, and R. K. Gangwar, "Dielectric resonator based two-port dual band antenna for MIMO applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[54] S. S. Singhwal, B. K. Kanaujia, A. Singh, J. Kishor, and L. Matekovits, "Multiple input multiple output dielectric resonator antenna with circular polarized

adaptability for 5G applications," *J. Electromagn. Waves Appl.*, vol. 34, no. 9, pp. 1180–1194, Feb. 2020. doi: 10.1080/09205071.2020.1730984.

[55] G. Varshney, R. Singh, V. S. Pandey, and R. S. Yaduvanshi, "Circularly polarized two-port MIMO dielectric resonator antenna," *Progr. Electromagn. Res. M*, vol. 91, pp. 19–28, 2020. doi: 10.2528/PIERM20011003. **<AU: Please provide the issue number or month.>**

[56] M. Belazzoug, I. Messaoudene, S. Aidel, T. A. Denidni, and A. A. Kishk, "Quarter cylindrical dielectric resonator antenna with circular polarisation for wideband MIMO systems," *Electron. Lett.*, vol. 56, no. 4, pp. 174–176, 2020. doi: 10.1049/el.2019.3138.

[57] L. Zou and C. Fumeaux, "A cross-shaped dielectric resonator antenna for multifunction and polarization diversity applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 10, pp. 742–745, 2011. doi: 10.1109/LAWP.2011.2162479. **<AU: Please provide the issue number or month.>**

[58] N. K. Sahu, R. K. Gangwar, and P. Kumari, "Dielectric resonator based circularly polarized MIMO antenna for WLAN applications," in *Proc. Int. Conf. Microw. Photonics (ICMAP 2018)*, 2018. **<AU: Please provide the page range.>**

[59] N. K. Sahu, G. Das, and R. K. Gangwar, "L-shaped dielectric resonator based circularly polarized multi-input-multi-output (MIMO) antenna for wireless local area network (WLAN) applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2018. **<AU: Please provide the volume, issue number or month, and page range.>**

[60] G. Das, A. Sharma, and R. K. Gangwar, "Dielectric resonator based circularly polarized MIMO antenna with polarization diversity," *Microwave Opt. Technol. Lett.*, vol. 60, no. 3, pp. 685–693, 2018. doi: 10.1002/mop.31033.

[61] S. S. Singhwal, B. K. Kanaujia, A. Singh, J. Kishor, and L. Matekovits, "Dual-band circularly polarized MIMO DRA for sub-6 GHz applications," *Int. J. RF Microw. Comput.-Aided Eng.*, July 2020. **<AU: Please provide the volume and page range.>**

[62] J. B. Yan and J. T. Bernhard, "Design of a MIMO dielectric resonator antenna for LTE femtocell base stations," *IEEE Trans. Antennas Propag.*, vol. 60, no. 2, pp. 438–444, 2012. doi: 10.1109/TAP.2011.2174021.

[63] L. Z. Thamae and Z. Wu, "Dielectric resonator based multiple-input-multiple-output antennas and channel characteristic analysis," *IET Microw. Antennas Propag.*, vol. 6, no. 9, pp. 1084–1089, 2012. doi: 10.1049/iet-map.2012.0024.

[64] L. Zou, D. Abbott, and C. Fumeaux, "Omnidirectional cylindrical dielectric resonator antenna with dual polarization," *IEEE Antennas Wireless Propag. Lett.*, vol. 11, pp. 515–518, 2012. doi: 10.1109/LAWP.2012.2199277. **<AU: Please provide the issue number or month.>**

[65] Y. X. Sun and K. W. Leung, "Dual-band and wideband dual-polarized cylindrical dielectric resonator antennas," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, 2013. **<AU: Please provide the issue number or month and page range.>**

[66] J. Nasir, M. H. Jamaluddin, M. Khalily, M. R. Kamaruddin, I. Ullah, and R. Selvaraju, "A reduced size dual port MIMO DRA with high isolation for 4G applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2014. **<AU: Please provide the volume, issue number or month, and page range.>**

[67] X. S. Fang, K. W. Leung, and K. M. Luk, "Theory and experiment of three-port polarization-diversity cylindrical dielectric resonator antenna," *IEEE Trans. Antennas Propag.*, vol. 62, no. 10, pp. 4945–4951, 2014. doi: 10.1109/TAP.2014.2341698.

[68] N. M. Nor and M. H. Jamaluddin, "A dual band MIMO dielectric resonator antenna for WLAN application," *Jurnal Teknologi (Sci. Eng.)*, pp. 1–4, 2015. **<AU: Please provide the volume and issue number or month.>**

[69] J. Nasir, M. H. Jamaluddin, M. Khalily et al., "Design of an MIMO dielectric resonator antenna for 4G applications," *Wireless Pers. Commun.*, vol. 88, no. 3, pp. 525–536, 2016. doi: 10.1007/s11277-016-3174-3. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[70] A. A. Khan, M. H. Jamaluddin, J. Nasir et al., "Design of a dual-band MIMO dielectric resonator antenna with pattern diversity for WiMAX and WLAN applications," *Progress Electromagn. Res. M*, vol. 50, pp. 65–73, 2016. doi: 10.2528/PIERM16070102. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the issue number or month.>**

[71] A. A. Khan, M. H. Jamaluddin, S. Aqeel et al., "Dual-band MIMO dielectric resonator antenna for WiMAX/WLAN applications," *IET Microw., Antennas Propag.*, vol. 11, no. 1, 2017. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the page range.>**

[72] A. A. Khan, R. Khan, S. Aqeel et al., "Dual-band mimo rectangular dielectric resonator antenna with high port isolation for LTE applications," *Microw. Opt. Technol. Lett.*, vol. 59, no. 1, 2017. **<AU: Please provide the names of

the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the page range.>

[73] A. A. Khan, R. Khan, S. Aqeel et al., "Design of a dual-band MIMO dielectric resonator antenna with high port isolation for WiMAX and WLAN applications," *Int. J. RF Microw. Comput.-Aided Eng.*, 2017. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the page range.>**

[74] A. Abdalrazik, A. S. Abd El-Hameed, and A. B. Abdel-Rahman, "A three-port MIMO dielectric resonator antenna using decoupled modes," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, 2017. **<AU: Please provide the issue number or month and page range.>**

[75] H. Tang, J. X. Chen, W. W. Yang, L. H. Zhou, and W. Li, "Differential dual-band dual-polarized dielectric resonator antenna," *IEEE Trans. Antennas Propag.*, vol. 65, no. 2, pp. 855–860, 2017. doi: 10.1109/TAP.2016.2630661.

[76] G. Das, A. Sharma, and R. K. Gangwar, "Dual port aperture coupled MIMO cylindrical dielectric resonator antenna with high isolation for WiMAX application," *Int. J. RF Microw. Comput.-Aided Eng.*, 2017. **<AU: Please provide the volume, issue number or month, and page range.>**

[77] J. Nasir, M. H. Jamaluddin, A. A. Khan et al., "Throughput measurement of a dual-band MIMO rectangular dielectric resonator antenna for LTE applications," *MDPI Sensors*, vol. 17, no. 12, 2017. doi: 10.3390/s17010148. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the page range.>**

[78] A. Sharma, A. Sarkar, A. Biswas, and M. J. Akhtar, "Equilateral triangular dielectric resonator based co-radiator MIMO antennas with dual polarization," *IET Microw., Antennas Propag.*, vol. 12, no. 14, pp. 2161–2166, 2018. doi: 10.1049/iet-map.2018.5035.

[79] Z. Y. Zhang and K. L. Wu, "A wideband dual-polarized dielectric magneto-electric dipole antenna," *IEEE Trans. Antennas Propag.*, vol. 66, no. 10, pp. 5590–5595, 2018. doi: 10.1109/TAP.2018.2859914.

[80] H. Tang, C. Tong, and J. X. Chen, "Differential dual-polarized filtering dielectric resonator antenna," *IEEE Trans. Antennas Propag.*, vol. 66, no. 8, pp. 4298–4302, 2018. doi: 10.1109/TAP.2018.2836449.

[81] M. S. Sharawi, S. K. Podilchak, M. U. Khan, and Y. M. Antar, "Dual-frequency DRA-based MIMO antenna system for wireless access points," *IET Microw. Antennas Propag.*, vol. 11, no. 8, pp. 1174–1182, 2017. doi: 10.1049/iet-map.2016.0671.

[82] A. Sharma, A. Sarkar, A. Biswas, and M. J. Akhtar, "Dual-band multiple-input multiple-output antenna based on half split cylindrical dielectric resonator," *J. Electromag. Waves Appl.*, 2018. **<AU: Please provide the volume, issue number or month, and page range.>**

[83] I. K. C. Lin, M. H. Jamaluddin, A. Awang et al., "A triple band hybrid MIMO rectangular dielectric resonator antenna for LTE applications," *IEEE Access*, vol. 7, pp. 122,900–122,913, 2019. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors. Also, please provide the issue number or month.>**

[84] G. Das, A. Sharma, R. K. Gangwar, and M. S. Sharawi, "Performance improvement of multi-band MIMO dielectric resonator antenna system with a partially reflecting surface," *IEEE Antennas Wireless Propag. Lett.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[85] A. K. Dwivedi, A. Sharma, A. K. Singh, and V. Singh, "Quad-port ring dielectric resonator-based MIMO radiator with polarization and space diversity," *Microw. Opt. Technol. Lett.*, 2020. **<AU: Please provide the volume, issue number or month, and page range.>**

[86] M. T. Hussain, O. Hammi, M. S. Sharawi et al., "A dielectric resonator based millimeter-wave MIMO antenna array for hand-held devices," in *Proc. IEEE Int. Symp. Antennas Propag. USNC/URSI Nat. Radio Sci. Meeting*, Vancouver, BC, 2015, pp. 3–4. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[87] M. T. Hussain, M. S. Sharawi, S. K. Podilchack et al., "Closely packed millimeter-wave MIMO antenna arrays with dielectric resonator elements," in *Proc. 10th European Conf. Antennas Propag. (EuCAP)*, Davos, 2016, pp. 1–4. **<AU: Please provide the names of the first six authors. Use "first author's name et al." if there are seven or more authors.>**

[88] A. Dadgarpour, B. Zarghooni, B. S. Virdee, T. A. Denidni, and A. A. Kishk, "Mutual coupling reduction in dielectric resonator antennas using metasurface shield for 60 GHz MIMO systems," *IEEE Antennas Wireless Propag. Lett.*,

vol. 16, pp. 477–480, 2016. doi: 10.1109/LAWP.2016.2585127. **<AU: Please provide the issue number or month.>**

[89] M. S. Sharawi, S. K. Podilchak, M. T. Hussain, and Y. M. M. Antar, "Dielectric resonator-based MIMO antenna system enabling millimetre-wave mobile devices," *IET Microw., Antennas Propag.*, vol. 11, no. 2, pp. 287–293, 2017. doi: 10.1049/iet-map.2016.0457.

[90] M. Farahani, J. Pourahmadazar, M. Akbari, M. Nedil, A. R. Sebak, and T. A. Denidni, "Mutual coupling reduction in millimeter-wave MIMO antenna array using a metamaterial polarization-rotator wall," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2324–2327, 2017. doi: 10.1109/LAWP.2017.2717404. **<AU: Please provide the issue number or month.>**

[91] J. Li, Q. Zeng, R. Liu, and T. A. Denidni, "Beam-tilting antenna with negative refractive index metamaterial loading," *IEEE Antennas Wireless Propag. Lett.*, vol. 16, pp. 2030–2033, 2017. doi: 10.1109/LAWP.2017.2694400. **<AU: Please provide the issue number or month.>**

[92] G. Das, A. Sharma, R. K. Gangwar, and M. S. Sharawi, "Triple-port, two-mode based two element cylindrical dielectric resonator antenna for MIMO applications," *Microw. Opt. Technol. Lett.*, vol. 60, no. 6, pp. 1566–1573, 2018. doi: 10.1002/mop.31202.

[93] G. Das, A. Sharma, R. K. Gangwar, and M. S. Sharawi, "Compact back-to-back DRA-based four-port MIMO antenna system with bi-directional diversity," *Electron. Lett.*, vol. 54, no. 14, pp. 884–886, 2018. doi: 10.1049/el.2018.0959.

[94] G. Das, N. K. Sahu, A. Sharma, R. K. Gangwar, and M. S. Sharawi, "Dielectric resonator-based four-element eight-port MIMO antenna with multi-directional pattern diversity," *IET Microw., Antennas Propag.*, vol. 13, no. 1, pp. 16–22, 2019. doi: 10.1049/iet-map.2018.5081.

[95] Y. M. Pan, X. Quin, Y. X. Sun, and S. Y. Zheng, "A simple decoupling method for 5G millimeterwave MIMO dielectric resonator antennas," *IEEE Trans. Antennas Propag.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[96] G. Das, N. K. Sahu, A. Sharma, R. K. Gangwar, and M. S. Sharawi, "FSS based spatially decoupled back to back four port MIMO DRA with multi-directional pattern diversity," *IEEE Antennas Wireless Propag. Lett.*, 2019. **<AU: Please provide the volume, issue number or month, and page range.>**

[97] Y. Zhang, J. Y. Deng, M. J. Li, D. Sun, and L. X. Guo, "A MIMO dielectric resonator antenna with improved isolation for 5G mm-Wave applications," *IEEE Antennas Wireless Propag. Lett.*, vol. 18, no. 4, pp. 747–751, Apr. 2019. doi: 10.1109/LAWP.2019.2901961.

[98] M. Belazzoug, I. Messaoudene, S. Aidel, M. Nedil, and A. A. Kishk, "Ultra-compact 4-port DR antenna for multi-input multi-output standards," *Int. J. RF Microw. Comput.-Aided Eng.*, Jan. 2020. **<AU: Please provide the volume and page range.>**

[99] G. Das, N. K. Sahu, A. Sharma, and R. K. Gangwar, "Dielectric resonator based multiport antenna system with multi-diversity and built-in decoupling mechanism," *Int. J. Electron. Commun. (AEÜ)*, 2020. **<AU: Please provide the volume, issue number or month, and page range.>**

[100] M. B. Kakhki, A. Dadgarpour, A. R. Sebak, and T. A. Denidni, "Twenty-eight-gigahertz beam-switching ridge gap dielectric resonator antenna based on FSS for 5G applications," *IET Microw. Antennas Propag.*, vol. 14, no. 5, pp. 397–401, 2020. doi: 10.1049/iet-map.2019.0591.

[101] M. S. Sharawi, "Printed multi-band MIMO antenna systems and their performance metrics," *IEEE Antennas Propag. Mag.*, vol. 55, no. 5, pp. 218–232, 2013. doi: 10.1109/MAP.2013.6735522.

[102] K. M. Luk and K. W. Leung, *Dielectric Resonator Antennas*. Research Studies Press Ltd., 2003. **<AU: Please provide the publisher location.>**

[103] A. Petosa, *Dielectric Resonator Antenna Handbook*. Norwood, MA: Artech House, 2007.

[104] R. K. Mongia and A. Ittipiboon, "Theoretical and experimental investigations on rectangular dielectric resonator antennas," *IEEE Trans. Antenna Propag.*, vol. 45, no. 9, pp. 1348–1356, 1997. doi: 10.1109/8.623123.

[105] Y. Hwang, Y. P. Zhang, and K. M. Luk, "Gain-enhanced miniaturised rectangular dielectric resonator antenna," *Electron. Lett.*, vol. 33, no. 5, pp. 350–352, 1997. doi: 10.1049/el:19970228.

[106] Y. M. Pan, K. W. Leung, and K. M. Luk, "Design of the millimeter-wave rectangular dielectric resonator antenna using a higher-order mode," *IEEE Trans. Antennas Propag.*, vol. 59, no. 8, 2011. **<AU: Please provide the page range.>**

A comprehensive review of all of the latest trends of MIMO DRAs is presented along with seven research gaps/ challenges that need to be addressed in the future.

Multiband antennas have always attracted researchers due to their diverse commercial applications in different bands simultaneously.

DRAs are good candidates for various applications across the spectrum, from the microwave to optical frequency bands.

The MIMO technology uses multiple radiators at the transmitter and receiver ends to increase the data rates for a given BW and power requirement.

CP antennas are preferred over their linearly polarized counterparts because of advantages such as the ease of reception and higher reliability of the signal.

The results of this study indicate that there is immense scope in the emerging domain of DRA-based MIMO antennas.

Compactness is one of the major attractions in multiport MIMO antennas.

Machine Learning Algorithms for Detection: A Survey and Classification

¹Deepa N. Reddy, ²Priyanka R, ³Sanjana S, ⁴Santrupti. M. Bagali, ⁵Sara Sadiya

¹Department of Electronics and Communication, BMS Institute of Technology & Management, Bengaluru, India.
Email: reddydeepa2680@gmail.com
²Department of Electronics and Communication, BMS Institute of Technology & Management, Bengaluru, India.
Email: priyankaraju9480@gmail.com
³Department of Electronics and Communication, BMS Institute of Technology & Management, Bengaluru, India.
Email: sanjanadathathri123@gmail.com
⁴Department of Electronics and Communication, BMS Institute of Technology & Management, Bengaluru, India.
Email: santruptibagali@gmail.com
⁵Department of Electronics and Communication, BMS Institute of Technology & Management, Bengaluru, India.
Email: mnsara2288@gmail.com

Abstract: There is an enormous amount of data being dealt with by the medical field on a daily basis. Using a conventional method for handling data can affect the accuracy of the results. Early recognition of the disease is crucial for the analysis of patient medicines and specialists. The objective of this paper is to provide a comprehensive review of the techniques used in disease detection. Machine learning algorithms can be used to find out facts in medical research, particularly disease prediction. Machine learning algorithms such as Support vector machine [SVM], Decision trees, Bayes classifiers, K-Nearest Neighbours [KNN] Ensemble classifier techniques, etc. are used to determine different ailments. The use of machine learning algorithms can lead to fast and high accuracy prediction of diseases. This research paper analyses how machine learning techniques and algorithms are used to predict different diseases and their types. This paper provides an extensive survey of the machine learning techniques used for the prediction of chronic kidney disease, liver disease, haematological diseases, Alzheimer's disease, and urinary tract infections.

Index Terms: Machine Learning Algorithm, Disease, Neural Networks, Decision tree.

1. Introduction

Machine learning is a subset of artificial intelligence that contains algorithms or methods, for naturally creating models from data. Machine Learning is the technique of making computers learn and act like human beings by feeding data sets and information without being specifically programmed[1][2]. A machine learning system learns from experience, unlike a machine that performs a task by following clear-cut rules[3][4]. A rule-based system will perform a task the same way every time, whereas the performance of a machine learning system can be made better through training and testing by exposing the algorithm to more data. The Machine Learning model can be broadly classified into three categories.

2. Machine Learning:Classification
2.1. Supervised Learning Model

Supervised learning as the name suggests works under supervision, that is the machine predicts after being trained by the data that is labelled. Data for which the target answer is already known is called a labelled data. [5] The labelled data is fed to the machine which analyses and learns the relation of these images with its labels, based on its features. Now when a new image is fed to the machine without any label, with the aid of only the past data set, the machine is able to predict accurately and give the output. Example algorithms include: The Back Propagation Neural Network and Logistic Regression[6][7].
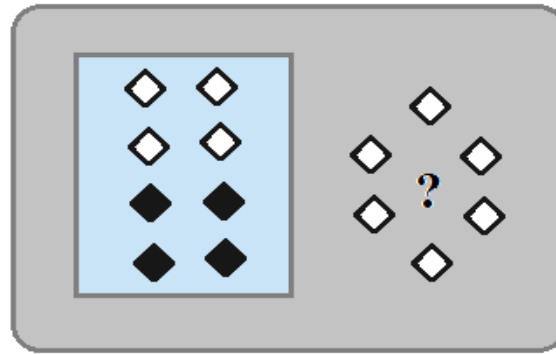
Fig.1. Supervised Learning Algorithm

2.2. Unsupervised Learning Model

Unsupervised machine learning Model lacks supervision, by not training the machine nor allowing it to act on the data which is not labeled. Therefore, machine tires to link the patterns and give the response. The machine recognizes the patterns from given set of data and clusters them based on their similarities, patterns, etc. [8] Unsupervised learning can be additionally grouped into association and clustering. Example problems include association role learning and dimensionality reduction clustering. Example algorithms are: the Apriori algorithm and K-Means [9].Unsupervised Learning models are extensively used over real life data sets[10].
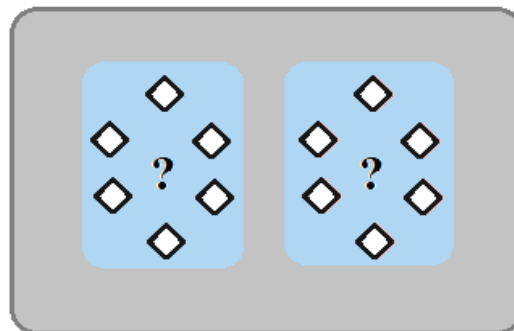
Fig.2. Unsupervised Learning Algorithm

2.3. Semi-supervised Learning Model

The amalgamation of supervised learning and unsupervised learning is referred to as Semi Supervised learning. This learning has a combination of labeled and unlabeled data set. Manually labeling all the data set is not practically possible, but can label some portions of the data and use that portion to train our model. The labeled data can be used as a training set for our model. We use our model to predict on the unlabeled part of the data set and label them. This operation of labelling the unlabeled data in tandem with the output that was forecasted by our neural network is called pseudo labelling. After designating the unlabeled data, then we train our model with the full data set [11][12].
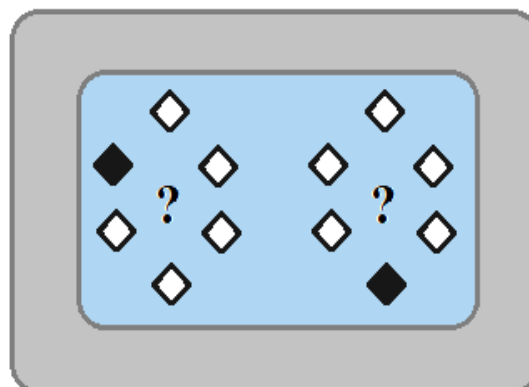
Fig.3. Semi-supervised Learning Algorithm

Table 1. Summary of the Classification Models [13][14]

TYPES OF LEARNING		
Supervised Machine learning	**Unsupervised Machine learning**	**Semi-Supervised Machine learning**
Models/Method 1. Linear Regression 2. Support Vector Machine (SVM) 3. Nearest Neighbour 4. Random Forest 5. Naïve Bayes 6. Decision Trees	Models/Method 1.Heirarchical clustering 2.Principal Component Analysis 3.Independent Component Analysis 4.K Nearest Neighbours	Models/Method 1.Continuity Assumption 2.Manifold Assumption 3.Cluster Assumption

3. Machine Learning Techniques for Disease Predictionbased Self-adaptation with Reusability
3.1. Liver Disease

Liver disease (LD) is one of those uncertain diseases which are difficult to diagnose even though the symptoms are seen at an early stage. The reason it is difficult to diagnose is that the symptoms are not prominent in the early stage as the liver is capable of functioning at a partially damaged state. Early diagnosis can be life-saving. Even though the diagnosis of this disease at an early stage is a challenge to the medical industry, it can still be detected. Diagnoses at an early stage can increase the patient's life span substantially.

The most effective ML algorithms Back Propagation, Support Vector Machine (SVM), Naïve Bayes, Random Forest, and K- Nearest Neighbors (KNN) are used in the study and analysis of this disease [15].

The outcome of the analysis is as follows: Back-propagation algorithms give the accuracy of 73.2%[15], SVM gives an accuracy of 71% [15], Random forest provides an accuracy of 74%[16], KNN gives an accuracy of 62%[16], Naïve Bayes gives an accuracy of 95.1% [17].

3.2. Chronic Kidney Disease

Kidney is one of the vital organs in human body. They perform major functions like excretion, filtration of blood and osmoregulation. In other words, we can say that it helps in removing all the unnecessary and toxic material from the body. Every year in India, around a million cases of renal failure are diagnosed. Chronic kidney disease (CKD) is also called as renal failure. CKD is a slow and dangerous disease that leads to periodic loss of kidney function with the duration of time. It will develop to permanent kidney failure. The following Symptoms are seen if it's not diagnosed and cured at early stage: week bones, Blood Pressure, anemia, decreased immune response, electrolytes accumulation, poor nutrition health and nerve damage, and built-up wastes in blood and body. The symptoms of CKD develop slowly and are unpredictable and aren't confined and specific to only CKD. Sometimes the symptoms are not even observed with some patients. Machine learning can be used here in order to predict whether or not the person has CKD. This is done by using the old records of CKD affected patients in order to train the predicting model of machine learning algorithm. The stage of chronic kidney disease can be analyzed using the technique called Glomerular Filtration Rate (GFR). This is also used to find the level of kidney function. It majorly uses the patient's blood creatinine for its calculation.

By using data of CKD patients with machine learning algorithms such as Support Vector Machine, Decision Tree, Random Forest a model that provides maximum accuracy for predicting CKD can be built.

From the outcome of the algorithms, it is noticed that, SVM provides an accuracy of 96.75%[18], decision tree algorithm gives the accuracy of 91.75% [18], Random Forest gives 99.16% [19].

3.3. Haematology Diagnosis

Blood affects human life in many different ways. It circulates throughout the body like a postman and visits all organs. Blood should reflect the growth in change. Different parameter values in the blood analysis tests can be used to detect this change.

The pathological conditions that affect blood producing organs or the blood are termed as haematological diseases(HD). This group includes a variety of blood cancers, different types of anemia such as severe aplastic

anemia (SAA), thalassemia, iron deficiency anemia, sickle cell disease, and hemorrhagic conditions. These also include idiopathic thrombocytopenic purpura, congenital neutropenia etc.

Several parameters such as age, genders, symptoms and other health conditions are considered by the doctors to choose the specific test for detecting that disease. For successful treatment of these diseases, quick and accurate medical diagnosis is very important. In the survey, using laboratory blood test results and machine learning algorithms two models were built to predict the hematological diseases. The first model used a dataset containing different blood test parameters to detect the diseases and in the second model a reduced data set was used whose parameters were measured once the patient is admitted [20].

When the top five most likely haematological diseases were considered, the accuracy of the first and second model were 88% and 86% respectively. For determining the most likely disease the accuracy was 59% and 57% respectively. On conducting a clinical test, the accuracies of both the predictive models were very much similar to the results of haematology specialists. This was the first study which concluded that we can detect a haematological disease successfully with blood test samples alone, by using machine learning algorithms [20]. Usually, the data set used in haematological surveys contained a set of 50-60 parameters, a few of them were WBC (white blood cell concentration), PLT (platelet concentration), MCV (mean corpuscular volume), HCT (haematocrit), LYMPH % (lymphocyte concentration), NEUT% (percentage of neutrophils), MONO% (percentage of monocytes), IMI# (immature WBC concentration), PLT-X (average platelet concentration) [21]. From these studies we get to know that the results of the blood test contain a lot more information that is generally not recognized by the physicians. This remarkable result opens up exceptional possibilities in the field of medical diagnosis [22].

Among all the machine learning algorithms the ones used in this study include used:

Support Vector Machine (Linear and RBF)- the scikit implementation of SVC (Support Vector Classifier) is used; Naïve Bayesian Classifier-the experiment above used the scikit lean implementation, Gaussian NB; Decision Tree; Random Forest. On comparing different machine learning algorithms that were applied to a large number of data sets, it was observed that overall Random Forest was the best algorithm to be used.

When the results of different classifiers were examined, the accuracies of these classifiers ranged between 71.2% and 98.16%. The accuracies were found to be NaiveBayes-81.60%, Bayesian network-92.86%, Multilayer Perception-91.8-%, Decision Tree-97.00%, SVM-71.20%, Random Forests-97.12% [22].

3.4. Alzheimer's Disease

Alzheimer's Disease (AD) is a neurodegenerative disorder. The cause of this is uncertain and it is mostly seen in aged people. It is one of the most common cause that leads to dementia. Selective memory impairment is seen as the early symptom of this. As of now, only treatments to amend some of the symptoms are available and no cure is available. The most commonly used model and their analysis and accuracy is as follows:

The most frequently applied algorithm is Support Vector Machine (SVM) and the second most frequently used is the Naïve Bayes algorithm. From the results it is observed that the generalized linear model surpasses the other classifiers with an accuracy of 88.24% during the test period. Accuracies seen for Naive Bayes algorithms and deep learning are 74.65% and 78.32% respectively. Accuracies for KNN (K- Nearest Neighbors) and Decision Tree are 43.26 and 74.22 respectively[23].

96 out of the total of 1454 showed the instruments psychometric properties. 89 papers explain the paper and pencil test. From the studies it is observed that, the Montreal Cognitive Assessment showed an effective screening test for memory clinic testing [24].

3.5. Urinary Tract Infections

A collective term that describes any infection or abnormality affecting all or any part of the urinary tract namely kidneys, urinary bladder, urethra and the ureters is known as UTI or the Urinary Tract Infection. The urinary system of human beings can be split into two sections. The upper tract consists of the kidneys and the ureteric (ureter). The lower tract consists of the vesica urinaria (urinary bladder) and the urethra.

In the local primary care, Cystitis (commonly known as Urinary Tract Infection or UTI) is considered as one of the most common bacterial infections, affecting approximately 150 million people every year worldwide [25]. It is seen in older adults much more than the younger ones. It is very crucial to detect Urinary Tract Infections in early stages especially in older adults as delayed treatments might lead to further complications and can be catastrophic. An experiment was conducted whose aim was to validate, train and compare different models to detect Urinary Tract Infection's using a validation dataset.

Different models were developed in this experiment to predict the Urinary Tract Infection using seven machine learning algorithms namely- Support Vector Machine (SVC), Elastic Net, Logistic Regression, Extreme Gradient Boosting, Random Forest, Neural Network and Adaptive Boosting. The machine algorithm- Logistic Regression,

which is regularly used in the field of medicine was named as a side-by-side comparison of baselines. The other algorithms were chosen for relative ease in implementation, resiliency to overfitting, their ability to model non-linear associations and since they were widely accepted by the machine learning community.

The accuracies of different models were found to be: Random Forests-87.4%, Adaboost-85.6%, Support Vector Machine-86.3%, Elastic Net-86.4%, Logistic Regression-86.4%, Neural Networks-86.3% and XG Boost-87.5% [26].

4. Discussion and Conclusion

Machine learning is the science of making computers gain an understanding of and behave like humans, providing data and information without being specifically programmed. It can be classified into Supervised learning model, Semi-Supervised learning model and Unsupervised learning model. In this paper, a detailed review of machine learning classification is discussed. Machine learning uses various algorithm or models to perform a specific task. The primary use of these algorithms is in the medical prediction field. The focus is on the use and amalgamation of different algorithms for predicting different types of diseases using machine learning. Various research works with some effective techniques done by different people is studied.

In this paper, five such diseases of high priority are considered and different machine learning models are used for its prediction. Models like Back Propagation, Support vector machine, Naive bayes, Random forest, and K-Nearest Neighbors are used for the study and analysis of Liver disease, among these, Naive Bayes gave the highest accuracy of 95.1%. Chronic Kidney Disease is the second disease and models like Decision tree, support vector machine, random forest are used for analysis. The maximum accuracy of 99.16% is observed by the Random forest model. The analysis and prediction of Hematology included models like support vector machine, Naive Bayesian Classifier, Decision tree, Random Forest and the highest accuracy of 97.12 was obtained by Random Forest making it the most efficient model for the prediction of this disease. The study of Alzheimer's Disease included models such as Support vector machine, Naive Bayes, Random Forest. Reasonable accuracy of 74.65% is obtained for Naive Bayes. Different models were developed for UTI prediction employing seven machine learning algorithms namely- Neural Network, Random Forest, Support Vector Machine, Logistic Regression, Extreme Gradient Boosting, Adaptive Boosting and Elastic Net. Accuracy of 90.4% is obtained by XG Boost which proved to be the highest among all other models. From this study a wide overview of the relative performances of different variants of supervised machine learning for disease prediction is provided. The information provided on relative performances by this study can be used by researches in selecting appropriate machine learning algorithms for their studies.

Table 2. Comprehensive Survey on the machine Learning Techniques for Disease Detection

MODEL/DISEASE	LIVER	CHRONIC KIDNEY	ALZHEIMER'S	HAEMATOLOGY DIAGNOSIS	UTI
SVM	71%[15]	96.75%[18]	-	71.2% [15]	86.3% [20]
RANDOM FOREST	74%[16]	99.16%[19]	-	97.12% [15]	87.4% [20]
DECISION TREE/ XGBOOST	-	91.75%[18]	74.22%[23]	97% [15]	87.5% [20]
BACK PROPAGATION	73.2%[15]	-	-	-	-
KNN	62%[16]	-	43.26%[23]	-	-
NAÏVE BAYES	95.1%[17]	-	74.65%[23]	81.6% [15]	-
LINEAR REGRESSION	-	-	88.24%[23]	-	-
DEEP LEARNING	-	-	78.32%[23]	-	-
ADABOOST	-	-	-	-	85.6% [20]
LOGISTIC REGRESSION	-	-	-	-	86.4% [20]
NEURAL NETWORK/ MULTILAYER PERCEPTRON	-	-	-	91.8% [15]	86.3% [20]
BAYESIAN NETWORK	-	-	-	92.86% [15]	-

References

1. Nithya, B., and V. Ilango. "Predictive analytics in health care using machine learning tools and techniques." In 2017 International Conference on Intelligent Computing and Control Systems (ICICCS), pp. 492-499. IEEE, 2017.
2. Sharmila, S. Leoni, C. Dharuman, and Perumal Venkatesan. "Disease classification using machine learning algorithms-a comparative study." International Journal of Pure and Applied Mathematics 114, no. 6 (2017): 1-10
3. Das, Kajaree, and Rabi Narayan Behera. "A survey on machine learning: concept, algorithms and applications." International Journal of Innovative Research in Computer and Communication Engineering 5, no. 2 (2017): 1301-1309.
4. Kaur, Sunpreet, and Sonika Jindal. "A survey on machine learning algorithms." Int J Innovative Res Adv Eng (IJIRAE) 3, no. 11 (2016): 2349-2763
5. Narayanan, Uma, Athira Unnikrishnan, Varghese Paul, and Shelbi Joseph. "A survey on various supervised classification algorithms." In 2017 International Conference on Energy, Communication, Data Analytics and Soft Computing (ICECDS), pp. 2118-2124. IEEE, 2017
6. Dahiwade, Dhiraj, Gajanan Patle, and Ektaa Meshram. "Designing disease prediction model using machine learning approach." In 2019 3rd International Conference on Computing Methodologies and Communication (ICCMC), pp. 1211-1215. IEEE, 2019
7. Osisanwo, F. Y., J. E. T. Akinsola, O. Awodele, J. O. Hinmikaiye, O. Olakanmi, and J. Akinjobi. "Supervised machine learning algorithms: classification and comparison." International Journal of Computer Trends and Technology (IJCTT) 48, no. 3 (2017): 128-138.
8. Jamgade, Akash C., and S. D. Zade. "Disease prediction using machine learning." International Research Journal of Engineering and Technology 6, no. 5 (2019): 6937-6938.
9. Lim, Sunghoon, Conrad S. Tucker, and Soundar Kumara. "An unsupervised machine learning model for discovering latent infectious diseases using social media data." Journal of biomedical informatics 66 (2017): 82-94.
10. Sathya, Ramadass, and Annamma Abraham. "Comparison of supervised and unsupervised learning algorithms for pattern classification." International Journal of Advanced Research in Artificial Intelligence 2, no. 2 (2013): 34-38.
11. Reddy, Y. C. A. P., P. Viswanath, and B. Eswara Reddy. "Semi-supervised learning: A brief review." Int J Eng Technol 7, no. 1.8 (2018): 81.
12. Obulesu, O., M. Mahendra, and M. ThrilokReddy. "Machine learning techniques and tools: A survey." In 2018 International Conference on Inventive Research in Computing Applications (ICIRCA), pp. 605-611. IEEE, 2018.
13. Uddin, Shahadat, Arif Khan, Md Ekramul Hossain, and Mohammad Ali Moni. "Comparing different supervised machine learning algorithms for disease prediction." BMC medical informatics and decision making 19, no. 1 (2019): 1-16.
14. Gupta, Shubham, Vishal Bharti, and Anil Kumar. "A survey on various machine learning algorithms for disease prediction." Int. J. Recent Technol. Eng 7, no. 6c (2019): 84-87.
15. Sontakke, Sumedh, Jay Lohokare, and Reshul Dani. "Diagnosis of liver diseases using machine learning." In 2017 International Conference on Emerging Trends & Innovation in ICT (ICEI), pp. 129-133. IEEE, 2017.
16. Rahman, AKM Sazzadur, FM Javed Mehedi Shamrat, Zarrin Tasnim, Joy Roy, and Syed Akhter Hossain. "A Comparative Study On Liver Disease Prediction Using Supervised Machine Learning Algorithms." International Journal of Scientific & Technology Research 8, no. 11 (2019): 419-422.
17. El-Shafeiy, Engy A., Ali I. El-Desouky, and Sally M. Elghamrawy. "Prediction of liver diseases based on machine learning technique for big data." In International conference on advanced machine learning technologies and applications, pp. 362-374. Springer, Cham, 2018.
18. Tekale, Siddheshwar, Pranjal Shingavi, Sukanya Wandhekar, and Ankit Chatorikar. "Prediction of chronic kidney disease using machine learning algorithm." International Journal of Advanced Research in Computer and Communication Engineering 7, no. 10 (2018): 92-96..
19. Revathy, S., B. Bharathi, P. Jevanthi, and M. Ramesh. "Chronic Kidney Disease Prediction using Machine Learning Models." . International Journal of Engineering and Advanced Technology (IJEAT) 9, no. 1 (2019).
20. Gunčar, Gregor, Matjaž Kukar, Mateja Notar, Miran Brvar, Peter Černelč, Manca Notar, and Marko Notar. "An application of machine learning to haematological diagnosis." Scientific reports 8, no. 1 (2018): 1-12.

21. Pekelharing, J. M., O. Hauss, R. De Jonge, J. Lokhoff, J. Sodikromo, M. Spaans, R. Brouwer, S. De Lathouder, and R. Hinzmann. "Haematology reference intervals for established and novel parameters in healthy adults." Sysmex Journal International 20, no. 1 (2010): 1-9.
22. Alsheref, Fahad Kamal, and Wael Hassan Gomaa. "Blood diseases detection using classical machine learning algorithms." International Journal of Advanced Computer Science and Applications (IJACSA) 10, no. 7 (2019).
23. Shahbaz, Muhammad, Shahzad Ali, Aziz Guergachi, Aneeta Niazi, and Amina Umer. "Classification of Alzheimer's Disease using Machine Learning Techniques." In DATA, pp. 296-303. 2019.
24. De Roeck, Ellen Elisa, Peter Paul De Deyn, Eva Dierckx, and Sebastiaan Engelborghs. "Brief cognitive screening instruments for early detection of Alzheimer's disease: a systematic review." Alzheimer's research & therapy 11, no. 1 (2019): 1-14.
25. Flores-Mireles, Ana L., Jennifer N. Walker, Michael Caparon, and Scott J. Hultgren. "Urinary tract infections: epidemiology, mechanisms of infection and treatment options." Nature reviews microbiology 13, no. 5 (2015): 269-284.
26. Taylor, R. Andrew, Christopher L. Moore, Kei-Hoi Cheung, and Cynthia Brandt. "Predicting urinary tract infections in the emergency department with machine learning." PloS one 13, no. 3 (2018): e0194085.

RECENT ADVANCES IN GRAPH THEORY AND ITS APPLICATIONS

Dankan V. Gowda[1], K.S. Shashidhara, M. Ramesha, S.B. Sridhara, and S.B. Manoj Kumar

ABSTRACT. In different fields the field of mathematics plays a key role. In mathematics, graph theory is one of the important fields used in structural models. This structural structure of different objects or technologies leads to new developments and changes in the current world in these areas. The field graph theory began in 1735 with the Koinsberg Bridge problem. This paper provides a description of implementations of graphical theory in a number of heterogeneous areas but focuses mostly on information science, electrical engineering, linguistics, physics and chemistry, computer network science, biotechnology and graphical theoretical applications. Several articles focused on graph theory have been studied concerning scheduling principles, engineering technology implementations and an outline.

1. INTRODUCTION

A diagram consisting of many points and lines that unite several pairs of these points can be easily represented for several real-world contexts. The points might, for example, show individuals with lines who join couples with friends; or the points could be contact centers with lines showing connection connections. Notice that one is primarily concerned in such diagrams whether a line

connects two defined points or not; the way they are connected is immaterial. The definition of a graph is a statistical abstraction of conditions of this kind.Graph theory principles are commonly used in various fields to research and model different applications. This includes studying molecules, building chemical bonds and studying atoms. In sociology, for instance, graph theory is used to calculate the popularity of actors or to investigate processes of diffusion. The theory of graphs is used for biodiversity and conservation, where a vertex represents areas in which some species live and where edges represent migratory or moving paths between areas. This data is important for examining the breeding habits of disease, parasites and for investigating the effect of migration on other animals. This knowledge is important. In the field of computer science, graph theory concepts are widely used [1]. The graph theory uses algorithms such as Breadth First Search, Depth First Search, Topological Sort, Bellman-Ford, the algorithme of Dijkstra, Minimum Trees, the Algorithm of Kruskal and the Prim's.

2. HISTORY OF GRAPH THEORY

The root of the graphic principle began with the Koinsber bridge dilemma in 1735. This dilemma leads to the Eulerian graph principle. Euler analyzed the Koinsberg Bridge problem and created a structure to solve the problem known as the Eulerian graph. A.F Mobius offered the concept of a total graph and a bipartisan graph in 1840, and Kuratowski showed that they were planar of leisure problems. The principle of tree (Gustav Kirchhoff introduced a linked graph without cycles in 1845 and used graphical technical concepts for the measurement of current in electrical networks or circuits. In 1852, the popular four-color issue was discovered by Thomas Gutherie. Then in 1856, Thomas, P. Kirkman and William Hamilton, researched polyhydra cycles and developed, by observing trips which visited a number of locations exactly once, the idea called the Hamiltonian graph. In 1913, H. Dudeney spoke about an issue of puzzles. Eventually, Kenneth Appel and Wolfgang Haken addressed the four-color dilemma only after a century. This period the birth of graph theory is considered [2]. To research the trees Caley learned specific analytical forms from the differential calculus. And has several consequences for theoretical chemistry. This leads to enumerative graph theory being invented. Anyway, in 1878, Sylvester introduced "Graph," where he drew an analogy from "quantum invariants" to

algebra and molecular-diagram covariants [3]. In 1941 Ramsey experimented on the colours, leading to the identification of a subset of graphic science named severe graphic theory. In 1969, Heinrich's computers solved the four-color mystery. The analysis of asymptotic graph connectivity has led to a random principle of graphics.

3. Applications of Graph Theory

Graph theory principles are commonly utilized in diverse fields to research and model different applications. This includes studying compounds, building bonds in chemistry and studying atoms. In sociology, graph theory is similarly used for example to calculate the popularity of performers or to investigate processes of diffusion. Graphic theory is used in biology and conservation where the vertex describes the areas in which animals occur and the edges reflect the direction of migration or travel through regions. This knowledge is critical for examining breeding trends or monitoring the propagation of diseases and parasites and for investigating the effect of migration on other animals[4,5]. Theoretical graphic principles are commonly utilized in research operations. For example the dilemma of the tour sales person, the shortest stretch in a weighted graph, obtains optimal work and men match and finds the shortest route from two vertices in a diagram. It is also used for modelling transport networks, networks of operation and game theory[6]. A digraph is used to describe the finite game method. The vertices here mark the locations and the edges represent the movements. Graph theory is widely employed in research and technology. Any of the following are given:

3.1. **Computer Science.** For the analysis of algorithms such as: Dijkstra Algorithm, Prims Algorithm, Kruskal Algorithm theory is used in computer graphics. Anwendungsbereiche such as Graphs are used to describe the calculation flow. Graphs are used to portray contact networks. Graphs reflect the organization of results. Graph transformation schemes operate on the manipulation of graphs on the basis of rules. Graph databases ensure secure, continuous storage and querying of organized graph data. Graph theory is used for finding the shortest route or network direction. Google Maps shows different places as vertices or points, and the roads are seen as corners and the idea of the chart is used to find the shortest path between two nodes.

3.2. **Electrical Engineering.** Graph theory is used in electrical engineering in the construction of circuit links. This relations are referred to as topologies. Certain topologies include sequence, bridge, star and parallel topologies.

3.3. **Linguistics.** Graphs are mostly used for the parsing of a language tree and a language tree grammar in linguistics. In lexical semantitics the semantitic networks are used, particularly for computers, and the modelization of word sense is simpler when a word is interpreted in relation to the word. Phonological methods (e.g. optimal theory that is based on grid diagrams) and morphology (e.g. finite state morphology utilizing finite-state transducers) are popular in linguistic research as a diagram.

3.4. **Physics and Chemistry.** Chemistry graphs are used to model chemical compounds. Any sequences of cell samples may be omitted in statistical bio-chemistry in order to overcome contradictions between two sequences. This is modelled as a graph in which the vertices reflect the sample sequences. An edge is drawn between two vertices where there is a conflict between the sequences. The goal is to delete potential vertices (sequences) in order to remove all disputes. In short, graphic theory has a special influence in several areas and is already spreading over many days. The following section analyzes the uses of graph theory in computational sciences in particular. Chart theory is used in physics and chemistry to analyze molecules. The 3D layout of complex artificial atomic systems can be quantitatively analyzed by collecting statistics on graph-theoretical features in relation to atom topology. Graphs are also used in statistical mechanics. In this area, diagrams may describe local relations between the interacting sections of a system and the physical process dynamics on those structures. Graphs also express porous media micro channels in which the vertices reflect the pores and the borders represent the smaller pores. Graph is also useful in building both the molecular structure and the molecular grid. It also allows us to demonstrate the connection between atoms and molecules and helps us to compare the structure of a molecule with another.

3.5. **Computer Network.** The ties between linked computers in the network obey the concepts of graph theory in the computer network. Graph theory is often used for protection of the network. We will use the vertex coloring algorithm to paint the map in four colours. Vertex Coloring Algorithm may be used

to delegate a maximum of four distinct frequencies to any mobile network GSM (Grouped Special Mobile).

3.6. **Social Sciences.** In sociology, graph theory is also used. For example, to explore the dissemination of rumor or to calculate the credibility of actors by the use of tools for social network analysis. Friendship and knowledge graphs describe whether or not individuals meet one another. Some individuals may affect the behavior of others in influential diagrams. In collaborative graphs model, two individuals operate together in a similar context, for example participating in a film together.

3.7. **Biology.** Nodes in biological networks are bimolecular such as chromosomes, proteins or metabolites and edges that link the nodes signify interactive, physical or chemical interactions between the bimolecular concerned. In transcriptional regulatory networks, graph theory is used. It is seen in metabolic networks as well. Graph theory is also useful in PPI (protein interaction) networks. Characterizing drug goal partnerships. drug target interactions.

3.8. **Mathematics.** Operational analysis is the essential area of mathematics. Graph theory offers numerous practical organizational analysis uses. Like: Minimum route expenses, A issue with the schedule. Graphs reflect the roads between the towns. We may construct hierarchically organized details such as a family tree with the aid of a sort of graph.

4. CONCLUSION

Programmers and designers, graph theory is an extraordinarily rich field. Graphs can help solve some very complicated issues, such as lower costs, visualization, program analysis, etc. To calculate an optimum traffic routing, network devices, such as routers and switches use graphics. This paper focuses mainly on presenting the recent developments in the field of graph theory and its various applications in the field of engineering. In particular, the concept of graph theory is outlined in an overview. Researchers in different streams, such as engineering, social science, general sciences etc., benefit from this. There is a wide discussion of each domain application, which is very beneficial to any researchers.

<div align="center">REFERENCES</div>

[1] G. MARCIALIS, F. ROLI: *Graph Based and Structural Methods for Fingerprint Classification*, Springer verlag, Berlin Heidelberg, **9**(1) (20018), 1–202.

[2] S. DICKINSON, R. ZABIH: *Introduction to the special section on graph algorithms in computer vision* , IEEE on pattern analysis, **23**(10) (2016), 114–122.

[3] B. HONG LIU, W. CHIEH KE: *Constructing a message pruning tree with minimum cost for tracking moving objects in wireless sensor networks*, IEEE, **57**(6) (2017), 16–22.

[4] S. SKIENA, S. PEMMARAJU: *Implementing Discrete Mathematics-Combinatorics and Graph Theorywith Mathematica*, Addison-Wesley Publishing Company, **3**(9) (2019), 1–448.

[5] X. ZHOU, T. NISHIZEKI: *Edge-coloring algorithms*, Technical report, Graduate School of Information Sciences, Tohoku University, Sendai, **1**(3) (2020), 120–142.

[6] A. VINCE, C. HAALEY: *Star chromatic number*, Journal of Graph Theory, **12**(4) (2020), 551–559.

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, B.M.S INSTITUTE OF TECHNOLOGY , VISVESHWARAYA TECHNOLOGICAL UNIVERSITY , ADDRESS: BANGALORE, KARNATAKA STATE, INDIA

Email address: dankan.v@bmsit.in

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, NITTE MEENAKSHI INSTITUTE OF TECHNOLOGY , VISVESHWARAYA TECHNOLOGICAL UNIVERSITY , ADDRESS: KARNATAKA STATE, INDIA

Email address: sks.nmit@gmail.com

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, GITAM SCHOOL OF TECHNOLOGY , GITAM (DEEMED TO BE UNIVERSITY) , ADDRESS:BANGALORE, KARNATAKA STATE, INDIA

Email address: rameshmalur037@gmail.com

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, VIJAYA VITTALA INSTITUTE OF TECHNOLOGY , VISVESHWARAYA TECHNOLOGICAL UNIVERSITY , ADDRESS: BANGALORE, KARNATAKA STATE, INDIA

Email address: sridharasb@gmail.com

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING, B.G.S INSTITUTE OF TECHNOLOGY , ADICHUNCHANAGIRI UNIVERSITY , ADDRESS: KARNATAKA STATE, INDIA

Email address: manojkumarsb.ece@gmail.com

Energy-Efficient Hybrid Hierarchical Routing scheme for overall Efficiency in WSN

Raghunandan G H[a]*, Dr. A ShobhaRani[b]

[a]*Department of Electronics and Telecommunication Engineering, BMS Institute of Technology and Management, Bangalore
India
[b]Department of Electronics and Communication Engineering, BMS Institute of Technology and Management, Bangalore, India

Abstract: Wireless Sensor Network are spatially distributed sensors intended to monitor different physiological conditions. Sensing and communicating data from one place to another consumes more energy, therefore the management of sensor energy is a very important factor. Energy utilization, synchronization, and a lifetime of the network are the main criteria in WSN. More energy is utilized by sensors that are distant from the base station. The gateway nodes are deployed to collect and relay information from nodes to the base station to resolve this problem. To decrease energy consumption, gateway nodes are deployed in the network. In this paper, a hybrid approach is used to increase the overall efficiency of the network in WSNs with time synchronization which increases the throughput of the network. The efficiency of proposed protocol based network has shown improvements in network lifetime, residual energy, data packets, and the throughput of the network. The performance of WSN of the proposed scheme is compared to other classical routing schemes and the proposed algorithm has proved its merit.
Keywords: Gateway Node , Wireless Sensor Network , Routing protocols, Base Station , Centrality, Cluster head

_____

1. Introduction

Today due to an increase in the demand for continuous monitoring of physiological changes of different environments has led to major developments and growth of micro-sensors. Previously few industries use wired sensors; implementation provides deployment of sensor nodes more viable than before. Previously, there has been a study regarding applications of WSN such as monitoring of environmental changes, the agricultural field, military surveillance, and home automation. Deep research on routing protocols efficiency is being done considering the power constraints in WSNs .Recently there has been more scope in the field of WSNs and their applications because they are easy to deploy and are of low cost, have flexibility. Algorithms are used to solve the problem of power constraint without altering the standard. Local collaboration among sensors, suppression, data compression, redundant data, and avoidance of direct transmission too far distant sensors is of the major factors that influence algorithm designers to device unique distributed, scalable and energy-efficient solution for Wireless Sensor Networks. In common, the sensor nodes measure environmental conditions. The sensor node extracts some useful information by processing the raw sensor signals. The output of this processed signal is transmitted through direct communication or multi-hop communication with access points across other sensor nodes. In some situations, repeaters (RPs) are used for multi-hops, to support sensors installed outside the radio range. One of the components of the WSN is the base station which has more energy, computational, communication resources. Forwarding of data from wireless sensor network onto a server is done by the BS which acts as a gateway. has less energy consumption when compared to the communication subsystem. Figure 1 shows the arrangement of WSN. The energy required for transmitting one bit may consume as same as executing a few thousand instructions. Hence, communication must be traded for computation. In different environments, together with remote and hostile regions, where ad-hoc communications are a key element, WSNs are to be deployed in large numbers. For this reason, algorithms and protocols should concentrate on subsequent issues:

Lifetime enhancement: Sensor nodes should be energy efficient and the consumption of energy of the device should be less due to the limited energy resources. The radio power supply when not in use should shut off to conserve power of the node

- Fault tolerance and Robustness.
- Self-configuration.
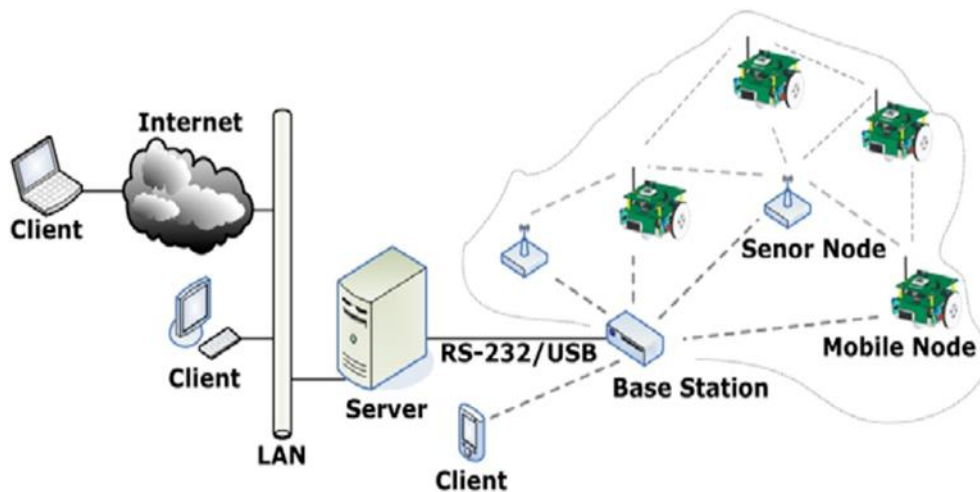
Figure 1: Schematic of WSN

Hierarchical routing algorithms are more efficient algorithms. In this, CHs are responsible for collecting the data and send it to GN and then the GN sends the information to the BS. If the members of the cluster are far, then there is wastage of energy. To solve this problem, the Centrality approach is used. In this method, the distance between every node to other nodes in the cluster is calculated using the Pythagoras theorem. The node which is equidistance to all other nodes along with high energy will be selected as the cluster head. This will reduce energy consumption. Error during communication can be reduced by using time synchronization in communication. This will reduce the packet collision during data transmission. In this paper, a new synchronization method considering a new broadcast sequence is used. This determines the order of the nodes that transmit timing information, this Eliminates collision. Section II presents a summary of the existing systems. The framework of the proposed system is explained in Section III. The system is evaluated in Section IV based on simulation and implementation. The paper concludes with section V.

2. Existing System

In [2] network is divided into areas as per geographic locations based on clustering for Cluster Head selection and formation in WSNs. CH. After a few iterations, CH's energy level decreases due to more energy usage as it processes the data obtained before forwarding it to the Base Station (BS). It is also no longer worthy of being the head of a cluster. Thus, the formation of the cluster takes place when CH residual energy is below the threshold goes; hence cluster reformation is not that efficient comparatively. In [3] It aims to minimize energy consumption at the level of the sensor node and the level of the network in the WSN. Before the available transmission, and then, the lowest transmission power required for transmitting, the distance between the transmitter and the receiver is calculated, but here the neighbor status information is less. In protocol [14] balance energy consumption is implemented for Cluster Heads (CHs) by creating clusters in each round with an almost equal number of nodes. But this is may not be applicable in all conditions as the nodes are randomly distributed based on applications. For transmission of data to the base station majority of the energy of sensor nodes is used. Thus, there is a fast depletion of energy. In [15] only energy conservation is concentrated which is the modified form of LEACH delay is not considered as a major factor.

Hierarchical Protocols opens up several technological challenges and enormous implementation possibilities as sensor networks interconnect many other nodes when large networks are set up. Using multi-hop wireless communications, these wireless sensor networks communicate. Due to the limited computational resources and finite power available to each sensor node, standard ad hoc routing methods cannot be directly applied to the domain of sensor networks. Protocols for Routing. In [6] for wireless sensor networks are used to ensure stable multi-hop communication and to maintain routes on the network. It provides us with an idea on routing protocols for Wireless Sensor Networks and compares their relative strengths and limitations to provide better energy efficiency or increase wireless sensor networks. Routing in WSNs is generally, classified into 3 types depending on the network structure. Data-centric routing protocols are often referred to as the first routing protocol. In this, each sensor node works together to perform the task of sensing. In the second kind of routing is hierarchical routing, higher energy nodes are used to process and send information, and low energy nodes are used to perform the sensing in the closeness of the target. In the third category location based routing, the sensor node's area is addressed through its location. By sharing such information between neighbors, relative co-ordinates of neighboring nodes can be obtained. In the current method, there are some routing problems and design concerns. One should assume the uniform distribution of sensor nodes when developing routing protocols; otherwise, we should go for clustering. Production costs should also be lower and service quality should be high.

3. Material Method

In the recent past, wireless sensor network has attracted substantial research attention as WSN's are fast-growing In hierarchical-based routing algorithms, Cluster Heads are in charge of compressing, forwarding and, gathering, data to the BS. Therefore, they play a major role. For the dissemination of all information to the members of that cluster. It is also critical that the clusters are developed for efficient WSN.CHs is also elected based on certain criteria. Thus efficiency, Network lifetime, and network stability of WSN are increased. In the Proposed algorithm, the sensor member senses the data and sends the data to the respective Cluster Heads. Cluster Head processes this data and sends it to the gateway node which in turn forwards to sink. Hence the efficiency of the CH increases which increases the network lifetime.
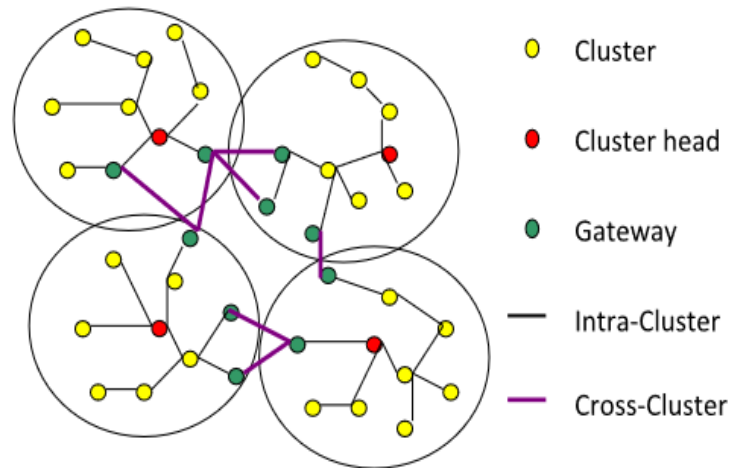
Figure 2: Clustering in WSN

Time synchronization is an elementary part of any network-oriented organization and system. In our proposed system grouping of sensor nodes is done called clustering. The transmitter's energy dissipated for the transmission of bit messages to a distance is provided by,

If the distance between the nodes is less than the threshold distance then the following equation will be used

$$\text{Energy} = ((E_{TX}+E_{DA}) \text{ (packet length)} + E_{mp} \text{ (packet length) (distance)}^2) \tag{1}$$

If the distance is greater than the threshold then the following equation will be used

$$\text{Energy} = ((E_{TX}+E_{DA})(packetLength) + E_{fs} \text{ (packetLength)(distance)}^4) \tag{2}$$

The energy dissipated by a receiver to receive bit message is given by

$$E_R(k)=k(E_{RX}) \tag{3}$$

Threshold distance is given by

$$r_0 = \sqrt{\frac{E_{fs}}{E_{mp}}} \tag{4}$$

Clustering is done to construct the appropriate topology of the network. The clustering-based network reduces the cost of the routing algorithm and the flooding broadcast. In this cluster-based routing approach, sensor nodes that are in a particular radio range are grouped which forms a cluster. Each group or cluster has one cluster head (CH). This CH collects all the data from sensor nodes in a cluster and it performs data fusion and sends it to the nearby gateway node. CH is selected based on the centrality approach. In this method distance of nodes between each other is determined; the node which is equidistant from the other node and the node having more energy is selected as CH. After CH selection aggregation of data is performed by removing redundant data. Data aggregation is the process of removing redundant data during transmission. This will increases the lifetime of the network. In our system base station (BS) is fixed. The same algorithm can also be applied to a movable base station. During the data communication phase, the data packets are transmitted from CH to the gateway node and from the gateway node to BS. This process is called multi-hop communication. The proposed technique uses the receiver-to-receiver concept implemented

by the reference broadcast Synchronization which minimizes the time-critical path when compared to the sender-to-receiver method. The following assumptions are made for the proposed algorithm:

- The nodes have equivalent initial energy and are left unattained after deployment.But limitations on energy, memory, and computation are not considered for Base Station The nodes can limit the power required for transmission to the distance.
- Using wireless radio signal strength distance can be calculated.
- Usually, one of the causes of node failure is energy depletion.
- Sensor nodes are immobile.

Clustering is a process of connecting nodes using a specific topology to perform certain tasks as per the requirements. For wireless sensor networks, the algorithm used finds a collection of distinguished nodes to create the required network topology. The next step after the deployment of the sensor nodes is grouping the sensors into the cluster. In the proposed algorithm, cluster formation is the same as that employed in the LEACH algorithm. At the beginning of each round, after grouping nodes into clusters, the cluster head (CH) is chosen. A threshold is set, up to which the proposed algorithm is used for the cluster head selection along with the centrality based approach. Data transmission is continuously monitored with energy updating after each round. The distance between one node and every other node is determined here, and the equidistant node is chosen as the head of the cluster.CH is selected based on certain criteria as mentioned below

1) CH's residual energy: The CH is selected based on the maximum amount of energy it has.
2) The distance of the cluster head to the base station. The more distance from each other the more energy required for data transmission
3) Depends on the number of nodes in the cluster that is in a cluster if there are more members data processing required is also more.

In the Proposed Algorithm Agent gateway node approach is applied for the network. The Sensors are connected to corresponding CH and CH to nearby Gateway nodes.The Gateway nodes (GN) will be connected to the Base station (BS).Time Synchronization algorithm is applied between Sensors, CH, and GN. A reference node for time synchronization is selected for the first time. Later Source Receiver synchronization is applied and Beacon messages are sent to the nodes. Based on the radio energy dissipation model, the energy requirement of sensors are determined. The Routing of Data from nodes to CH, CH to GN, and the base station will be carried out.

4. The Research Findings And Discussion

The proposed system is simulated using MATLAB software. Here consider 100 nodes are distributed randomly along with gateway nodes. MATLAB Simulation is done considering the following parameters as shown in table 1. The proposed algorithm results are compared with DR-LEACH , EPEGASIS , and Mod-LEACH. Figure 4 shows the initial topology of the hierarchical approach. Figure 4 shows LEACH topology. Figure 5 shows the topology of the proposed system. In the proposed system gateway node along with time synchronization is applied. Figure 6 illustrates the network's residual energy which is applied with the proposed algorithm. Figure 7 shows the presence of dead sensors after around 1350 rounds. The number of dead nodes of other classical routing schemes is more compared to the proposed system.

Table 1: Network Considerations

Parameter	Considerations
Initial Energy of sensor	10J
Energy consumed to transmit one bit	50nJ
Number of bits per packet	6400 bits per packet
Energy for E_{fs}	10 pJ/bit/sq.m
Energy for E_{mp}	5 nJ/bit/signal
Number of Sensor Nodes	100
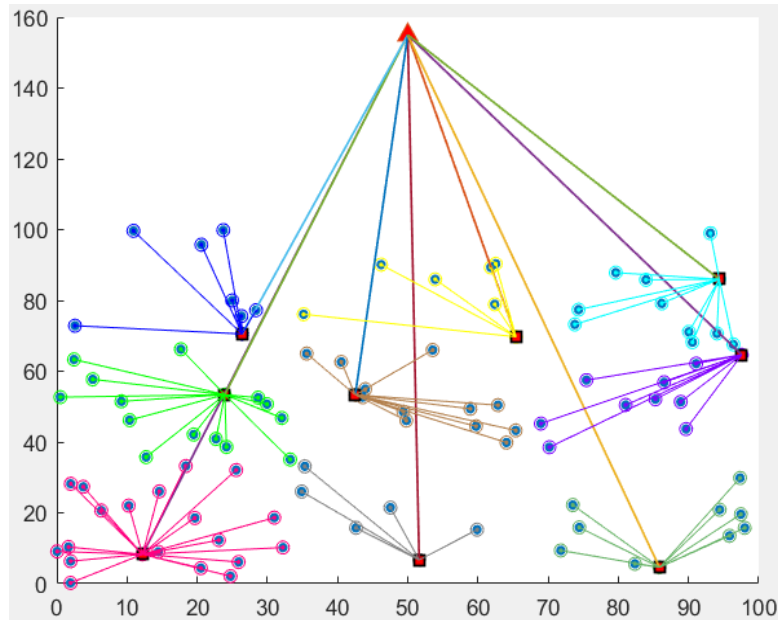
Figure 4: Initial Topology of LEACH
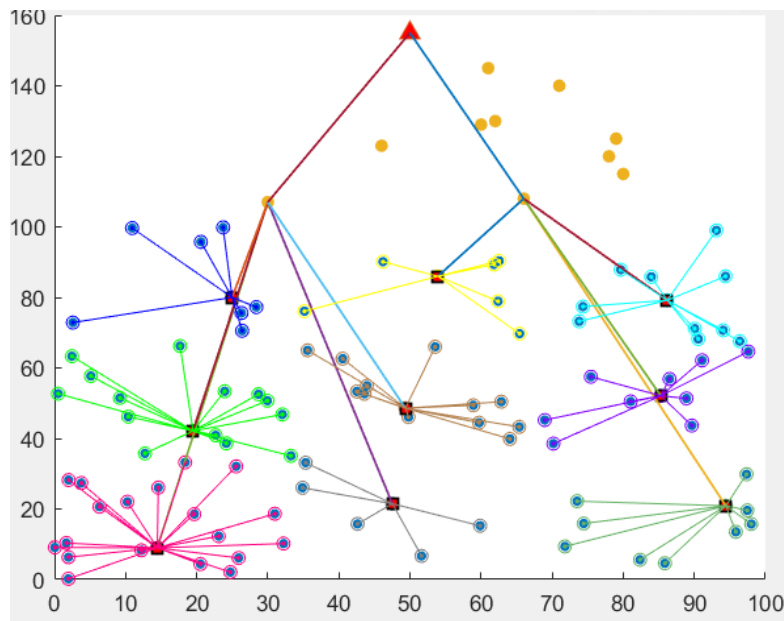
Figure 5: Proposed System

Figure 8 shows the number of packets transmitted to the base station to rounds. Here also the improvement in the number of data packets to the base station is more in the proposed algorithm compared to other routing schemes. The is a considerable lifetime enhancement in the proposed system compared to other schemes as shown in Figure 9.Figure 10 shows the delay in transmitting the data to the base station the proposed system shows less delay compared to other routing schemes. The delay is reducing gradually as synchronization among the nodes improved. The throughput of the network using different routing schemes is shown in Figure 11.here also the proposed algorithm shows a considerable improvement. This illustrates that the proposed system is more efficient compared to the existing protocols which are considered for Comparision. Throughput is calculated using the formula given below,

Throughput =[Sum (number of successful packets)*(average packet length)]/ (Time taken by each round)

(5)
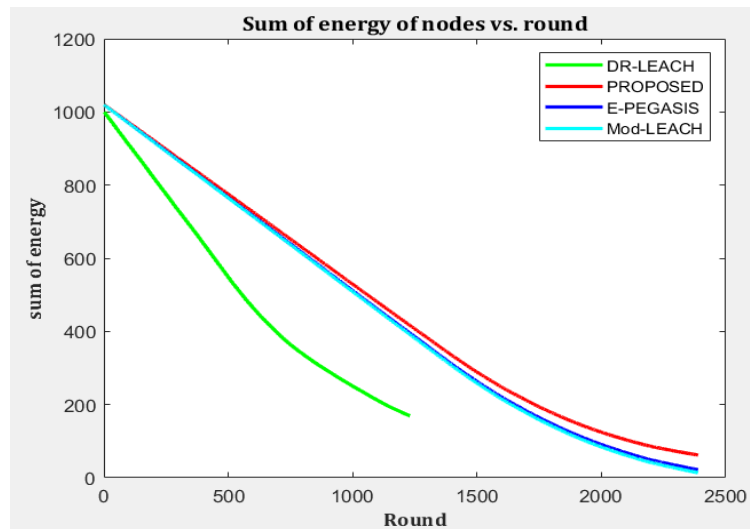
Figure 6: Comparision of Residual Energy of proposed system DR-LEACH, EPEGASIS, and Mod-LEACH
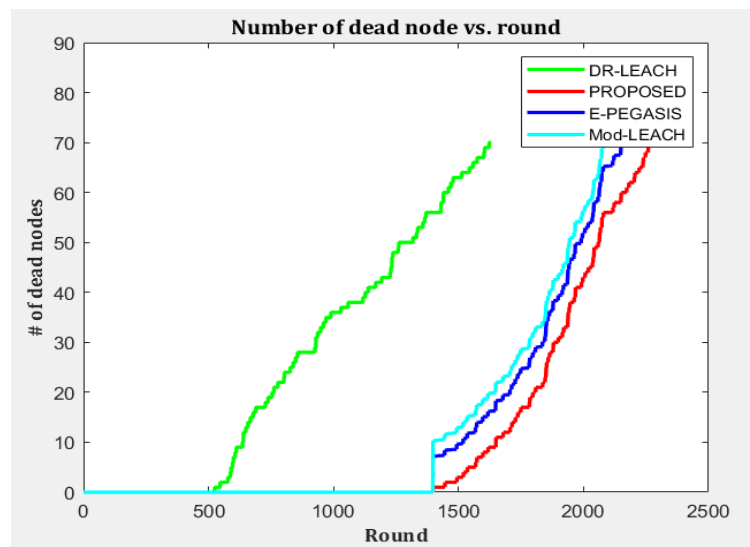
Figure 7: Comparision of number of dead nodes of the proposed system with DR-LEACH, EPEGASIS, and Mod-LEACH
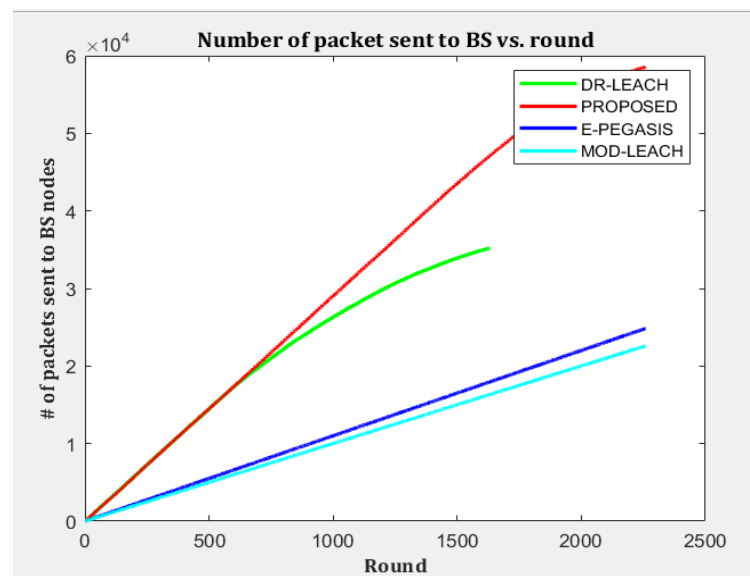
Figure 8: Comparision of packet delivered by the proposed system with DR-LEACH, EPEGASIS, and Mod-LEACH
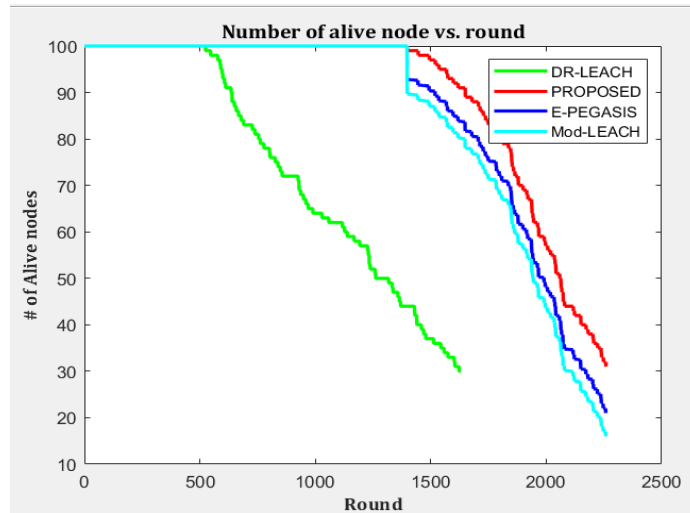
Figure 9: Comparision of Network Lifetime of Proposed system with DR-LEACH, EPEGASIS, and Mod-LEACH
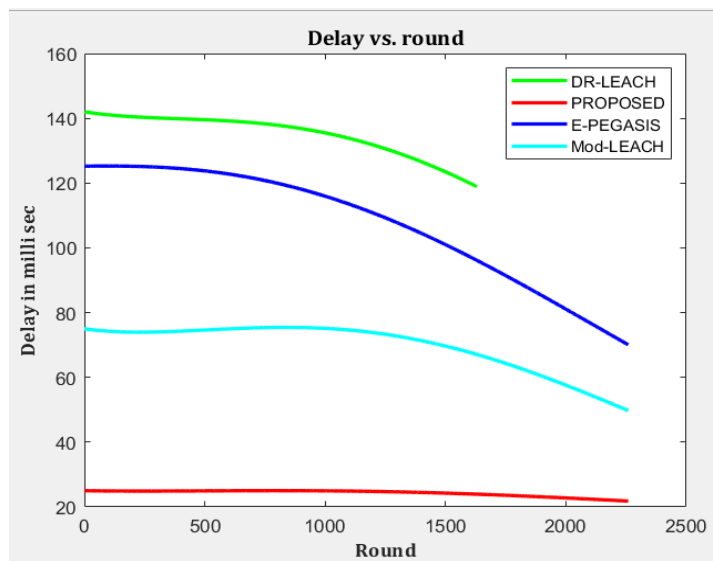
Figure 10: Comparision of delay by the proposed system with DR-LEACH, EPEGASIS, and Mod-LEACH
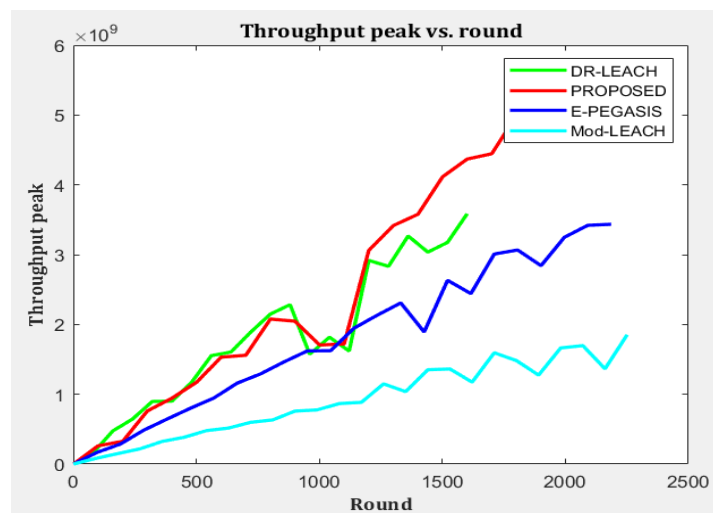
Figure 11: Comparision of Throughput by the proposed system with DR-LEACH, EPEGASIS, and Mod-LEACH

Here we are comparing our proposed system with leach protocol. Figure 9 indicates the presence of alive nodes after corresponding iterations. The number of dead sensors appears nearly around 1000 iteration as shown in Figure 10. Hence the network lifetime is improved and the residual energy after each round provides us the information about the amount of energy consumed as in figure 11. Hence there is an improvement in the lifetime of the network of our proposed system.

5. Results

Energy consumption by sensors is a major parameter for network lifetime in WSN. In hierarchical clustering methodology, the cluster head selection plays an important role. Power consumption is one of the main factors for network lifetime in the wireless sensor network. By alternating this cluster head within the individual clusters, power consumption can be optimized by considering whether the network is taken as a whole. The choice of the cluster head between the nodes of the wireless sensor is based on the method of centrality. The implementation of gateway nodes would also reduce the load on the cluster head, which in turn extends the lifetime of the network. The outcome of the proposed approach is an improvement in the network lifespan and the number of nodes that remain alive for the maximum period. An efficient technique that is required for reducing the energy consumption by the nodes in the wireless sensor network is discussed. Along with that time synchronization is achieved using reference broadcast synchronization. The implemented protocol shows an improvement in the lifetime network, more data packets are sent with less delay. The throughput of the network using the proposed system is more compared to classical routing schemes. Here the proposed system is compared with the DR-LEACH, EPEGASIS, and Mod-LEACH protocol and it proves that the proposed system is more efficient than DR-LEACH, EPEGASIS, and Mod-LEACH.

Conflicts Of Interest

No conflict of interest was declared by the authors..

References

Salah M., Boulouz, "Energy Efficient Clustering based on LEACH", International Conference on Engineering and MIS, (2016).

D Djenouri, "R4 Time Synchronization in WSN", IEEE signal processing letters, Vol. 19, issue 4, (2012).

Sang H. K, Nguyen, "Distance-Based Thresholds for CH selection in WSN", IEEE Communications Letters, Vol 16, Issue 9, (2012).

Raghunandan, Sagar, "A Novel Approach to increase overall efficiency in WSN ",2012 IEEE ICCEET proceedings, pp 699-703, March 2012.

Raghunandan, Lakshmi, "A Comparative Analysis of Routing Techniques for WSN", IEEE NCOIET-2011, IEEE Conference Proceedings, (2011).

Akkaya, Younis, "A Survey of Routing Protocols in WSN" Journal of Ad Hoc Network, Elsevier, vol. 3, pp. 325-349, (2005).

Ying H, Chin Y, Ting C, " AER Protocol for mobile sink in WSN," IEEE Conference on Media Computing, conference proceeding, pp. 44-49, (2008).

Raghunandan.G H, Lakshmi "Secure Routing protocols for WSN", International Conference on FCS proceedings, (2011).

Yunhe, Zhihua, Zhang Q, "Efficient Load Balance Data Aggregation methods for WSN based on compressive network coding ", IEEE International Conference on EICT, (2016).

Amit R, R Singh, A Nandini "WSN – Challenges and Possibilities", IJCA, volume 140, no 2, (2016).

Moumita, Paramita C, Sarmistha, Chowdhury "Analysis of EE WSN Routing Schemes", International Conference of EAIT, (2014).

Abhishek C, Sumedha, "Minimization of Avg Energy Consumption to Prolong Lifetime of WSN", IEEE Global Conference on WCN, (2014).

Zhenpengp,wuxiongzhang et.al," Energy-efficient improvements in mod-LEACH protocol on optimal cluster head selection", IEEE conference on IoT, Green computing and communication, (2018).

JinWang, Yu Gao, Xiang Y, Feng,"An Enhanced PEGASIS Algorithm with Mobile Sink Support for WSNs" Hindawi, Wireless Communications and Mobile Computing, Volume 9,(2018).

Abdallah I, Abdalraheem I, Huthaifa Al-Issa," Dynamic Re-Clustering LEACH-Based Protocol For WSNs", IJCNC, Vol.7, No.6, (2015).

Structural Health Monitoring of Bridges Using WSNs

Vinutha B[1]*, Prathiba N[2], Raghunandan.G.H[3]

[1]Departmentof Electronics and Communication Engineering,
[2,3]Departmentof Electronics and Telecommunication Engineering,
[1,2,3]BMS Institute of Technology and Management, India
[1]vinuthabacademia@bmsit.in, [2]pratibha.yashas@bmsit.in, [3]raghunandangh@bmsit.in

Abstract— Wireless sensor networks (WSNs) for Structural health monitoring (SHM) has earned immense attention in research as it has potential to decrease expenses associated with the maintenance cum installation of these monitoring systems. SHM systems, traditionally are being utilized to screen perilous substructures such as bridges, tall buildings, fields and is capable to increase structure lifetime and safety of public. Earlier wired networks were used by SHM. But this resulted in heavy cost in terms of maintenance and deployment. This project introduces a wireless sensor network using Bluetooth nodes for detection of damage if any in any type of structure in which the data is processed locally in real time and helps in sending alerts once the damage is identified.

Keywords— GSM, Sensors.WSN, agriculture, precision

1. INTRODUCTION

Bridges are endlessly exposed to very harsh environmental effects such as aging of materials, extensive rust of steel bars in tangible erections, rusting of steel erections and its parts, growing heaps of traffic and overfilling, or just complete weakening and old. All these elements together with negatives of proposal, creation, and unintentional damage, triggers the worsening of bridges and results in the injury of bridges in terms of carrying sufficient load. The state of extensively used built-up bridges is even inferior, one in four is categorized as not able to adapt modern vehicle masses or as aging and volumes of congestion. Therefore, a major amount of these erection needs firming, restoration, or replacement, but due to lack of public funds mostly offered for the essential renewal of standing buildings or creation of innovative ones. Bridges can undergo physical deprivation due to aging, mismanagement, or less maintenance. One major factor that has been not considered is the insufficient checking and supervising of prevailing structures, among the many factors which have led to the substandard state of bridge structures. In order to sanction some of the design parameters made, also to deliver real-time opinion at the time of construction (specifically right for novel bridges), and to estimate the actual present status of the bridge, it should allow the engineers to take well-versed choices regarding upcoming aspects and repair actions or to propose preservation (specifically for enduring bridges).Obtaining the quantitative data with respect to the structural behavior, is the most important objective for monitoring a bridge. In the subsequent part, in order to safeguard the structure, the monitoring structure is used that which offers early warning of an acceleration of the notorious deterioration that are being noticed.

And then to carry out a controlled lifetime allowance of the bridges with known problems, application of SHM to existing bridges has significantly amplified in current years.

2. EXISTING WORKS

The existing bridge monitoring system for public safety uses MBM (Monitoring Based Maintenance) technology that qualifies the engineer to screen the status of bridge, but not in real time. The sensors are clinged on load cables and everything is connected via wired. Highway bridge systems and flyovers are life-threatening in various areas, as they are being used since many years. The health of these poor conditioned bridges is important to be monitored and report should be generated when there are maintenance operations required. Recent advancements in sensor technology have made it possible for the automation of factual-time bridge health checking system. However, prevailing structure uses complex and high rate wired network amid sensors in the bridge and high rate optical cable between the management hub and bridge. In this venture, an idea of autonomous structural health checking of bridges using wireless sensor network is proposed.

Table.1: Comparison of Various Wireless protocols

Wireless protocols	Advantages	Disadvantages

1. Wi-Fi	High speed and higher data rates.	High power consumption.
2. ZigBee	Low power consumption	Not compatible with smartphones
3. Bluetooth	Compatible through smartphones	Low data rates (but sufficient for sending alerts)

We propose the Bluetooth technology to implement the WSN for SHM because of its low cost and complexity, medium range (around 10m), considerably good data transfer rate (around 1Mbps) and low power consumption. It is easy to implement a WSN using Bluetooth protocol.

The main reason we have chosen Bluetooth for our project is because it is present in almost all mobile devices and it is easy to send alerts to the commuters using this protocol.

3. DESIGN METHODOLOGY

3.1 Design of Detection Module

1. Arduino Uno R3: A dual-inline package with ATmega328 microcontroller at core with an operating voltage of 5V. It comprises of 14 digital input/output pins. Has a Flash memory of 32KB and SRAM of 2 KB. It communicates through SPI, I2C and UART serial communications. It weighs 25g. It acts as the core of the system which helps to coordinate data movement from the sensor and the SIM808 module.

2. Flex Sensors: Variation in resistance is directly proportional to sum of curve on the device. They alter the variation in curve to electric confrontation. They are usually a shrill band of 1''-5'' lengthy variation in confrontation. It is to detect the depth of bend in the structure.

3. LM35: An electronic device that helps in heat measure of its surroundings and alters the input data to electronic data to record, watch or indicate heat variations. LM35 out pin is connected to any one analog input pin of the Arduino.

4. SW420: This sensor safeguards piezoelectric transducer. Voltages are generated when the mechanical neutral axis of transducer is displaced, and this displacement results in tension of piezoelectric element. It detects unfamiliar quivering in structure during earthquake.

5. SIM900: Is a quad-band GSM/GPRS module which combines GPS technology. Works on voltage supply in the range of 3.4-4.4V and consumes less energy. It is also Bluetooth compliant and the GPS receiver has 22 channels and 66 tracking channels for procurement. It helps to send messages and update the cloud using the GSM/GPRS technology.

6. HC-05: It is a Bluetooth module which is designed for wireless communication. This is used in Master-slave configuration. It interconnects with microcontrollers using serial communication (UART).

3.2 Considerations for a Crack in bridge

Inertial Measurement Unit (IMU) sensors integrate combinations of flex sensor, temperature sensor and vibration sensors. Standard parameter values are -55 degree Celsius to 150 degree Celsius for temperature senor and 60K -110K ohms for the flex sensor any variation above or under these values can be taken into consideration.
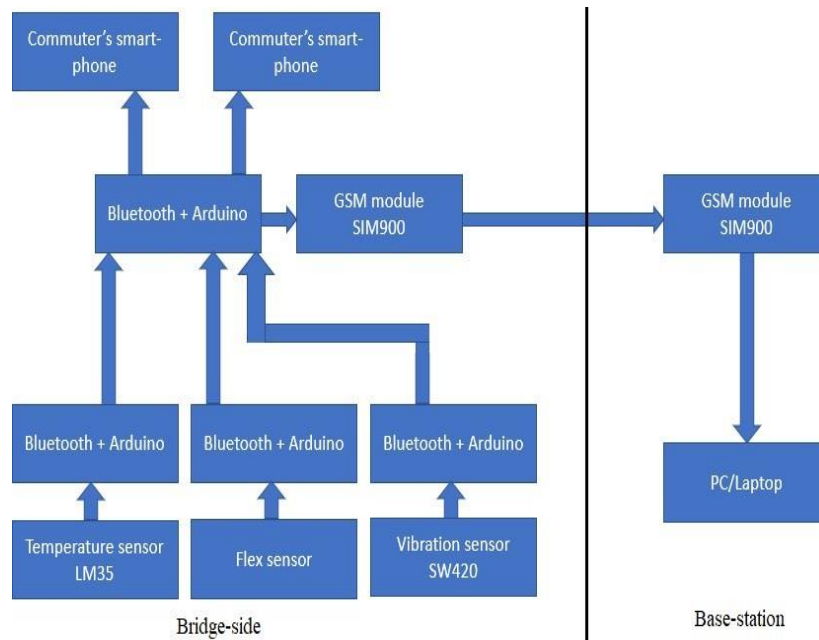
Fig 1. Basic Block Diagram

Various parts of the bridge are installed with sensors as depicted in the above block diagram. This continuously observes the bend, vibrations, temperature of the structure, etc. At any provided fact of time, the communication system notifies the base station giving an alarm for taking necessary deterrent measures, if any of these constraints cross their verge of edge value. The communication between the nodes is established using Bluetooth modules. The receiver module and sends an alert message to the base station that takes the data from the transmitter. The communication is established between the transitional Bluetooth node and the base station using GSM module. The sensory inputs are processed to signify the condition of the structure.
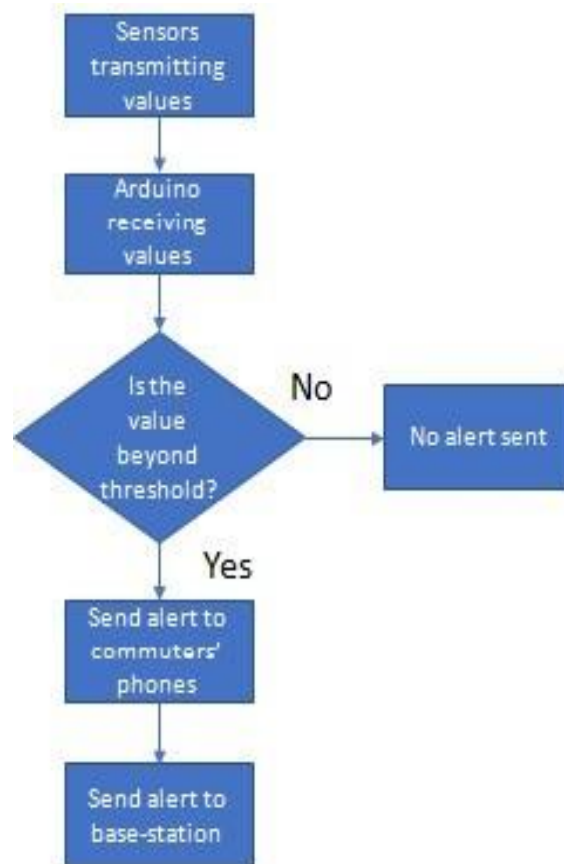
Fig 2. Flowchart depicted of SHM system

3.3 Alert System

The Arduino code is developed to do all tasks of a alert system that includes constantly checking if   the crack has occurred, reading the location, sensor data from the UART communication, making Places Search. Notify the team to repair the location. Sending SMS to all nearby people traveling on that road. Sending the alert message to all nearby essential services. In case, of emergency everyone will be ready.

4. RESULTS
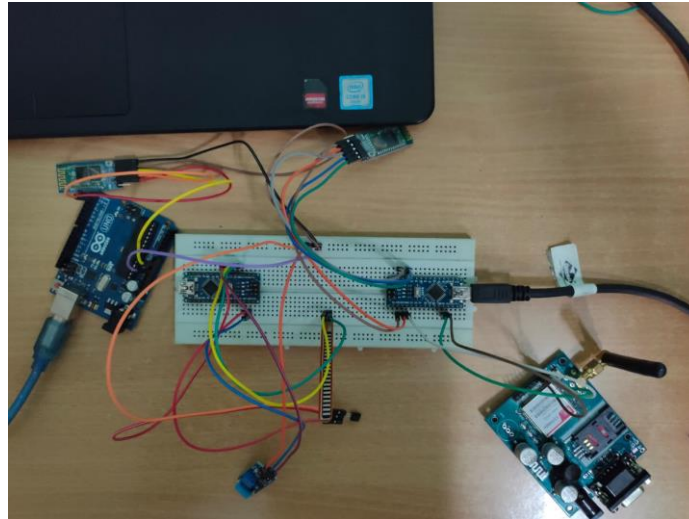
Fig.3 AT commands bluetooth master and slave
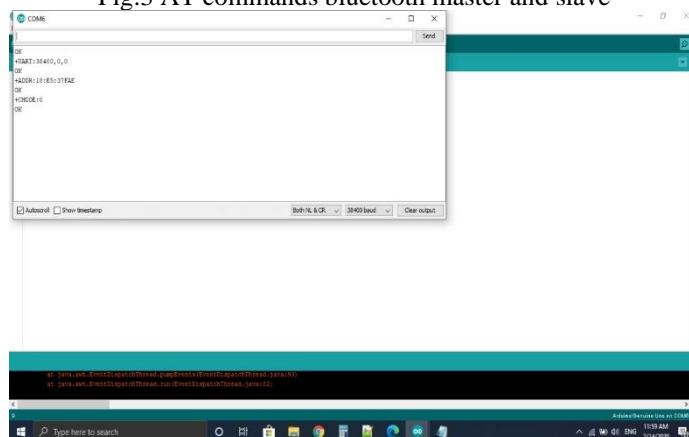
Fig.4 CMODE=0 means slave is active
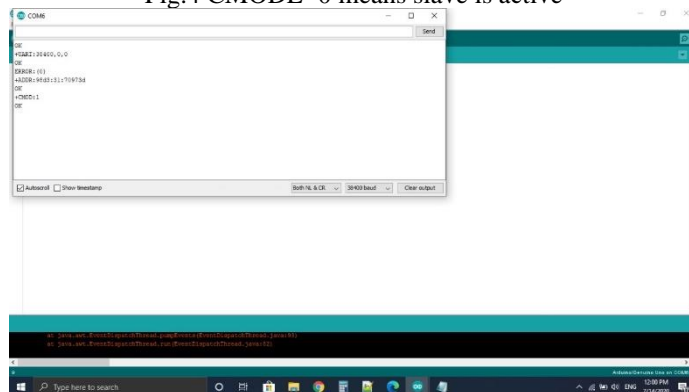
Fig.5 CMODE=1 means master is active
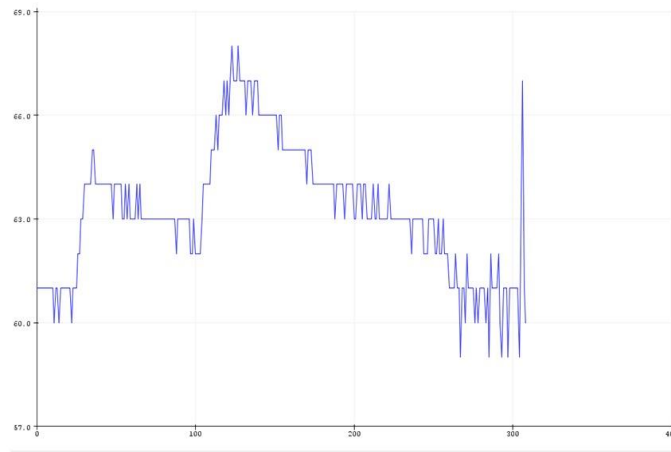
Fig.6 Temperature sensor graph

5. CONCLUSION

The Wireless Sensor Network for autonomous Structural Health check of a bridge is built and its functionality is verified in the simulation software. The network is built using Bluetooth protocol and the alerts are transmitted to the base-station using GSM module. This system can help in saving life and property by informing about the damages beforehand. The total system cost is very less, and the power consumption is also very less as the data is being transmitted only when there are damages detected.

6. FUTURE SCOPE

In future, all the smart phones will have BLE (Bluetooth Low Energy) chips inside them. The SHM systems can be built around this new technology which will lead to significant decrease in power consumption. It will also be compatible to send alerts to the commuters' devices.

The concept of 5G and smart cities which are believed to be a reality in future is also responsible for driving interest in this field. In a world where every "thing" will be connected to every other "thing", WSN will form an integral part of our lives. We can incorporate this technology in SHM throughout the world.

REFERENCES

1. Fatemeh Mansourkiaie, Loay Sabry Ismail, Tarek Mohamed Elfouly And Mohamed H. Ahmed, "Maximizing Lifetime in Wireless Sensor Network for Structural Health Monitoring with and Without Energy Harvesting", 2017, IEEE access, volume 5.
2. Pooja Krishnath Patil, Prof. Dr. S. R. Patil, "Review on Structural Health Monitoring System using WSNBridges"2017 International Conference on Electronics, Communication and Aerospace Technology, 978-1-5090-5686-6.
3. Chengyin Liu, Zhaoshuo Jiang, Fei Wang and Hue Chen, "Energy-Efficient Heterogeneous Wireless Sensor Deployment with Multiple Objectives for Structural Health Monitoring", MDPI Journal, November 2016.
4. N. Elamathi, R. Nusrath, S. Jayanthi, "An Energy Aware Fault Tolerance Scheme with Delay Reduction in WSN for Structural Health Monitoring" IJIRCCE, Vol. 4, Issue 5, May 2016.
5. Adam B. Noel, Abderrazak Abdaoui, Tarek Elfouly, Mohamed Hossam Ahmed, Ahmed Badawy, and Mohamed S. Shehata, "Structural Health Monitoring using Wireless Sensor Networks: A Comprehensive Survey", 2016 IEEE Translations and content mining are permitted for academic research.
6. Bruno Srbinovski, "Energy Aware Adaptive Sampling Algorithm for Energy Harvesting Wireless Sensor Networks", IEEE Instrumentation and Measurement Society, ©2015 IEEE.
7. Mohammed Najeeb A, and Vrinda Gupta, "Energy Efficient Sensor Placement for Monitoring Structural Health", Conference Proceedings Paper – Sensors and Applications, June 2014.
8. S. K. Ghosh, M. Suman, R. Datta, and P. K. Biswas, "Power efficient event detection scheme in wireless sensor networks for railway bridge health monitoring system," in 2014 IEEE International Conference on Advanced Networks and Telecommunications Systems (ANTS), 2014, pp.1-6: IEEE.

9. L. Dos Santos et al., "A localized algorithm for Structural Health Monitoring using wireless sensor networks," Information Fusion, vol. 15, pp. 114-129, 2014.
10. Sumudu Wijetunge, Upul Gunawardana, Ranjith Liyanapathirana "Wireless Sensor Networks for Structural Health Monitoring: Considerations for Communication Protocol Design 2010 17th International Conference on Telecommunications.

IMPLEMENTATION OF 3-DIMENSIONAL DISCRETE WAVELET TRANSFORMS (3D-DWT) USING LIFTING BASED ALGORITHM

Hamsavahini R[1]*,
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Mamatha K R[2]
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Abstract— **In this paper we implemented 3D discrete wavelet Transform(DWT) using lifting based algorithm .The lifting based DWT architecture has the advantage of low power computational complexities transforming the signals with extension and regular data flow. We also provide a survey on these architectures for both 1-D and 2-D DWT. 3-D DWT uses a cascade combination of three 1-D wavelet transforms.**

KEYWORDS: DWT, LIFTING, IMAGE COMPRESSION

I.INTRODUCTION

Wavelet transforms is one of the most modern areas of mathematics. In the last few decades, there has been an enormous increase in the applications of wavelets, in which it is used as an analytical tool in various areas of technical research such as electronics, computers and telecommunication systems that require efficient speed, resolution & real time memory and optimization with maximum hardware utilization. A wavelet, as the name suggests, is a little piece of a wave. They are used in representing data and other functions that satisfy certain mathematical requirements. The fundamental idea behind wavelets is to analyze according to scale. Wavelet algorithms process data at different scales or resolution. Wavelet transforms offer a wide variety of useful features in contrast to other transforms such as Fourier or cosine transform. Fourier transform is a powerful tool for analyzing the components of stationary signal, but it failed for analyzing non-stationary signal, whereas wavelet transforms allows the components of a non-stationary signal to be analyzed by using a set of damped oscillatory functions known as wavelet basis. Wavelet transforms in its continuous form is represented as CWT. A discrete and fast implementation of CWT is known as standard DWT. It decomposes the signal into different sub-bands with both time and frequency and facilitates to arrive at high compression ratio.

Recent advances in implementation of image coding and progressive image transmission through DWT demand both a large number of computations and large storage features that are not desirable for either high speed or low power applications. A methodology for implementation of the above complex applications have been proposed known as lifting based DWT that often has far fewer computations than DWT. In lifting scheme, the signal is divided like a zipper and a series of convolution-accumulate operations across the divided signals is applied, since this technique applies to each of this individual divided signals. Finally, brief summaries are given in below sections to conclude the paper.

II.WAVELETS

2.1. Wavelet Definition: A 'wavelet' is a small wave which has its energy concentrated in time. It has an oscillating wavelike characteristic and is a suitable tool for transient, non-stationary or time-varying phenomena.

2.2 Wavelet Characteristics: The difference between wave (sinusoids) and wavelet is shown in figure 1.1. Waves are smooth, predictable and everlasting, whereas wavelets are of limited duration, irregular and may be asymmetric. Waves are used as deterministic basis functions in Fourier analysis for the expansion of functions (signals), which are time-invariant, or stationary. The important characteristic of wavelets is that they can serve as deterministic or non-deterministic basis for generation and analysis of the most natural signals to provide better time-frequency representation, which is not possible with waves using conventional Fourier analysis

(a)   (b)

Fig.2. Representation of a (a) wave (b) wavelet

2.3 Wavelet Analysis: The wavelet analysis procedure is to adopt a wavelet prototype function, called an 'analyzing wavelet' or 'mother wavelet'. Temporal analysis is performed with a contracted, high frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low frequency version of the same wavelet. Mathematical formulation of signal expansion using wavelets gives Wavelet Transform (WT) pair, which is analogous to the Fourier Transform (FT) pair. Discrete-time and discrete-parameter version of WT is termed as Discrete Wavelet Transform (DWT).

2.4. Advantages of Wavelets:
*(a).*Wavelet analysis is an exciting new method for solving difficult problems in mathematics, physics, and engineering, with modern applications as diverse as wave propagation, data compression, signal processing, image processing, pattern recognition, computer graphics, the detection of aircraft and submarines and other medical image technology.
*(b).*Wavelets allow complex information such as music, speech, images and patterns to be decomposed into elementary forms at different positions and scales and subsequently reconstructed with high precision.
*(c).*Wavelets are a powerful statistical tool which can be used for a wide range of *applications*, namely: Signal processing, Data compression, Fingerprint verification, Blood-pressure, heart-rate and ECG analyses, DNA analysis, and protein analysis.

III.DISCRETE WAVELET TRANSFORM

3.1. Wavelet Transforms: Wavelet transform has gained widespread acceptance in speech, image and video processing, and in signal processing in general because of its attractive characteristics to represent non-stationary signals in both frequency and time domains. The wavelet transform is computed separately for different segments of the time-domain signal at different frequencies. *Multi-resolution analysis*: analyzes the signal at different frequencies giving different resolutions. Multi-resolution analysis is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Good for signal having high frequency components for short durations and low frequency components for long duration, e.g. Images and video frames.
3.2. Discrete Wavelet Transforms: It is a wavelet transform for which the wavelets are discretely sampled. DWT of a signal x(n) is calculated by passing it through a series of filters. First the samples are passed through a low-pass filters with impulse response g(n) resulting in a convolution of the two.

$$\mathbf{Y[n]=(x*g)[n]= \sum_{k=-\infty}^{\infty} x[k]g[n-k]} \quad (1)$$

The signal is also decomposed simultaneously using a high-pass filter h(n). The outputs giving the detail coefficients(from the high pass filter)and approximate coefficients(from

low pass filter) as shown in fig.3.These 2 filters are related to each other and they are known as a quadrature mirror filters. The frequency of the signal have now been removed, half the samples can be discarded according to Nyquist'srule. The filter output's are then sub sampled by 2(g-high pass and h-low pass).

$$\mathbf{Y_{low}[n]=\sum_{k=-\infty}^{\infty} x[k]g[2n-k]} \qquad (2)$$

$$\mathbf{Y_{high}[n]=\sum_{k=-\infty}^{\infty} x[k]h[2n-k]} \qquad (3)$$

This decomposition has halved the time resolution since only half of each filter output characterizes the signal.Each output has the frequency band of the input so the frequency resolution has been doubled.

Fig.3. One-Dimensional DWT

IV.TWO-DIMENSIONAL DWT

One-Dimensional DWT can be easily extended to two dimensions which can be used for the transformation of two dimensional images. A two dimensional digital image which can be represented by a 2-D array x[n1,n2] with n1 rows and n2 columns, where n1 and n2 are positive integers.First, a one-dimensional DWT is performed on rows to get low frequency 'L' and high frequency 'H' components of the image.The fundamental block diagram of a 2-D DWT is as shown in fig 4a.

Fig.4 Two-Dimensional DWT

At each scale, an image f(x,y) is decomposed into an approximation image a_j of a low-pass band, and three detail images d_j^x, d_j^y, d_j^{xy} corresponding to a horizontal high-pass

band d_j^x, a vertical high-pass band d_j^y, and a diagonal high-pass band d_j^{xy} as given below:

$$a_j(x,y) = \left\langle f(x,y), \phi_j(x)\phi_j(y) \right\rangle$$
$$d_j^x(x,y) = \left\langle f(x,y), \psi_j(x)\phi_j(y) \right\rangle$$
$$d_j^y(x,y) = \left\langle f(x,y), \phi_j(x)\psi_j(y) \right\rangle$$
$$d_j^{xy}(x,y) = \left\langle f(x,y), \psi_j(x)\psi_j(y) \right\rangle$$

Compare with single dimension case

$$Wf(b,s) = \int f(t)\frac{1}{\sqrt{s}}\psi\left(\frac{t-b}{s}\right)dt = f * \overline{\psi}_s(b)$$

$$= <f(t), \psi(t)> \tag{6}$$

Due to down-sampling of the coefficients, the size of the image at lower level is half of that of the higher level.In two-dimensional image analysis, the three wavelets are given by:

$$\psi^1(x,y) = \varphi(x)\psi(y),$$
$$\psi^2(x,y) = \psi(x)\varphi(y)$$
$$\psi^3(x,y) = \psi(x)\psi(y) \tag{7}$$

Extract image details at different scales and orientations. At each scale, we end up with three "detail" images:(a) Low-pass filtered in the x-direction and high-pass filtered in the y direction, $\psi^1(x,y) = \varphi(x)\psi(y)$, yielding detail D^x. (b)Low-pass filtered in the y direction and high-pass filtered in the x direction, $\psi^2(x,y) = \psi(x)\varphi(y)$ yielding detail D^y, and, (c)finally, high-pass filtered in both x and y directions yielding detail D^{xy}

Fig.5 Components of decomposed image, three details and one approximation at each scale.

So, we have three orientations for details: Horizontal, Vertical and Diagonal. To cover the entire image using one dimensional wavelets, we consider image composed of rows or columns which are one dimensional in rows or columns. We apply wavelet transform for each row (i.e. keeping x

constant but changing y) starting from the top row. Then apply wavelet transform on the results of row operations but now move column-wise starting from the left column where we use wavelet or scaling function depending on whether we want LL, LH, HL or HH signal components. Since we are using one dimensional wavelet transform at each scan of x or y direction, we have only two filters low pass 'L' and high pass 'H' where we use them in sequence in x and y directions. To obtain low frequency band in x and y direction, we need to use to low pass filters as shown in Fig.4c Similarly we use LH, HL and HH for other frequency bands.

Fig.6. Different levels of decomposition of an image

V.3-DIMENSIONAL DWT

The 3D DWT can be considered as a combination of three 1D DWT in the x, y and z directions as shown in the fig. 5. The preliminary work in the DWT processor design is to build 1D DWT modules, which are composed of high-pass and low-pass filters that perform a convolution of filter coefficients and input pixels. After a one-level of 3D discrete wavelet transform, the volume of image is decomposed into HHH,HHL,HLH,LHL,LHH,LLH and LLL signals as shown in fig.5.The process results in 8-data stream. The approximate signal resulting in scalar operations only goes to the next octave of the 3D transform. It has roughly 90 percent of the total energy. Meanwhile, the seven other streams contain the detail signals. Note that though conceptual drawing of the 3D DWT for one octave has 7 filter phases, this doesn't mean that the process needs 7 physical pairs. For example, a folded architecture maps multiple filters onto one filter pair.

Fig.7. 3-Dimensional DWT

VI.LIFTING BASED DWT

The development of the lifting scheme was proposed for the construction of bi-orthogonal wavelets. The main feature of the lifting scheme is that all constructions are derived in the spatial domain. It does not require complex mathematical calculations that are required in traditional methods. Lifting scheme is simplest and efficient algorithm to calculate wavelet transforms. It does not depend on Fourier transforms. Lifting scheme is used to generate second - generation wavelets, which are not necessarily translation and dilation of one particular function.

The memory modules are divided into multiple banks to accommodate high computational bandwidth requirements .a number of architectures have been proposed for calculation of the convolution based DWT. The architectures are mostly folded and can be broadly classified into series (where the inputs are supplied to the filters in a serial manner) and parallel architectures (where the inputs are supplied to the filters in a parallel manner). The basic principle of lifting scheme is to factorize the poly-phase matrix of a wavelet filter into a sequence of alternating upper and lower triangular matrices and diagonal matrix. This leads to the wavelet implementation by means of a banded-matrix multiplications.

Constructing wavelets using lifting scheme consists of three steps: (a)Split step: where the data is split-up into odd and even sets. (b)Predict step: in which odd set is predicted from even set. Predict phase ensures polynomial cancellation in high pass. (c)Update phase: that will update even set using wavelet coefficient to calculate scaling function. Update stage ensures preservation of moments in low pass. Block diagram of forward Lifting scheme is as shown in fig.6a.

6.1. Working of lifting scheme:

The basic idea behind lifting scheme is that, it tries to use the correlation in the data to remove redundancy. First split the data into 2 sets(split phase) i.e., odd sample and even sample. Because of the assumed smoothness of the data, we predict that odd samples have a value that is closely related to their neighboring even samples. We use N even samples to predict the value of a neighboring odd value (predict phase).With a good prediction method, the chance is high that the original odd sample is in the same range as its prediction. We calculate the difference between the odd sample and its prediction and replace the odd sample with this difference.

6.2. For Image Compression:

In each lifting stage, the predicting or updating operations are carried out in the direction of image edges and textures in a local window, and are not necessarily horizontal or vertical. This adaptation can significantly reduce the signal energy of high-pass sub-bands. High angular resolution in prediction is achieved by the use of fractional pixels in prediction and update operations. The fractional pixels can be calculated by any existing interpolation method. In order to guarantee perfect reconstruction, the predicted and updated samples are always in integer pixel positions.

Fig.8. Block diagram of forward Lifting scheme

Fig.9. Architecture of DWT based on lifting scheme

6.3.Reasons for the choice of Lifting scheme

We have used lifting scheme of wavelet transform for the digital speech compression, because lifting scheme is having following advantages over conventional wavelet transform technique. (1)It allows a faster implementation of the wavelet transform. It requires half number of computations as compare to traditional convolution based discrete wavelet transform. This is very attractive for real time low power applications. (2)The lifting scheme allows a fully in-place calculation of the wavelet transform. In other words, no auxiliary memory is needed and the original signal can be replaced with its wavelet transform. (3)Lifting scheme allows us to implement reversible integer wavelet transforms. In conventional scheme it involves floating point operations, which introduces rounding errors due to floating point arithmetic. While in case of lifting scheme perfect reconstruction is possible for loss-less compression.

VII. RESULTS

Consider an color RGB image of size 225x225 and convert it into a gray scale image which is then resized into 256x256 as shown in fig 10.

Fig 10. Resized Input Image

We extract the required wavelet co-efficients and combine it with corresponding filter coefficients. Divide the image coefficients into odd and even indexed values. Convolute filter coefficients with image coefficients taking two adjacent pixel values at a time. After performing predict and update step column wise first and row wise next we get 128x128 pixel image which is the output of 1D DWT as shown in fig b.

fig 11. Output Of 1D DWT

The output of 1D DWT is given as an input to the cascaded filters of 2D DWT where the image is compressed further that leads to 64x64 pixel image as shown in fig c.

Fig 12. Output of 2D DWT

 The output of 2D DWT is given as an input to the cascaded filters of 3D DWT. The final image is the compressed image(32x32) of the original input image due to normalization/down sampling which is shown in fig d.

Fig 13. Output of 3D DWT

VII.COMPARISIONS

At first, the comparision between different stages of DWT implemented in this paper using lifting based algorithm is tabulated below:

Levels of dwt	Extent of relative compression
1st level DWT	2:1
2nd level DWT	4:1
3rd level DWT	8:1

Next, the time difference in computation speed between normal dwt and discrete wavelet transforms implemented using lifting based algorithm is shown below in fig 14 and fig 15.

As we have already seen the advantages of using lifting based algorithm that it takes less time to compute than normal DWT. As we see in fig 14. that normal dwt takes a total time of 1.108s to process an image. Where as lifting based algorithm shown in fig 15. Takes the total time of 1.044s to process same image ,there is a time difference of 0.064s for just one image but in real world we deal with millions of images at a time where time constraint is more effective and hence lifting based occupies high priority.

Fig 14. Time simulation results of normal DWT

Fig 15. Time simulation results of lifting based DWT

CONCLUSION

In this paper, we implemented the lifting based implementations of 1-dimensional, 2-dimensional and 3-dimensional Discrete Wavelet Transform. We briefly described the principles behind the lifting scheme in-order to better understand the different implementation styles and structures.

The result of this implementation is the compressed image whose relative compression is 8 times the input image as shown in simulation figures.

REFERENCES

[1] Senthilkumar.M ,Uma .S " High speed 3d DWT VLSI architecture for image processing using lifting based wavelet transform".International journal of communication and computer technologies. Vol 01-no.36.

[2] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecturefor Image Processing".IMACST: vol 3.

[3] Kishore Andra, Chaitali Chakrabarti, member IEEE, and Tinku Acharya, sentor member IEEE" A VLSIfor lifting based forward and inverse wavelet transform".IEEE transactions on signal processing. Vol 50,no. 4.

[4] Malay Ranjan Tripathy, Kapil Sachdeva and Rachid Talhi "3D Discrete wavelet transform VLSI architecture for Image processing" progress in Electromagnetics research symposium proceedings Moscow, Russia.

[5]Patrick Longa " An Optimized architecture for 2D discrete wavelet transform on FPGA's usingdistributed arithmetics 2006.

[6] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," IEEE Signal Processing, vol. 8, issue: 4,pp. 14-38, Oct. 1991.

[7] ISO/IEC. International Standard, 15444-1: 2000(E), JPEG2000 Image Coding System – Part I Core coding system.

[8] C. Chakrabarti, M. Vishwanath, and R. M. Owens, "Architectures for wavelet transforms: A survey," J. VLSI Signal Process., vol. 14, pp.171–192, 1996.

[9] W. Sweldens, "The lifting scheme: A construction of second-generation wavelets," SIAM J. Mathematical Analysis, vol. 29, no.2, pp. 511–46, 1997.

[10] P. Chen, "VLSI implementation for one-dimensional multilevel lifting-based wavelet transform," IEEE Trans. on Computers, vol.53, no.4, pp.386-398, April 2004.

[11] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecture for Image Processing".IMACST: vol 3.

[12] N.D. Zervas, G.P. Anagnostopoulos, V. Spiliotopoulos,Y. Andreopoulos, and C.E. Goutis, "Evaluation of Design Alternatives for the 2-D Discrete Wavelet Transform," IEEE Trans. Circuits and Syst. Video Tech., Vol. 11, No. 12, pp. 1246-1262, December 2001.

[13] C. Chrysafis and A. Ortega, "Line-Based, Reduced Memory, Wavelet Image Compression," IEEE Trans.Circuits and Syst. Video Tech., Vol. 9, No. 3, pp.378-389, March 2000.

Gateway Based Hybrid Hierarchical Routing Scheme with Time Synchronization in WSN

Raghunandan G. H[1] and A Shobha Rani[2]

[1]*Department of Electronics and Telecommunication Engineering*
[2]*Department of Electronics and Communication Engineering*
[1,2]*BMS Institute of Technology and Management, Bengaluru, India*

ABSTRACT

Wireless Sensor Network are spatially distributed sensors intended to monitor different physiological conditions. Sensing and communicating data from one place to another consumes more energy, therefore the management of sensor energy is very important factor. Energy utilization, synchronization and lifetime of the network is the main criteria in WSN. More energy is lost by sensors which are far from the base station.The cluster head is deployed to collect and relay information from nodes to the base station or gateway nodes to resolve this problem. To decrease energy consumption, gateway nodes are deployed between the cluster head and the base station. In this paper, a hybrid approach is used to increase the overall efficiency of the network in WSNs with time synchronization which increases the throughput of the network. The efficiency in terms of network lifetime, residual energy, data packets, throughput of the network has been improved as shown in simulation results. The performance of WSN of the proposed scheme is compared to other classical routing scheme and proposed algorithm has proved its merit.

KEY WORDS: BASE STATION (BS), CLUSTER HEAD (CH), GATEWAY NODE (GN), ROUTING PROTOCOLS, WIRELESS SENSOR NETWORK (WSN).

INTRODUCTION

Presently there have been development in Micro-Electro-Mechanical Systems (MEMS) in tandem with major developments in digital signal processing (DSP) which has. led to growth of micro-sensors. Previously few industries use wired sensors, implementation provides deployment of sensor nodes more viable than before. Previously, there has been study regarding applications of WSN such as environmental monitoring, agricultural field, military surveillance and home automation (Pavithra et al., 2019).

Deep research on routing protocols efficiency was done considering the power constraints in WSNs and there was vast deployment of nodes on large-scale , to ensure reliable and real-time data transmission. Recently there has been exposure in the field of WSNs and their applications because they are easy to deploy and are of low cost, have flexibility. a WSN have distinctive set of resource curtailment like finite on-board battery power, limited processing ability and limited communication bandwidth. Since sensors are battery-powered, energy efficiency is of vital importance in WSNs. Algorithms are used to solve the problem of power constraint without altering the standard. Local collaboration among sensors, suppression, data compression, redundant data ,avoidance of direct transmission to far distant sensors are of the major factors that influence algorithm designers to device unique distributed, scalable and energy efficient solution for Wireless Sensor Networks.

.In common, the sensor nodes measure environmental conditions,. The sensor node extract some useful information by processing the raw sensor signals . The output of this processed signal is transmitted the through direct communication or multi-hop communication with access point across other sensor nodes. In some situations, repeaters (RPs) are used for multi-hops, to support sensors installed outside the radio range. One of the components of the WSN is the base stations which has more energy, computational ,communication resources. Forwarding of data from wireless sensor network on to a server is done by the BS which acts as a gateway. Energy is a limited resource of WSN , and it determines the lifetime of WSNs. The computation subsystem has less energy consumption when compared to the communication subsystem has. The energy required for transmitting one bit may consume as same as executing a few thousands instructions. Hence, communication must be traded for computation. In different environments, together with remote and hostile regions, where ad-hoc communications are a key element WSNs is to be deployed in large numbers .For this reason, algorithms and protocols should concentrate on the subsequent issues:

- Lifetime maximization: Sensor nodes should be energy efficient and consumption of energy of the device should be less due to the limited energy resources. The radio power supply when not in use should shut off to conserve power of the node
- fault tolerance and Robustness .
- Self-configuration.

Figure 1 shows the arrangement of WSN.

Figure 1: Schematic of WSN

Hierarchical routing algorithms are more efficient algorithm. In this, CHs are responsible for collecting the data and send it to GN and then the GN sends the information to the BS. If the members of the cluster are far, then there is a wastage of energy. To solve this problem, the Centrality approach is used. In this method, the distance between one node to every other node is calculated using the Pythagoras theorem. The node which is equidistance to all other nodes along with high energy will be selected as the cluster head. This will reduce energy consumption. Error during communication can be reduced by using time synchronization in communication. This will reduce the packet collision during data transmission. In this paper, a new synchronization method considering a new broadcast sequence is used. This determines the order of the nodes that transmit timing information, this Eliminates collision. Section II presents a summary of the existing systems. The framework of proposed system is explained in Section III . The system is evaluated in Section IV based on simulation and implementation. Section V, addresses the future scope and demands. The paper concludes with section VI.

Existing Systems: A wireless sensor network consists of distributed sensor nodes. Sensor nodes include of sensing unit based on application, processing levels. In the existing system, many limitations or difficulties cause less efficiency of the system. It includes Energy limitations storage or environmental limitations, communication constraints, and other limitations. Transfer of information to nodes needs more energy. Hence sensor nodes are affected by energy limitations. Sensor nodes are small devices and hence their capacity to store the data is less. Nodes of the network may get affected by many environmental conditions physical obstacles, unpredictable errors, and also communication interferences.

Communication constraints are limitation of bandwidth, frequent routing changes, channel error rates, and also unreliable communication. LEACH (Salah et al., 2016) is the main protocol under hierarchical routing protocols which includes set-up phase, and steady-state phase. In the first phase some nodes are selected at random probability as cluster heads (CHs) and nodes are ordered into clusters. In the second phase, the data is transmitted to the BS. CHS has to lose more energy when compared to other normal nodes. A drawback of this system is cluster head consumes more energy and sensor nodes die faster.

In (Djenouri et al., 2012) they proposed a protocol that mainly deals with the technique of the cluster head (CH) selection that makes balanced energy consumption among the sensors, and which leads to enhancement of lifetime of the sensor network. In (Sang et al., 2012) for transmission of data to the Base Station majority of the energy of sensor nodes is used. Thus, there is a fast depletion of energy. Here agglomerative to limit the energy utilization of cluster heads a portable base station is utilized along with cluster approach. But the movement of the Base Station is not always feasible. In (Ruqiang et al., 2014) this it discusses up several technical challenges and many application possibilities that occur when the sensor networks interconnect several nodes when wide networks are established. These wireless sensor networks communicate using multiple -hop wireless communications systems.

To (Raghunandan et al., 2011) guarantee reliable multi-hop communication and to maintain the routes in the network routing protocols for WSN are used. It provides us with an idea of routing protocols for WSN and compares their relative strengths and limitations to provide better energy efficiency or increase wireless. In (Raghunandan et al., 2012) mainly concentrates on synchronization schemes where a multichip extension is used which use a final local estimates, without any forwarding of synchronization signals. In the existing system, there are many routing challenges and design issues. In designing routing protocols one should consider the uniform distribution of sensor nodes; otherwise, we should go for clustering. Production costs also should be less and the quality of service should be good.

Proposed System: In a hierarchical approach, some sensors are classified as lower energy nodes and some sensors are grouped as higher energy nodes. The higher energy nodes become cluster heads (CHs) and lower energy nodes become the normal sensors that perform the ensuing operation only.

CH is decided with a different probability (D Djenouri et al., 2012) by following equation.

$$T(n) = [p/(1-p)] \times (r \bmod p-1) \qquad (1)$$

Figure 2: Concept of proposed techniques

But the cluster heads deal with the collection of information from sensors, data aggregation of these data and transmission of processed data to the next level. Figure 2 shows the concept of the proposed technique. The main aims of this protocol are to reduce power consumption, data aggregation, and time synchronization. The main aim of the proposed technique is to increase energy efficiency, improving the time synchronization in communication between each node and the receiver, and enhance the network lifetime of sensor nodes. This paper is mainly motivated to overcome the problem of higher energy consumption of sensor nodes and extending the lifetime of wireless sensor networks to provide good transmission and data sense. Since Sensor nodes away

from base stations hence they consume more energy during transmission of data to a base station. Hence gateway nodes are introduced. Gateway nodes receive the data from sensor nodes and then send them to the base station. Time synchronization is an elementary part of any network-oriented organization and system. In our proposed system grouping of sensor nodes is done called clustering. The energy consumed by the transmitter for transmission of message to a distance is given by,

$$E_T(k,r) = \begin{cases} k\left(E_{TX} + E_{fs} * r^2\right) & \text{if } r < r_o, \\ k\left(E_{TX} + E_{mp} * r^4\right) & \text{if } r \geq r_o. \end{cases} \qquad (2)$$

The energy dissipated by a receiver to receive bit message is given by

$$E_R(k) = k * E_{RX}, \qquad (3)$$

Threshold distance is given by,

$$r_o = \sqrt{\frac{E_{fs}'}{E_{mp}}}. \qquad (4)$$

Clustering is done to construct the appropriate topology of the network. The clustering-based network reduces the cost of the routing algorithm and the flooding broadcast. In this cluster-based routing approach, sensor nodes that are in a particular radio range are grouped which forms a cluster. Each group or cluster has one cluster head (CH). This CH collects all the data from sensor nodes in a cluster and it performs data fusion and sends it to the nearby gateway node. CH is selected based on the centrality approach. In this method distance of nodes between each other is determined; the node which is equidistant from the other node and the node having more energy is selected as CH. After CH selection aggregation of data is performed by removing redundant data. Data aggregation is the process of removing redundant data during transmission. This will increases the lifetime of the network.

In the proposed system base station (BS) is fixed. The same algorithm can also be applied to a movable base station. During the data communication phase, the data packets are transmitted from CH to gateway node and from gateway node to BS. This process is called multi-hop communication. The proposed technique uses the receiver-to-receiver concept implemented by the reference broadcast Synchronization which minimizes the time-critical path when compared to the sender-to-receiver method.

Implementation: The proposed system is simulated using MATLAB software. We consider 100 sq.m area and 100 nodes are distributed randomly. MATLAB Simulation is done for 4500 rounds. The proposed algorithm is compared with DR-LEACH. Fig.3 Initial network topology for LEACH. The figure 3 shows deployment of sensor and clustering of the network area in DR-LEACH.

Figure 3: Initial network topology for LEACH

Figure 4: Network framework for Proposed Technique

Figure 5: Comparision of packets sent of Proposed System vs. DR-LEACH in multiples of 10^4

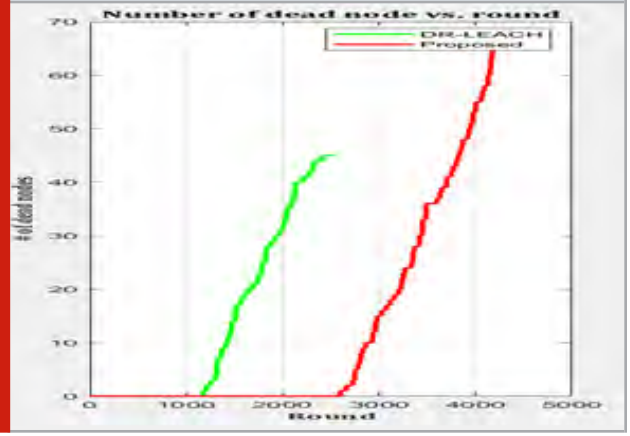| Figure 6: Comparison of number of dead sensors of Proposed System vs DR-LEACH |
| --- |

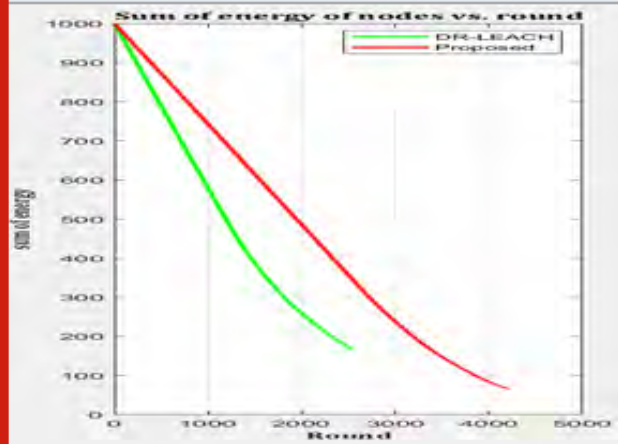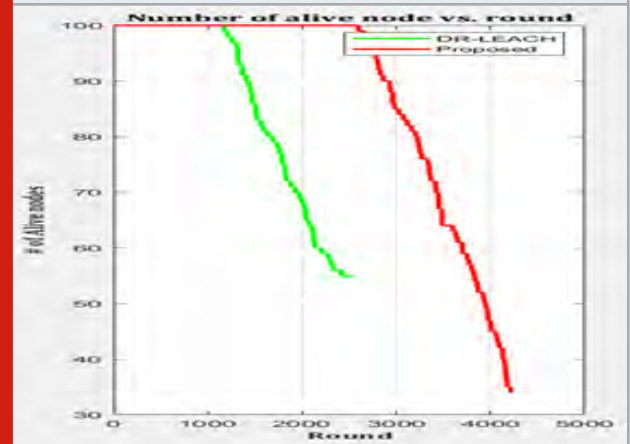| Figure 7: Comparison of Residual energy of results of Proposed System vs DR-LEACH |
| --- |

Figure 8: Comparison of Network Lifetime of Proposed System vs DR-LEACH

212

The connectivity between individuals to CH, CH to the gateway node, and GN to Base station are shown in figure 4 of proposed system. Figure 4 shows network framework of Proposed Technique. Figure 5 packets sent of Proposed System vs. DR-LEACH. Figure 6 shows the Comparison of number of dead sensors of Proposed System vs DR-LEACH. Figure 7 Comparison of Residual energy of results of Proposed System vs DR-LEACH. The implementation results in a hybrid model comprising of energy conservation, data aggregation, and time synchronization. The results in the graph show improvements in the lifetime of the network along with time synchronization. With respect to simulation results numbers of dead nodes are less on comparison with rounds of DR-LEACH protocol. This shows that the lifetime of network using proposed system is enhanced compared to DR-LEACH which leads to maximum transmission of data. Hence the WSN using the proposed technique works with more powerfully on comparison to the DR-LEACH.

CONCLUSION

Energy consumption by sensors is a major parameter for network lifetime in WSN. In hierarchical clustering methodology, the cluster head selection plays an important role. When the complete network is considered, then the optimization of energy consumption can be done by the changing of the cluster head based on priorities at the cluster level. In this proposed system, we establish a connection from sensor nodes to CH, CH to GN, and from the gateway node to the base station. The energy utilization can be decreased by properly designing the CH selection mechanism.

The selection of cluster head proposed in this paper is a new technique discussed concerning distance from the base station and centrality approach at the cluster level. Along with that time synchronization is achieved using reference broadcast synchronization. The implemented protocol shows an improvement in the lifetime network, and the life of nodes will be extended to the maximum period. Here we are comparing the proposed system with the DR-LEACH protocol and it proves that the proposed system is more efficient than DR-LEACH.

REFERENCES

Abhishek , Sumedha, (2014) ,Minimization of Avg Energy Consumption to Prolong Lifetime of WSN, IEEE Global Conference on WCN.

Akkaya, Younis,(2005) ,A Survey of Routing Protocols in WSN, Journal of Ad Hoc Network ,Elsevier, vol. 3, pp. 325-349.

Amit , R Singh, A Nandini(2016) WSN – Challenges and Possibilities, IJCA, volume 140-No.4

D Djenouri(2012), R4 Time Synchronization in WSN, IEEE signal processing letters, Vol. 19, No. 4.

Salah, A. Boulouz(2016) ,Energy Efficient Clustering based on LEACH, International Conference on Engineering and MIS.

Moumita, Paramita , Sarmistha (2014),Analysis of EE WSN Routing Schemes, International Conference of EAIT.

Pavithra, G.S. and Babu, N.V., 2019. Energy efficient hierarchical clustering using HACOPSO in wireless sensor networks. International Journal of Innovative Technology and Exploring Engineering, 8(12).

Raghunandan, Shobha Rani, Nanditha (2014), Hierarchical clustering based routing Technique for overall efficiency of WSN, In Proceeding of the IEEE 2017 International Conference on RTECIT.

Raghunandan, Shobha.Rani, Nanditha, et al.,(2017),Hierarchical agglomerative clustering based routing algorithm for overall efficiency of WSN, In Proceeding of IEEE ICICT -2017.

Raghunandan, Lakshmi(2011) ,Secure Routing protocols for WSN, International Conference on FCS proceedings.

Raghunandan, Sagar (2012) ,A Novel Approach to increase overall Efficiency in WSN ,2012 IEEE ICCEET proceedings, pp 699-703.

Raghunandan, Lakshmi.(2011) ,A Comparative Analysis of Routing Techniques for WSN, IEEE NCOIET-2011, IEEE Conference Proceedings.

Ruqiang, Sun,Yuning(2013) ,EAS Node Design with Its Application in WSN, IEEE Transactions on instrumentation and measurement, Vol. 62.

Sang, Nguyen (2012) ,Distance Based Thresholds for CH selection in WSN, IEEE Communications Letters ,Vol 16 , Issue 9 .

Ying, Chin , Ting (2008)" AER Protocol for mobile sink in WSN," IEEE Conference on Media Computing, conference proceeding, pp. 44-49.

Yunhe, Zhihua,Q Zhang(2016) "Efficient Load Balance Data Aggregation methods for WSN based on compressive network coding ", IEEE International Conference on EICT.

BBRC
Bioscience Biotechnology
Research Communications

Gateway Based Hybrid Hierarchical Routing Scheme with Time Synchronization in WSN

Raghunandan G. H[1] and A Shobha Rani[2]
[1]*Department of Electronics and Telecommunication Engineering*
[2]*Department of Electronics and Communication Engineering*
[1,2]*BMS Institute of Technology and Management, Bengaluru, India*

ABSTRACT

Wireless Sensor Network are spatially distributed sensors intended to monitor different physiological conditions. Sensing and communicating data from one place to another consumes more energy, therefore the management of sensor energy is very important factor. Energy utilization, synchronization and lifetime of the network is the main criteria in WSN. More energy is lost by sensors which are far from the base station.The cluster head is deployed to collect and relay information from nodes to the base station or gateway nodes to resolve this problem. To decrease energy consumption, gateway nodes are deployed between the cluster head and the base station. In this paper, a hybrid approach is used to increase the overall efficiency of the network in WSNs with time synchronization which increases the throughput of the network. The efficiency in terms of network lifetime, residual energy, data packets, throughput of the network has been improved as shown in simulation results. The performance of WSN of the proposed scheme is compared to other classical routing scheme and proposed algorithm has proved its merit.

KEY WORDS: BASE STATION (BS), CLUSTER HEAD (CH), GATEWAY NODE (GN), ROUTING PROTOCOLS, WIRELESS SENSOR NETWORK (WSN).

INTRODUCTION

Presently there have been development in Micro-Electro-Mechanical Systems (MEMS) in tandem with major developments in digital signal processing (DSP) which has. led to growth of micro-sensors. Previously few industries use wired sensors, implementation provides deployment of sensor nodes more viable than before. Previously, there has been study regarding applications of WSN such as environmental monitoring, agricultural field, military surveillance and home automation (Pavithra et al., 2019).

Deep research on routing protocols efficiency was done considering the power constraints in WSNs and there was vast deployment of nodes on large-scale , to ensure reliable and real-time data transmission. Recently there has been exposure in the field of WSNs and their applications because they are easy to deploy and are of low cost, have flexibility. a WSN have distinctive set of resource curtailment like finite on-board battery power, limited processing ability and limited communication bandwidth. Since sensors are battery-powered, energy efficiency is of vital importance in WSNs. Algorithms are used to solve the problem of power constraint without altering the standard. Local collaboration among sensors, suppression, data compression, redundant data ,avoidance of direct transmission to far distant sensors are of the major factors that influence algorithm designers to device unique distributed, scalable and energy efficient solution for Wireless Sensor Networks.

209

.In common, the sensor nodes measure environmental conditions,. The sensor node extract some useful information by processing the raw sensor signals . The output of this processed signal is transmitted the through direct communication or multi-hop communication with access point across other sensor nodes. In some situations, repeaters (RPs) are used for multi-hops, to support sensors installed outside the radio range. One of the components of the WSN is the base stations which has more energy, computational ,communication resources. Forwarding of data from wireless sensor network on to a server is done by the BS which acts as a gateway. Energy is a limited resource of WSN , and it determines the lifetime of WSNs. The computation subsystem has less energy consumption when compared to the communication subsystem has. The energy required for transmitting one bit may consume as same as executing a few thousands instructions. Hence, communication must be traded for computation. In different environments, together with remote and hostile regions, where ad-hoc communications are a key element WSNs is to be deployed in large numbers .For this reason, algorithms and protocols should concentrate on the subsequent issues:

- Lifetime maximization: Sensor nodes should be energy efficient and consumption of energy of the device should be less due to the limited energy resources. The radio power supply when not in use should shut off to conserve power of the node
- fault tolerance and Robustness .
- Self-configuration.

Figure 1 shows the arrangement of WSN.

Figure 1: Schematic of WSN

Hierarchical routing algorithms are more efficient algorithm. In this, CHs are responsible for collecting the data and send it to GN and then the GN sends the information to the BS. If the members of the cluster are far, then there is a wastage of energy. To solve this problem, the Centrality approach is used. In this method, the distance between one node to every other node is calculated using the Pythagoras theorem. The node which is equidistance to all other nodes along with high energy will be selected as the cluster head. This will reduce energy consumption. Error during communication can be reduced by using time synchronization in communication. This will reduce the packet collision during data transmission. In this paper, a new synchronization method considering a new broadcast sequence is used. This determines the order of the nodes that transmit timing information, this Eliminates collision. Section II presents a summary of the existing systems. The framework of proposed system is explained in Section III . The system is evaluated in Section IV based on simulation and implementation. Section V, addresses the future scope and demands. The paper concludes with section VI.

Existing Systems: A wireless sensor network consists of distributed sensor nodes. Sensor nodes include of sensing unit based on application, processing levels. In the existing system, many limitations or difficulties cause less efficiency of the system. It includes Energy limitations storage or environmental limitations, communication constraints, and other limitations. Transfer of information to nodes needs more energy. Hence sensor nodes are affected by energy limitations. Sensor nodes are small devices and hence their capacity to store the data is less. Nodes of the network may get affected by many environmental conditions physical obstacles, unpredictable errors, and also communication interferences.

Communication constraints are limitation of bandwidth, frequent routing changes, channel error rates, and also unreliable communication. LEACH (Salah et al., 2016) is the main protocol under hierarchical routing protocols which includes set-up phase, and steady-state phase. In the first phase some nodes are selected at random probability as cluster heads (CHs) and nodes are ordered into clusters. In the second phase, the data is transmitted to the BS. CHS has to lose more energy when compared to other normal nodes. A drawback of this system is cluster head consumes more energy and sensor nodes die faster.

In (Djenouri et al., 2012) they proposed a protocol that mainly deals with the technique of the cluster head (CH) selection that makes balanced energy consumption among the sensors, and which leads to enhancement of lifetime of the sensor network. In (Sang et al., 2012) for transmission of data to the Base Station majority of the energy of sensor nodes is used. Thus, there is a fast depletion of energy. Here agglomerative to limit the energy utilization of cluster heads a portable base station is utilized along with cluster approach. But the movement of the Base Station is not always feasible. In (Ruqiang et al., 2014) this it discusses up several technical challenges and many application possibilities that occur when the sensor networks interconnect several nodes when wide networks are established. These wireless sensor networks communicate using multiple -hop wireless communications systems.

To (Raghunandan et al., 2011) guarantee reliable multi-hop communication and to maintain the routes in the network routing protocols for WSN are used. It provides us with an idea of routing protocols for WSN and compares their relative strengths and limitations to provide better energy efficiency or increase wireless. In (Raghunandan et al., 2012) mainly concentrates on synchronization schemes where a multichip extension is used which use a final local estimates, without any forwarding of synchronization signals. In the existing system, there are many routing challenges and design issues. In designing routing protocols one should consider the uniform distribution of sensor nodes; otherwise, we should go for clustering. Production costs also should be less and the quality of service should be good.

Proposed System: In a hierarchical approach, some sensors are classified as lower energy nodes and some sensors are grouped as higher energy nodes. The higher energy nodes become cluster heads (CHs) and lower energy nodes become the normal sensors that perform the ensuing operation only.

CH is decided with a different probability (D Djenouri et al., 2012) by following equation.

$$T(n) = [p/(1-p)] \times (r \bmod p-1) \qquad (1)$$

Figure 2: Concept of proposed techniques

But the cluster heads deal with the collection of information from sensors, data aggregation of these data and transmission of processed data to the next level. Figure 2 shows the concept of the proposed technique. The main aims of this protocol are to reduce power consumption, data aggregation, and time synchronization. The main aim of the proposed technique is to increase energy efficiency, improving the time synchronization in communication between each node and the receiver, and enhance the network lifetime of sensor nodes. This paper is mainly motivated to overcome the problem of higher energy consumption of sensor nodes and extending the lifetime of wireless sensor networks to provide good transmission and data sense. Since Sensor nodes away

from base stations hence they consume more energy during transmission of data to a base station. Hence gateway nodes are introduced. Gateway nodes receive the data from sensor nodes and then send them to the base station. Time synchronization is an elementary part of any network-oriented organization and system. In our proposed system grouping of sensor nodes is done called clustering. The energy consumed by the transmitter for transmission of message to a distance is given by,

$$E_T(k,r) = \begin{cases} k\left(E_{TX} + E_{fs} * r^2\right) & \text{if } r < r_o, \\ k\left(E_{TX} + E_{mp} * r^4\right) & \text{if } r \geq r_o. \end{cases} \qquad (2)$$

The energy dissipated by a receiver to receive bit message is given by

$$E_R(k) = k * E_{RX}, \qquad (3)$$

Threshold distance is given by,

$$r_o = \sqrt{\frac{E'_{fs}}{E_{mp}}}. \qquad (4)$$

Clustering is done to construct the appropriate topology of the network. The clustering-based network reduces the cost of the routing algorithm and the flooding broadcast. In this cluster-based routing approach, sensor nodes that are in a particular radio range are grouped which forms a cluster. Each group or cluster has one cluster head (CH). This CH collects all the data from sensor nodes in a cluster and it performs data fusion and sends it to the nearby gateway node. CH is selected based on the centrality approach. In this method distance of nodes between each other is determined; the node which is equidistant from the other node and the node having more energy is selected as CH. After CH selection aggregation of data is performed by removing redundant data. Data aggregation is the process of removing redundant data during transmission. This will increases the lifetime of the network.

In the proposed system base station (BS) is fixed. The same algorithm can also be applied to a movable base station. During the data communication phase, the data packets are transmitted from CH to gateway node and from gateway node to BS. This process is called multi-hop communication. The proposed technique uses the receiver-to-receiver concept implemented by the reference broadcast Synchronization which minimizes the time-critical path when compared to the sender-to-receiver method.

Implementation: The proposed system is simulated using MATLAB software. We consider 100 sq.m area and 100 nodes are distributed randomly. MATLAB Simulation is done for 4500 rounds. The proposed algorithm is compared with DR-LEACH. Fig.3 Initial network topology for LEACH. The figure 3 shows deployment of sensor and clustering of the network area in DR-LEACH.

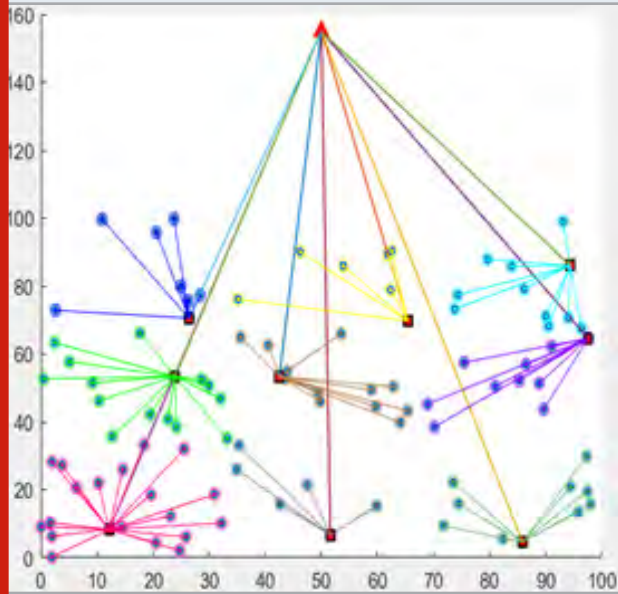Figure 3: Initial network topology for LEACH

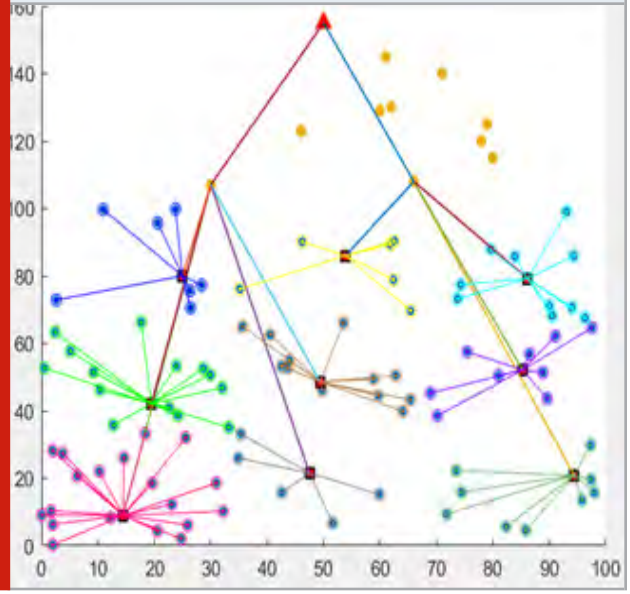Figure 4: Network framework for Proposed Technique

Figure 5: Comparision of packets sent of Proposed System vs. DR-LEACH in multiples of 10^4
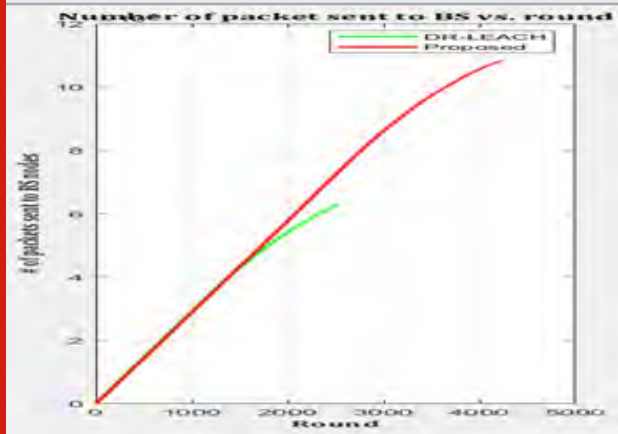

Figure 6: Comparison of number of dead sensors of Proposed System vs DR-LEACH
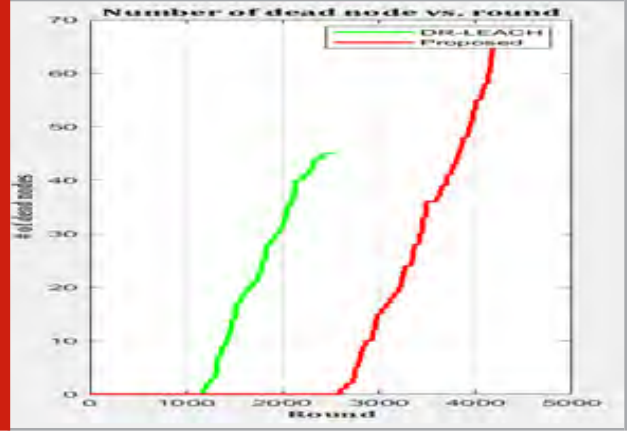

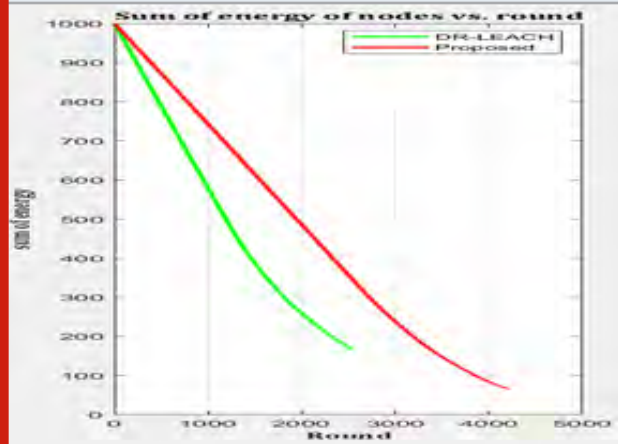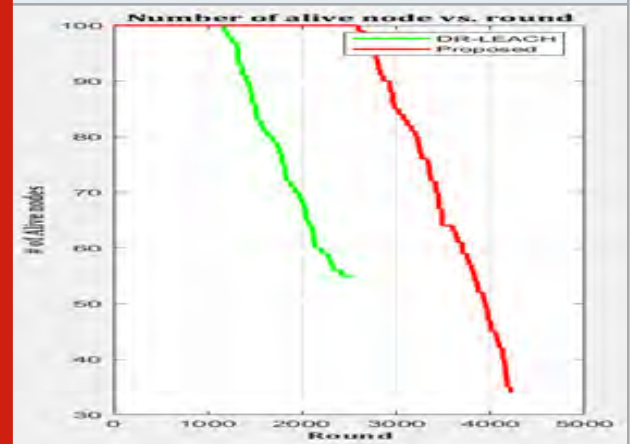Figure 7: Comparison of Residual energy of results of Proposed System vs DR-LEACH

Figure 8: Comparison of Network Lifetime of Proposed System vs DR-LEACH

The connectivity between individuals to CH, CH to the gateway node, and GN to Base station are shown in figure 4 of proposed system. Figure 4 shows network framework of Proposed Technique. Figure 5 packets sent of Proposed System vs. DR-LEACH. Figure 6 shows the Comparison of number of dead sensors of Proposed System vs DR-LEACH. Figure 7 Comparison of Residual energy of results of Proposed System vs DR-LEACH. The implementation results in a hybrid model comprising of energy conservation, data aggregation, and time synchronization. The results in the graph show improvements in the lifetime of the network along with time synchronization. With respect to simulation results numbers of dead nodes are less on comparison with rounds of DR-LEACH protocol. This shows that the lifetime of network using proposed system is enhanced compared to DR-LEACH which leads to maximum transmission of data. Hence the WSN using the proposed technique works with more powerfully on comparison to the DR-LEACH.

CONCLUSION

Energy consumption by sensors is a major parameter for network lifetime in WSN. In hierarchical clustering methodology, the cluster head selection plays an important role. When the complete network is considered, then the optimization of energy consumption can be done by the changing of the cluster head based on priorities at the cluster level. In this proposed system, we establish a connection from sensor nodes to CH, CH to GN, and from the gateway node to the base station. The energy utilization can be decreased by properly designing the CH selection mechanism.

The selection of cluster head proposed in this paper is a new technique discussed concerning distance from the base station and centrality approach at the cluster level. Along with that time synchronization is achieved using reference broadcast synchronization. The implemented protocol shows an improvement in the lifetime network, and the life of nodes will be extended to the maximum period. Here we are comparing the proposed system with the DR-LEACH protocol and it proves that the proposed system is more efficient than DR-LEACH.

REFERENCES

Abhishek , Sumedha, (2014) ,Minimization of Avg Energy Consumption to Prolong Lifetime of WSN, IEEE Global Conference on WCN.

Akkaya, Younis,(2005) ,A Survey of Routing Protocols in WSN, Journal of Ad Hoc Network ,Elsevier, vol. 3, pp. 325-349.

Amit , R Singh, A Nandini(2016) WSN – Challenges and Possibilities, IJCA, volume 140-No.4

D Djenouri(2012), R4 Time Synchronization in WSN, IEEE signal processing letters, Vol. 19, No. 4.

Salah, A. Boulouz(2016) ,Energy Efficient Clustering based on LEACH, International Conference on Engineering and MIS.

Moumita, Paramita , Sarmistha (2014),Analysis of EE WSN Routing Schemes, International Conference of EAIT.

Pavithra, G.S. and Babu, N.V., 2019. Energy efficient hierarchical clustering using HACOPSO in wireless sensor networks. International Journal of Innovative Technology and Exploring Engineering, 8(12).

Raghunandan, Shobha Rani, Nanditha (2014), Hierarchical clustering based routing Technique for overall efficiency of WSN, In Proceeding of the IEEE 2017 International Conference on RTECIT.

Raghunandan, Shobha.Rani, Nanditha, et al.,(2017),Hierarchical agglomerative clustering based routing algorithm for overall efficiency of WSN, In Proceeding of IEEE ICICT -2017.

Raghunandan, Lakshmi(2011) ,Secure Routing protocols for WSN, International Conference on FCS proceedings.

Raghunandan, Sagar (2012) ,A Novel Approach to increase overall Efficiency in WSN ,2012 IEEE ICCEET proceedings, pp 699-703.

Raghunandan, Lakshmi.(2011) ,A Comparative Analysis of Routing Techniques for WSN, IEEE NCOIET-2011, IEEE Conference Proceedings.

Ruqiang, Sun,Yuning(2013) ,EAS Node Design with Its Application in WSN, IEEE Transactions on instrumentation and measurement, Vol. 62.

Sang, Nguyen (2012) ,Distance Based Thresholds for CH selection in WSN, IEEE Communications Letters ,Vol 16 , Issue 9 .

Ying, Chin , Ting (2008)" AER Protocol for mobile sink in WSN," IEEE Conference on Media Computing, conference proceeding, pp. 44-49.

Yunhe, Zhihua,Q Zhang(2016) "Efficient Load Balance Data Aggregation methods for WSN based on compressive network coding ", IEEE International Conference on EICT.

FinFET based design and Performance Analysis of Nano-Processor for low area, low power and minimum delay using 32nm

Anil Kumar. D,
Department of Electronics and Communication Engg,
BMS Institute of Technology and Management, Bangalore -560064, India
anilkumard81@bmsit.in

Abstract— The recent technologies in VLSI chips has grown in terms of scaling of transistor and device parameters but still there is a challenging task for controlling of current between source and drain terminals. For effective control of device current, the FinFET transistors have come into VLSI chip manufacturing, through which current can be effectively controlled. This section addresses the issues present in CMOS technology and majorly concentrated on proposed 4-bit Nano processor using FinFET 32nm technology by using Cadence Virtuoso software tool. In the proposed Nanoprocessor design, the first portion of the design is done using 4bit ALU which includes all basic and universal gates, high speed adder, multiplier and multiplexer. The Carry Save Adder (CSA) and multiplier are the major sub component which can optimize the power consumption and area reduction. The second portion of the proposed Nano processor design is 4-bit 6T SRAM and encoder and decoder and also using Artificial Neural Network (ANN). All these sub components are designed at analog transistors (Schematic level) through which the Graphic Data System (GDS-II) is generated through mask layout design. Finally, the verification and validation are done using DRC and LVS and at the last chip level circuit is generated for chip fabrication. The ALU is designed by using CMOS inverters and the designed ALU schematic is simulated through 32nm FinFET using technological library and compared with CMOS technology which is simulated through 32nm CMOS library (without FinFET). The power consumption of AND, OR, XOR, NOT, NAND gates, SRAM, Encoder, Decoder and ANN are 36.09nW, 64.970nW, 61.13nW, 33.31nW, 37.45nW, 32.5% with optimization in power dissipation of 47% along with optimization in leakage current, with 2.68uW, 1.98uW and 7.5% improvement in power consumption and 0.5%

information loses are compressed subsequently respectively. The basic gates, universal gates, CSA, subtraction and MUX are integrated for 4-bit ALU design and its delay, power consumption and area are found to be 0.104nsec, 314.4uW and 56.8μsqm respectively.

Keywords—About four key words or phrases in alphabetical order, separated by commas.

I. INTRODUCTION

Rather than the dynamic force, static force from a transistor is devoured because of a little consistent release of current also called spillage. There are a few methods of spillage, however they can be extensively named as sub-edge, entryway and converse one-sided channel and-source substrate intersection band-to-band-burrowing (BTBT) [1, 2]. These spillage instruments have solid reliance with a few gadget parameters (e.g., oxide thickness, doping fixation, temperature), notwithstanding the flexibly and edge voltage of these gadgets. On account of these subsequent impacts, is done with the more testing to precisely evaluate transistor level patterns in static force [10]. In any case, the general static force because of spillage has been consistently becoming issue in the course of the last a few innovation ages [3]. Sub-limit spillage, which has an exponential reliance with the edge voltage, expanded quickly with the scaling of edge voltage and commonly overwhelmed the spillage flows [4]. The limit voltage has nearly little scaling to counter this sub-limit spillage, yet the scaling of other gadget parameters (e.g., oxide thickness) caused an ascent in door spillage and BTBT. In this manner, both these two spillages are presently used to establish a considerable division of the spillage [1]. Regardless of this consistent ascent, a few methods

have indicated huge potential in moderating the spillage (e.g., high-metal oxides, two fold door gadgets). Inside an innovation age, circuit planners pick the ideal methods for controlling the spillage vitality, dictated by the advancement of the general force execution of the framework [5]. The impact of door channel/source underlaps (Lun) on a thin band LNA execution has been contemplated, in 30 nm FinFET utilizing the gadget with blended mode recreations. Studies are done by keeping up and not keeping up the spillage current (Ioff) and edge voltage (Vth) of the different gadgets. To show signs of improvement commotion execution and addition, Lun in the scope of 3-5nm is suggested [13]. The structure of a chip must conjoin two key innovation patterns: i.e., to have drifts in semiconductor industry that underlies the fundamental structure square of microchips and to get patterns in programming that sudden spike in demand for these microchips [11]. The most significant pattern in the semiconductor innovation is the scaling of transistor gadgets, permitting to twofold the quantity of accessible transistors on a chip like clockwork [6]. Notwithstanding this exponential development in gadgets, the scaling additionally accompanies an improvement in the exhibition of every gadget, empowering all the more remarkable chip in progressive ages. Misusing the simultaneousness, and the calculation data transmission managed by these multicores is a significant plan challenge for the future frameworks, enveloping different layers of framework creators from the equipment modelers to the application engineers [9]. Then again, the predominant correspondence transmission capacity and inactivity between on-chip centers to have novel procedures to deal with the total access, in order to fulfill the interest for execution and vitality proficiency in the circuit. A large number of these strategies may have been altogether infeasible in multiprocessors worked from uniprocessor chips and subsequently neglected to legitimize genuine contemplations previously. One more test rotates around guaranteeing execution, particularly when the characteristic unwavering quality from the equipment segments is required to fall with the innovation scaling [10, 11]. A few current age multiprocessor chips are as of now working close to the building limit for power, in light of air cooling [12, 13]. With the exponential development of transistor gadgets, the capacity to put all the more handling centers on a chip is currently ready to effectively outpace the relating capacity to control up these centers at the same time.

Standard FinFET's properties are homogeneous to the planar technology, except it differs in its electrical performance and its arrangement of transistors. One or more FinFET fabric is the independent gate (IG)

FinFET. Each independent gate is separately controlled by another. By employing the IG device, the probability of obtaining new network topologies of transistors is more useful than using the planar or standard FinFET devices. These extended work provoked for the new research, to make use of IG FinFET technology cost-effectively in VLSI design [14]. By using IG FinFET, a balance between power utilization and delay can be achieved when compared with that of planar transistors or standard FinFETs. Hence by the use of IG FinFETs, there is a reduction in the number of devices used to construct logic gates. From the single device, the probability of logical OR and logical AND operation are performed which emerges from the reduction in the count of transistors in logic gates. By the usage of single device operations, it may be implemented by providing low threshold voltage to the transistor. Regardless of the other transistor by the usage of only one device by conducting channels are formed if and only if one gate is active and when both the gates are active and when both the gates are active there is a higher inversion level [17]. Previously for the OR operation, low threshold voltage was applied but for AND operation high threshold voltage is applied. Therefore for AND gate operation both the gates must be an inactive state for constructing the conducting channel. In this study, it is proposed that IG FinFET explore the reduction of transistor count in the logic gates and their principle of individualistic feature double gate devices will be taken into account. Furthermore, from electrical simulations, the effect of power consumed from different topologies obtained from IG FinFET which is used to implement logic functions are studied. However, the delay in the logic circuit's critical path is determined by different input ranges, transistor series output characteristics and also by the stacks of such series in circuit [18]. In a decade for the construction of IC's more planar technology was used, However to reduce the length of the channel short channel consequence in sub threshold operation region is degraded. By reducing the short channel effects, reductions in the parameters of current, voltage, electric field and transistor length can also be increased by the application of MUGFETs. FinFETs are one of the most significant types of MUGFETs which provides more effective characteristics for the design of digital IC. In this chapter, characteristic and conduction of transistor and its electrical stimulation of IG, FinFET's are presented and after that fundamental working principle of independent gate transistors is given with a different configurable arrangement, threshold voltage and utilization of logic gates [19].

The FinFET based microprocessor designs are the most fundamental trends in the latest technologies such as

demand of semiconductor technology which is a group of basic building blocks such as ALU, registers, buffers, encoder and decoders and software trends that simulates or compiles the microprocessors [20]. Behind a considerable lot of these innovation patterns, the demand for semiconductors plays a basic job and profoundly impacts in their development, including the state of microchips. The most significant pattern in semiconductor innovation is the scaling of transistors, permitting of multiplying the quantity of accessible transistors on a chip for every year [21]. Notwithstanding this exponential development in devices, the scaling likewise accompanies an improvement in the presentation of every transistor device, empowering all the more dominant microchips in progressive ages. For sure, no other industry in history has seen the wonderful development that the microchip business has exhibited in the previous a very long while. FinFET has double Gate, so it has better control of channel from transistor gates, reduced short channel effects, with better Ion/Ioff and improves the sub-threshold slopes [22]. The power consumption in both dynamic mode and static mode are less as compared to traditional CMOS circuits. Here, the IC's design becomes most important aspect, due to its low cost, low area and high speed process activity. For two decades, low force/vitality configuration has been a significant structure limitation[23]. The trend in advanced battery lifetime, improve framework unwavering quality and decrease cooling costs has pushed for broad research in low force/vitality computerized structure. The static commotion edge of FinFET static random access memory (SRAM) cells working in the subthreshold region was explored utilizing an explanatory arrangement of 3-D Poisson's condition [24]. A diagnostic SNM model for subthreshold FinFET SRAM was shown and approved by 3-D innovation PC structure with blended mode reenactments. When contrasted with mass SRAM, the standard 6T FinFET cell indicated bigger READ SNM (RSNM), better fluctuation insusceptibility and lesser temperature affectability of cell [15]. Since, Moore's law-driven scaling of planar MOSFETs faces imposing difficulties in the nanometer system, FinFETs and Trigate FETs have risen as their replacements. Attributable to the nearness of numerous (two/three) entryways, FinFETs/Trigate FETs can tackle short-channel impacts (SCEs) better than traditional planar MOSFETs at profoundly scaled innovation hubs and in this manner empower proceeded with a transistor scaling. In this work, we inquire about on FinFETs from the bottommost gadget level to the highest design level [25-27]. We review various kinds of FinFETs, different

conceivable FinFET asymmetries and their effect with novel rationale level and design level tradeoffs offered by FinFETs. We likewise show the investigation and enhancement that are accessible for portraying FinFET gadgets circuit models [16].

II. PROPOSED DESIGN OF 4-BIT NANO PROCESSOR DESIGN

Analysis of Power, area and delay for proposed 4-bit Nano-processor design: This research work mainly concentrates on the optimization of area, power and delay. The area is mainly depending on the number of transistors integrated to design the FinFET based processor using FinFET transistors and it can be expressed as $Area_{die} = \sum_{j=1}^{M} Size_j$, where $Area_{die}$ represents the silicon die area, M is the total number transistors integrated on single processor, $Size_j$ is the size of the each transistor. The total area of the FinFET based processor is designed by considering the various factors like total area die expected, wafer size and cost after fabrication for the market place. Afterward, the total area limitation or finalization of area for die, should then be met by the total sizes of incorporated each transistors. As the condition shows, one can either decrease the size of every transistor or essentially expel the necessary number of devices to fulfill this requirement. After the minimum size of the device is obtained using FinFET technology, rectifications of the violation for the area constraints are possible done and can be scaled down by transistor count. In this work, the area optimization is done by using "*fingering*" concept. Fingering is to upgrade the resistance of the gate poly terminal along the width of the transistor. Since, the gate poly is driven from one end and entryway poly is resistive, there might be motivation to have a rule that expresses the most extreme width of a solitary finger. But, the fingering requires more number of transistors but fingering concept will be the techniques to not only have area optimization, the active capacitance can also minimized because of the drain region is bounded with gate poly instead of the electric field.

A.Technique to optimize the area by using Fingering Concept

➤ Keep a fixed size transistor so that the height remains constant, but then add more transistors in parallel to deliver more current to the load.

➤ In this case, join in clamped design mode with the source and drain terminal so that they act as taps in parallel. Through particular, when 2x and 4x drive force buffers are used, it will deliver 8 times more current than the normal 1x buffer by using 8 or even more fingers.

➢ Another benefit of fingering is that, the resistance significantly lowers. Let us just assume you have a resistance of R from as given, now all these resistances come in parallel when fingering is done, therefore, the resistance reduces by a factor of N, this is another notable advantage to fingering.

The second most parameter is power consumption or dissipation of the proposed processor chip which is measured as similar case of area discussed in above. This power is mainly depending on the cost-effective capacity cooling of more number of transistors i.e., for more volume processor designs. In the recent research, the energy in term of efficiency is optimized in different ways on the silicon chip budget. The power measurement in general is $P_{dissipationdie} = \sum(j=1..M)*Power_j$, where $P_{dissipationdie}$, is the power constraint on the chip design level, $Power_j$ is the power dissipated by the complete processor and M is the total number transistors used for the designing of the FinFET based processor.

In this research work, the 4-bit Nano-Processor is designed; with one of the constraint of this processor is analytic model of optimizing the power which is sum of the dynamic and static powers. The dynamic power is mainly due to switching activity of the transistor which is due to charging and discharging of the capacitor load at the output of the each and every transistor and it can be expressed as $Power_{dy} = C_l.V^2dd.F_{signal}$. The static power is because of spillage that has been consistently becoming in the course of the last few innovation ages. Sub-edge spillage, which has an exponential reliance with the limit voltage has expanded quickly with the scaling of edge voltage and commonly ruled the spillage flows. The limit voltage has nearly quit scaling to counter this sub-edge spillage; however, the scaling of other gadget parameters (e.g., oxide thickness) has caused an ascent in entryway spillage. In any case, there are a few instances of over-provisioning in the smaller scale design segments of current usage (e.g., the physical register document and guidance window passages). A FinFET based Nano-Multicore System (FinFET-NMS) is a class of multicores that by configuration arrangements of more handling center assets that can be kept dynamic for the objective of Thermal Design Power (TDP). The key differentiation between an ordinary multicore and a FinFET-NMS originates from the characteristic objectives in their framework structures, separately. Authoritative multicore considers a preparing center as one of the most basic segments of the framework and thus targets full use of it. A FinFET-NMS considers power and warm attributes more basic than the handling centers. In this proposed wok, the design consists of transistor with

two controlling gates which is investigated as a circuit component and is utilized for the execution of various combinational basic components like gate, ALU, Multiplexers, Multipliers, CSA and encoder and decoders systems with utilizing these elements. The 4-bit Nano-Processor has been designed which depends on FinFET technology and is designed from Schematic level to layout using GDS-II level. The verification is done through physical design to validate the power, delay and area. With separately gates connected, series/parallel exhibit could be performed utilizing a solitary device. Various topologies got from the different gates activity were tried through electrical reenactments and the outcomes exhibit the current exchange off between these two parameters. Additionally, a diagnostic defer articulation was inferred for logic systems which use IG FinFETs, determining scientific articulations for the effect of diminishing varieties of arrangement of transistors in logic gates. The investigative model for IG-FINFET devices was tried in coherent information way and contrasted with SPICE reproduction results, demonstrating its utility for the planning examination of computerized circuit. In this work, mainly discussed about design of the ALU which includes all basic gates and arithmetic operation like multiplier, Carry Save Adder (CSA), MUX, 4-bit SRAM, Gilbert multiplier and buffers, through which the performance effects on the Nano-Processor is analyzed as shown in Fig.1. The complete processor design could be achieved and given in the following:

➢ The increase in source to drain current along with the simulation of electrical environment has been set-up and it is capable of controlling the nature of FinFET devices with gate controlling or coupling capacity.

➢ All transient electrical signals of different modules (discussed in the previous chapters) has been build using FinFET transistors and measured their performance metrics like delay and power dissipation.

➢ In order to measure the different equations of delay, transitions of the different modules based on FinFET logic structures, have been designed and compared with the existing work.

The FinFET device is a symmetrical transistor having same gate function as CMOS gate function i.e $G_1 = G_2 = G_{FinFET}$ and their oxide thicknesses are same i.e., $t_{ox1} = t_{ox2} = t_{FinFET\_ox}$. The gate and oxide thickness functions the control of the current flowing between source and drain which is very easy and can obtain the maximum current gain and voltage gain. The first part of design is 4 bit ALU using FINFET Technology. The CMOS Technology has the major issues of scaling of short channel effects like DIBL, sub thresholding and GIDL,

degradation of mobility and modulation of channel length. To address these issues with CMOS, the proposed design has been changed to mitigate technology formally called FINFET Technology. The main advantage of it are controlling of current since, it has two Gates, reduction of leakage current, reduction in

delay of each transistor, minimum VT and also FinFET transistors operates at a low power levels. The main concentration of 4 bit ALU design is to optimize the power delay, area, V_T and leakage currents.

| AND | OR | NOT |

| XOR | XNOR | NAND |

4-bit Carry Save Adder

4-bit Multiplier

FinFET transistors Schematic Level Design

4-bit Multiplexer

4-bit SRAM Cell

4-bit Encoder and decoder

4-bit Data Compression and Decompression using ANN

Analysis of AC, DC and Transients responses

Design of Layouts using Virtuoso Cadence tool

Verification of Physical Design using Lambda based rules

DRC, LVS andRC Extraction for Layout

Measurements of delay, area and power

GDS-II Design File Generation

Validation of the results

Fig.1 Proposed 4-bit Nano-Processor design using 32nm FinFET technology

The ALU consists of complex multiplier, comparator, substractor, carry save adder (CSA) which are the arithmetic operation and the seven logical gates are designed by 32 by using 32nm FinFET Technology for 4 bit as shown in Fig.2. The designed 4 bit ALU and its basic building block are used in the Nano processor. The 11 input multiplexer is designed to select one of the ALU outputs and given its input to MUX to get the desired output as shown in Fig.7.2. The first part of design is 4 bit ALU using FINFET Technology. The CMOS Technology has the major issues of scaling of short channel effects like DIBL, sub thresholding and GIDL, degradation of mobility and modulation of channel length. To address these issues with CMOS, the proposed design has been changed to mitigate technology formally called FINFET Technology. The main advantage of it are controlling of current since, it has two Gates, reduction of leakage current, reduction in delay of each transistor, minimum VT and also FinFET transistors operates at a low power levels. The main concentration of 4 bit ALU design is to optimize the power delay, area, V_T and leakage currents. The ALU consists of complex multiplier, comparator, substractor, carry save adder (CSA) which are the arithmetic operation and the seven logical gates are designed by 32 by using 32nm FinFET Technology for 4 bit as shown in Fig.2. The designed 4 bit ALU and its basic building block are used in the Nano processor. The 11 input multiplexer is designed to select one of the ALU outputs and given its input to MUX to get the desired output as shown in Fig.7.2.

III. RESULTS AND DISCUSSION

To meet power requirements, proposed Nano-processor must consolidate novel strategies for decrease in leakage current (LC), power decrease (PD), area optimization (AO) procedures. This work showed two explicit employments of LC, PR and AO contemplations utilizing scientific models in Nanoprocessor plan. In the first place, this work broadened the past work talked about in presentation and joined LC, PR and AO for unequivocally displaying power limitations in different Nano-Processors structures. This work found that with the expanding power imperative, demonstrated utilizing various degrees of LC, PR and AO, the uneven multicore moved towards the dynamic multicore. All structures of 4-bit Nano-processor cells of CMOS and FinFET were verified, validated and analyzed in term of

simulation waveforms in Cadence Virtuoso tool to have metric investigations like different delays, normalized power consumption/dissipation, power-delay-product (PDP), energy and delay product (EDP). In light of these, the 4-bit FinFET-based Nano-processor is demonstrated to be the most minimal and ideal tradeoff in every single metric execution contrasted with the CMOS-based existing processor as shown in Fig.3. This work mainly concentrated on the design of 4bit Nano-Processor and characterization of FinFET standard cell library of 32nm in Cadence Virtuoso bundle. The first part of this research work is to design of 4bit ALU, it includes all gates, multiplier, fast Carry Save Adder (CSA) and multiplexer. All individual logical and arithmetic modules are design from schematic level to generation of Graphic Data System (GDS-II) level through layout design, DRC and LVS; therefore, the 4-bit ALU can be fabricated as chip. The ALU is designed by using CMOS inverters and the designed ALU schematic is simulated through 32nm FinFET technological library and compared with CMOS technology which is simulated through 32nm CMOS library (without FinFET). The power consumption of AND, OR, XOR, NOT and NAND gates are 36.09nW, 64.970nW, 61.13nW, 33.31nW and 37.45nW respectively. The basic gates and universal gates, CSA, subtraction and MUX are integrated for 4-bit ALU design and its delay, power consumption and area are 0.104nsec, 314.4uW and 56.8usqm respectively.

This demonstrated that by utilizing FinFET innovation in 4-bit processor chip level circuitry, it will improve its characteristics parameters in terms of power, area and delay. In any case, the cell configuration likewise adds to how better the nano-processor performs, as examined with prior. The 4 bit FinFET based Nanoprocessor has a decreased propagation delay with normal energy dissipation, PDP along these lines giving FinFET innovation incredible points of interest in vitality effectiveness for 32 nm innovation. It was likewise confirmed that the 4-bit conventional pass-transistor logic (CPL) FinFET based nano processor performed very well with a decreased measure of PDP contrasted with other cell plans on account of its fast execution and full swing activity. The complete Nano-processor design which includes ALU, Multiplexer, encoder, decoder and ANN schematic in terms of analog transistors are shown in Fig.7.3. The results are validated through simulation of each and every sub components as shown in Fig.4.

The Fig.4 is ANN design, as it includes both compression, decompression and their input and output ports. From the result it is clear that, the delay and power simulations of the designed Nano-Processor were performed to prove that 32nm FinFET transistors is quite better to minimize the leakage currents and have the area optimization using Fingering concept. From the obtained results, it is observed that, there is reduction power of about 48%, reduction in delay is about 51% and 29% of

power reduction as compared to existing works.
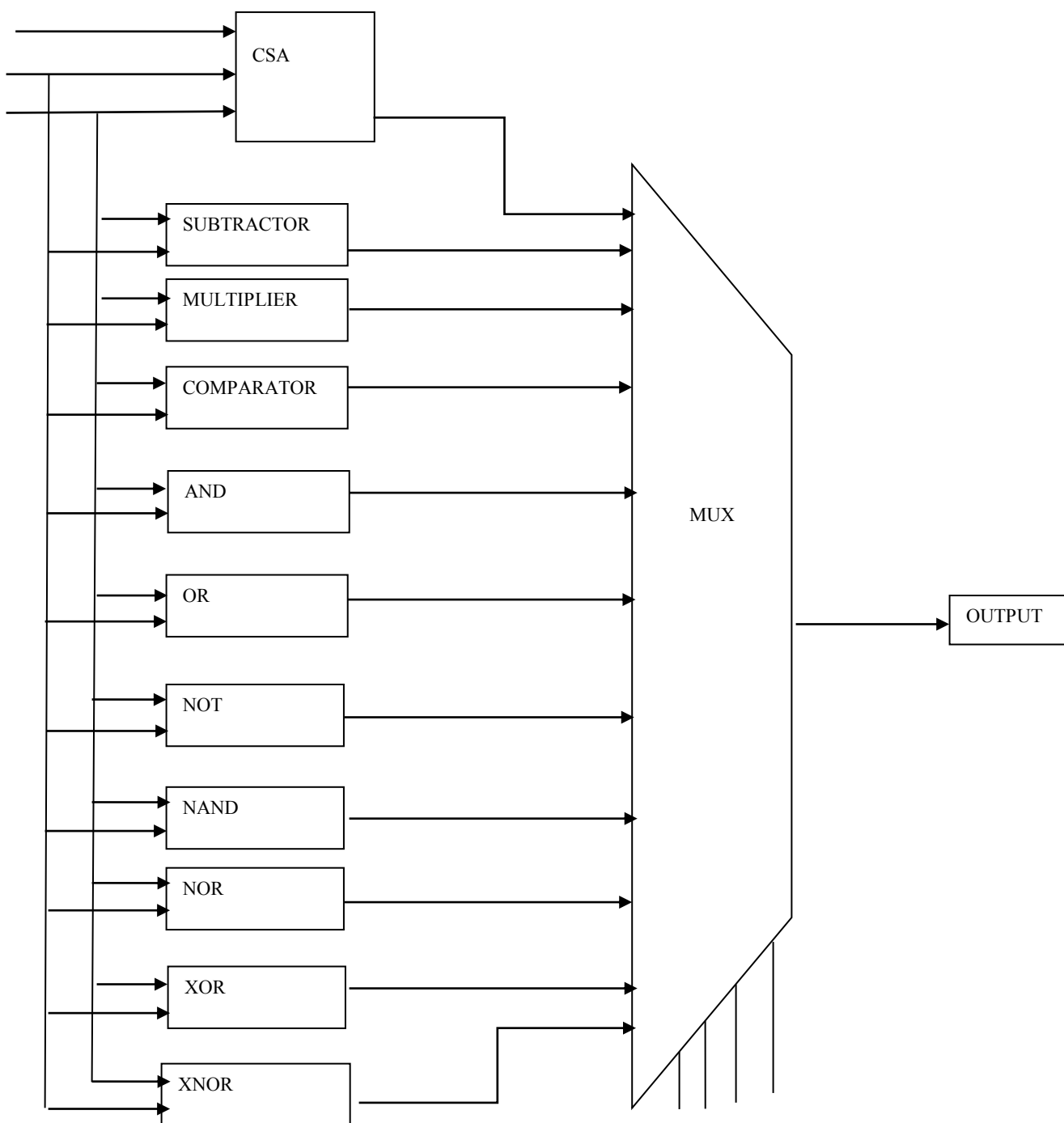
Fig.2 Architecture Of 11bit ALU with 11 Input MUX
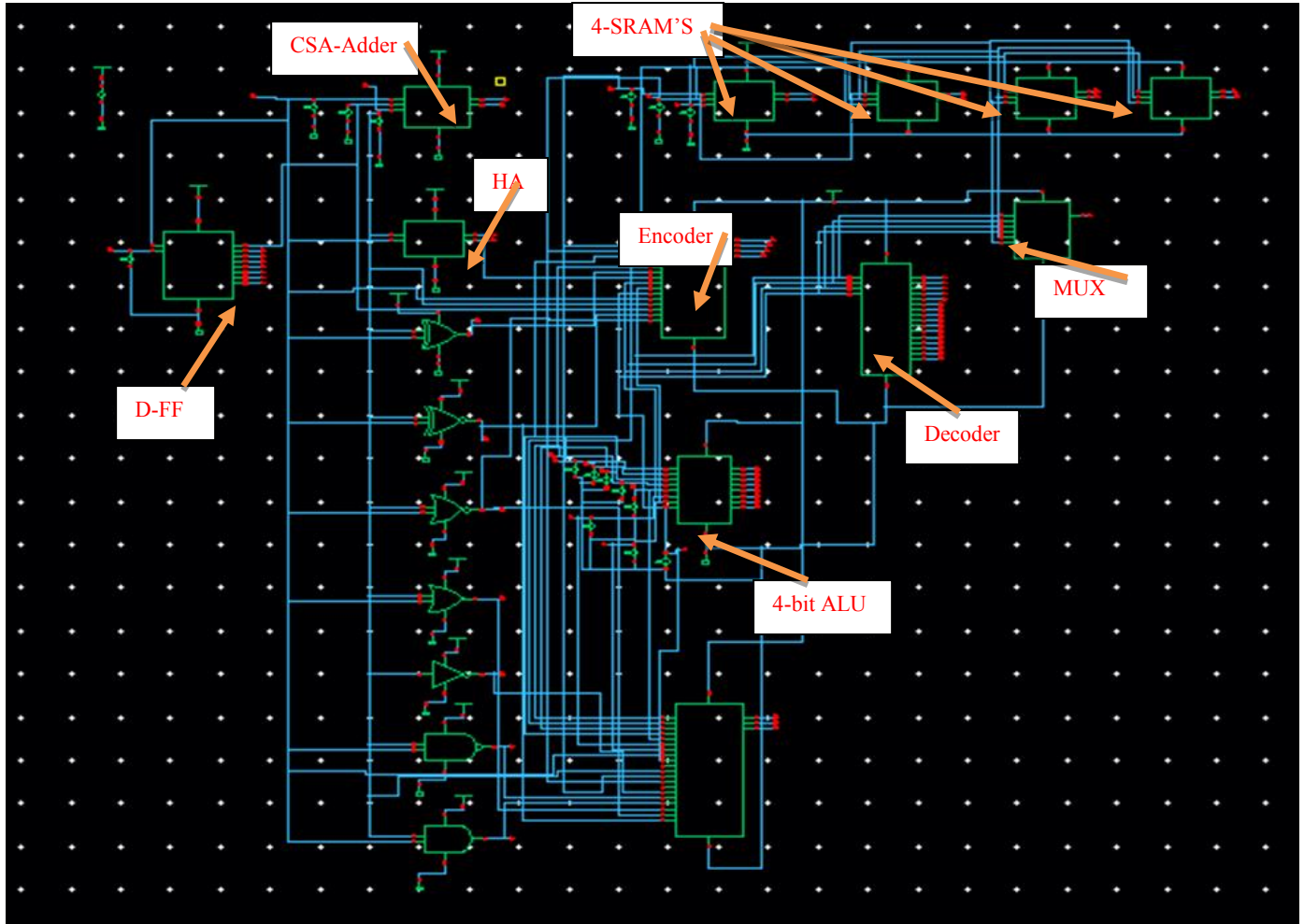
Fig.3 Overall proposed Nano-Processor design using FinFET 32nm technology and subcomponents.
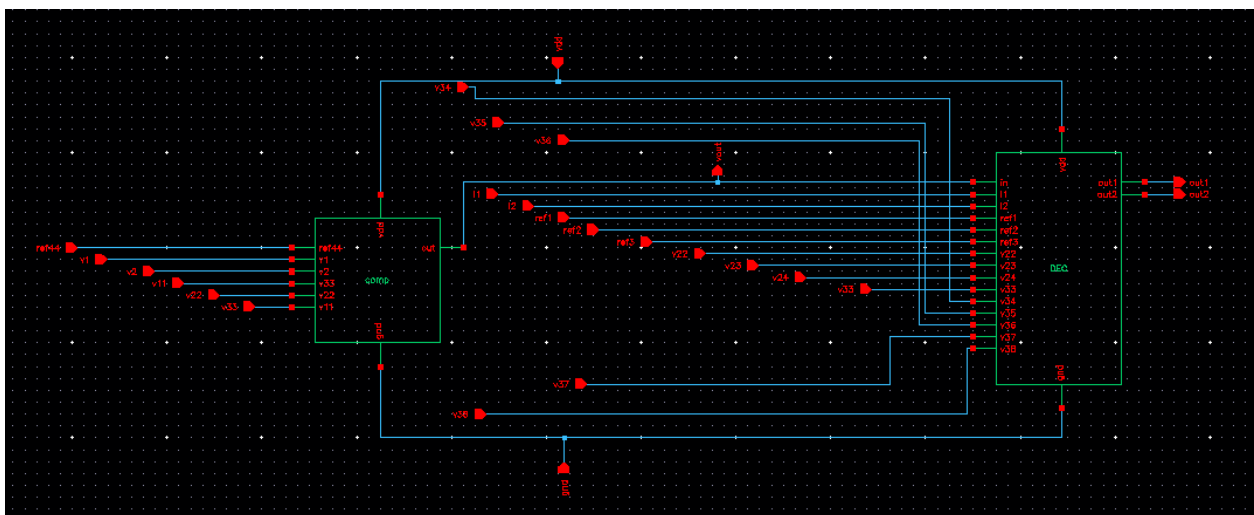
Fig.4 Top level schematic circuit of ANN

IV. CONCLUSION

In this work, the design having all basic operations with ALU, multipliers and using these basic elements, the 4-bit Nano-Processor is designed. This processor design, show the interested feature of FinFET devices for controlling of the current flowing between source and drain and controlling of the transistor operations to obtain better gain, low power, less delay and minimal area. Through dual gate operation, it is observed that, the threshold voltage (V_t) is minimum when compared to the case in the CMOS technology. The complex multipliers and other complex logical functions are simulated using Virtuoso back-end Cadence tool which include the 32nm FinFET technology. The simulated results using different topologies have shown that reduction in the number of transistors, low area and optimized power is obtained. Lastly, delay and power simulations of the designed Nano-Processor were performed to prove that 32nm FinFET transistors would be helpful to minimize the leakage currents and area optimization using Fingering concepts. From the obtained results, it is observed that, there is reduction power of about 48%, reduction in delay is about 51% and 29% of power reduction as compared to existing works.

REFERENCES

[1]. Pahuja, H., Tyagi, M., Panday, S., & Singh, B. (2018). A novel single-ended 9T FinFET sub-threshold SRAM cell with high operating margins and low write power for low voltage operations. Integration, the VLSI Journal, Vol. 60, PP. 99–116. doi:10.1016/j.vlsi.2017.08.004

[2]. Amuthavalli.G and Gunasundari.R, "Revisited Design of Short-pulse Power Gated Approach of Subthreshold Leakage Reduction Technique in Combinational Circuits," IEEE-International Conference on Systems, Computation, Automation and Networking (ICSCAN), 2018.

[3]. Lee, Chee Young, et al. "A Low Power Priority Encoding Technique with Address-Encoder and Reset-Decoder for an Improved Hierarchical Asynchronous Detector." 2018 15th International Conference on Synthesis, Modeling, Analysis and Simulation Methods and Applications to Circuit Design (SMACD). IEEE, 2018.

[4]. Saini, Sakshi, et al. "Design and Analysis of Priority Encoder with Low Power MTCMOS Technique." 2018 8th International Conference on Cloud Computing, Data Science & Engineering (Confluence). IEEE, 2018

[5]. Singh, Vanshikha, and Rajesh Mehra. "Design and Performance Analysis of Area Efficient CMOS Decoder Circuit." International Journal of Scientific Research Engineering and Technology PP.43-48, 2017

[6]. NOWAK, E. J. et al. Turning silicon on its edge (double gate CMOS/FinFET technology). IEEE Circuits and Devices Magazine, IEEE, Vol.20, PP. 20–31, 2004

[7]. QU, J. et al. Study of drain induced barrier lowering (dibl) effect for strained si nmosfet. Procedia Engineering, Elsevier, Vol. 16, PP.298–305, 2011

[8]. FOSSUM, J. G.; TRIVEDI, V. P. Fundamentals of Ultra-Thin-Body MOSFETs and FinFETs. [S.l.]: Cambridge University Press, 2013. ISBN 9781139343466

[9]. ZHAO, W.; CAO, Y. Predictive technology model for nano-CMOS design exploration. ACM Journal on Emerging Technologies in Computing Systems (JETC), ACM, Vol.3, PP.1-9, 2007

[10]. MARRANGHELLO, F. S.; REIS, A. I.; RIBAS, R. P. Design-oriented delay model for CMOS inverter. In: Integrated Circuits and Systems Design (SBCCI), 2012 25th Symposium on. [S.l.: s.n.], Vol.23, PP.1–6, 2012

[11]. SAKURAI, T.; NEWTON, A. R. Alpha-power law MOSFET model and its applications to CMOS inverter delay and other formulas. IEEE Journal of solid-state circuits, IEEE, Vol. 25, PP.584–594, 1990

[12]. NABAVI-LISHI, A.; RUMIN, N. C. Inverter models of cmos gates for supply current and delay evaluation. IEEE Transactions on Computer-Aided Design of Integrated Circuits and Systems, Vol. 13, PP. 1271–1279, Oct 1994.

[13]. DATTA, A. et al. Modeling and circuit synthesis for independently controlled double gate FinFET devices. IEEE transactions on computer-aided design of integrated circuits and systems, IEEE, v. 26, n. 11, p. 1957–1966, 2007

[14]. DAGA, J. M.; AUVERGNE, D. A comprehensive delay macro modeling for submicrometer CMOS logics. IEEE Journal of Solid-State Circuits, IEEE, Vol. 34, PP. 42–55, 1999.

[15]. ROSSELLÓ, J. L.; SEGURA, J. An analytical charge-based compact delay model for submicrometer cmos inverters. IEEE Transactions on Circuits and Systems I:Regular Papers, IEEE, Vol.51, PP. 1301–1311, 2004.

[16]. "http://www.itrs.net/," International Technology Roadmap for Semiconductors, 2013. 2, 44, 80, 81, 86

[17]. M. H. Abu-Rahma and M. Anis, Nanometer Variation-Tolerant SRAM Statistical Design for Yield. Vol.9, PP. 34-76, Springer, 2013

[18]. S. Ghosh and K. Roy, "Parameter variation tolerance and error resiliency:New design paradigm for the nanoscale era," Proceedings of IEEE, Vol. 98, PP. 1718–1751, October 2010. 3, 17

[19]. S. Ganapathy, R. Canal, A. Gonzalez, and A. Rubio, "Circuit propagation delay estimation through multivariate regression-based modeling under spatio-temporal variability," in Design, Automation and Test in Europe Conference and Exhibition, 2010. 3, 28

[20]. S. Tawfik, Z. Liu, and V. Kursun, "Independent-gate and tied-gate finfet SRAM circuits:design guidelines for reduced area and enhanced stability," in Microelectronics, ICM Internatonal Conference on (ICM), 2007.

[21]. M. Fan, Y. Wu, V. Hu, P. Su, and C. Chuang, "Comparison of 4t and 6t finfet sram cells for subthreshold operation considering variability, a model-based approach," Electron Devices, IEEE Transactions on, Vol. 58, PP. 609–616, March, 03, 2011

[22]. X. Wang, A. Brown, B. Cheng, and A. Asenov, "Statistical variability and reliability in nanoscale FinFET's," in Electron Devices Meeting (IEDM), IEEE International, 2011.

[23]. A. Asenov, A. Brown, J. Davies, S. Kaya, and G. Slavcheva, "Simulation of intrinsic parameter fluctuations in decananometer and nanometer-scale mosfets," IEEE transactions on electron devices, Vol. 50, PP. 1837–1852, September, 03, 2003.

[24]. E. Baravelli, M. Jurczak, N. Speciale, K. Meyer, and A. Dixit, "Impact of ler and random dopant fluctuations on finfet matching performance," IEEE transactions on nanotechnology, Vol. 7, PP. 291–298, March, 3, 2008.

[25]. N.Weste and D. Haris, CMOS VLSI Design. Pearson ddisonWesley, 2005

[26]. A. Agarwal, B. Paul, H. Mahmoodi, A. Datta, and K. Roy, "A processtolerant cache architecture for improved yield in nanoscale technologies," Very Large Scale Integration (VLSI) Systems, IEEE Transactions on, Vol. 13, PP. 27–38, January 2005.

[27]. Gennady Zebrev.et.al, "Static and Dynamic Oxide-Trapped Charge-Induced Variability in Nanoscale CMOS Circuits", IEEE TRANSACTIONS ON ELECTRON DEVICES, 0018-9383, 2019 IEEE.

HELPING AID FOR VISUALLY IMPAIRED

Shivarudraiah B | Pankaj Kumar Singh | Prem Prakash Tiwary | Karthik Raman | M S Vishal

Department of Electronics and Communication, BMS Institute of Technology, Visvesvaraya Technological University, India

Abstract: *On a day to day basis we witness people with visual impairment, struggling to distinguish people and things around them. This also limits them from interacting with the surroundings and hence, curbs their desire to step out of their house and explore the world. The World Blind Union has predicted that the cases will rise from 36 million to 115 million by 2050 if the treatment is not handled by better funding. Thus, using technology in order to make life easy for the visually impaired people, is the need of the hour. We were not technically advanced in the recent past, but as of now, the technology is ready and if used by like-minded people towards a specific goal, anything is achievable. In order to provide the blind people a hearable environment, this work focuses on the field of assistive devices for visually impaired people. It converts the visual data captured into an alternate hearable message that can be easily interpreted by the user.*

1. INTRODUCTION

Blindness is the condition of poor visual perception. Blind people are prone to risks as they are not aware of what is happening around them. A growing ageing population is behind the rising numbers. According to research, most number of cases related to blindness and vision impairment are in parts of South Asia and Africa. Although the percentage of the world's population with visual impairments is actually falling, according to the study, with the rise in global population, the scientists predict that number of people with sight problems will rise in the coming decades. Analysis of data from all over the world suggests there are more than 200 million people suffering with moderate to severe vision impairment. That figure is expected to rise to more than 550 million by 2050.

Hence, this work is a small dedication to the blind. For this work, we have integrated several technologies and combined them to form a system that recognizes a person or an object in-front of the user and converts the name of the detected person/object into an audio format, which forms the output and is read to the user through a set of earphones.

2. IMPLEMENATION

To implement this work in real time, we have used Raspberry Pi 4 board as our processing device. The Raspberry Pi 4 is powered by Broadcom BCM2711 which is based on Cortex A-72 core. A camera module is integrated with the board to capture either faces or objects in real time based on the requirement of the user. A set of earphones act as the output source through which the user hears the name of the detected person or the detected object. For the audio output we have used Google Text-to-speech (gTTS), which works as a command line tool to interface with Google Translator's text-to-speech Application programming interface. Additionally, we have developed a personal assistant for interacting with the user and switching between the modes (detecting person or detecting an object) based on the user's voice command.

Therefore, the main objective of this work is maintaining a comprehensive dataset of images of objects that are encountered on a day to day basis, as well as the images of close relatives/friends for accurate identification.

• Integrating a camera module with a Raspberry Pi microcontroller that will capture the images and will process them further.

- •   Using a Deep learning model and training it for image recognition.

- •   Converting the image identified into a voice message and transmitting it through a speaker/headphone.

- •   Building a personalized voice assistant that will take instructions from the user to enter different modes of working based on voice instructions provided by the user.

3. RELATED WORK

In paper [1], the author makes use of CIFAR10 classification used to classify 32x32 pixel RGB images. The reason for choosing CIFAR10 is that it is complex enough to handle most of TensorFlow's scalability for large models. At the same time, the model is small enough to learn quickly and is very suitable for trying new ideas and testing new technologies. This model provides the highest performance with an accuracy of approximately 86° in several hours of GPU training.

In paper [2] makes use of CNN to detect the object. The camera loops to take pictures at a speed of about 2 Hz, and the pictures pass face detection, and all restricted faces are drawn on the screen as a preview..

In paper [3] The main goal of the research is to develop a face recognition system with higher accuracy and improved face recognition system recognition time.) Linear discriminant analysis. The Jacob method is used to calculate the eigenvectors required by the PCA and LDA algorithms. The face recognition system is implemented on the Raspberry Pi 3 board based on the embedded system.

Paper [6] An overview of different combinations of readers (OCR), virtual assistants, and home automation systems with RaspberryPi, this will be a great system combination. It is a useful help for the visually impaired and the disabled. OCR stands for Optical Character Recognition, which recognizes the current text and converts it into audio through gTTS (Google Text to Speech) pre-processing and post-processing.

Therefore, in our work, we are trying to integrate the features of face recognition and object detection with the voice assistant which will make the job easy for the user.

4. PROPOSED MODEL

The Proposed model is divided into two stages in an additional stage of implementing the voice assistant system. The stages of the model are depicted below:
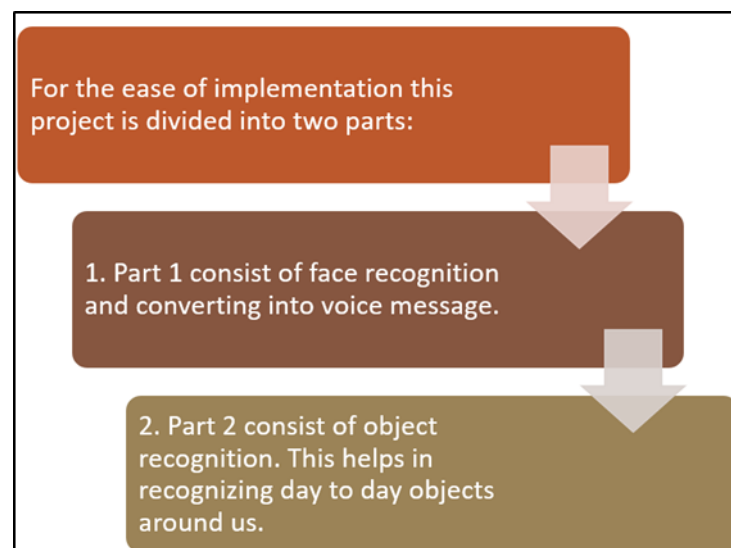
Fig 4.1 Stages in the proposed model

Initially when the system has been set up, the voice assistant guides the user to select the desired mode by giving a particular voice command. The user can choose whether he wants the system to detect a person, or whether he wants it to detect an object. When the camera module, which is integrated to the Raspberry Pi board detects a face, it immediately sends the data to the Raspberry Pi. Further, the Raspberry Pi checks if the detected face is matching with any of the faces that have already been stored in a custom database created for each individual. If the detected face matches with the entries in the database, the name associated with the face will be stored in a variable which would then, be converted into an audio source using gTTS and this audio is heard using the earphones. If the detected face is not matching with any of the entries in the database, the voice assistant asks the user if he wants to save the detected face in the database or let it pass. Based on the user's command the system either clicks multiple photos of the face and stores it under a name or forgets the face.

The model makes use of Haar cascade for the purpose of face detection. It is an effective approach based on machine learning where the cascade function is trained using a lot of positive and negative images. In order to train the classifier, we need to feed in a lot of images with faces, known as positive images and also images that do not have any face, known as negative images. Then we need to extract features from it.

Similarly, when the system switches to object detection mode, the camera detects the closest object to the user and sends this data to Raspberry Pi. Further, this image is matched with another database dedicated only for objects. The database used in our work is a combination of COCO dataset, YOLOv3 dataset and tensorflow lite. If the image is successfully matched with the entries of the database, the name associated with it is stored in a variable and will be converted into an audio format using gTTS, which will be produced as the output through the earphones. If the image does not match with any entry in the database, then the user can choose to save it in the database with a name given to it, or he/she can let it pass.

The raspberry pi is powered by using a portable power bank.
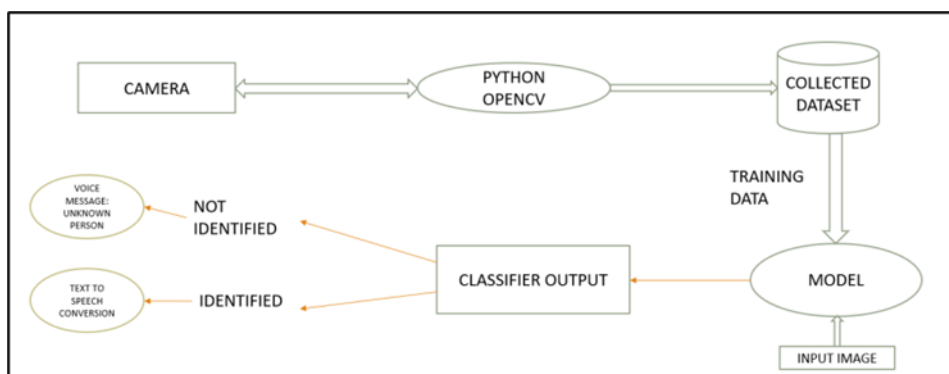
Fig 4.2 Control flow of the system

Important points to note:

1. Storing a number of images may require an external Memory.
2. The images are captured using OpenCV library for Python.
3. The Convolutional Neural Network (CNN) is trained for identification of images.
4. The captured images are analyzed using tensorflow library.
5. The expected accuracy of the model is around 90%.
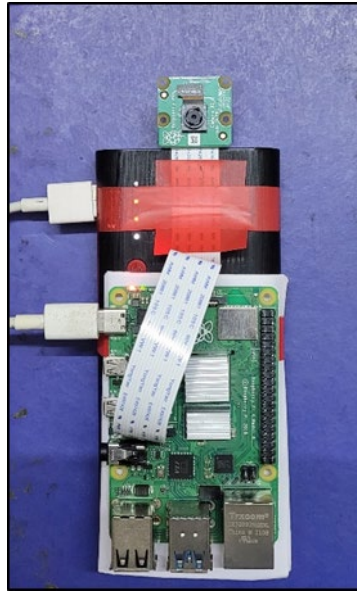
Fig 4.3 Helping assistant prototype

We have kept the size of the device as small as possible, since this is supposed to be used as a wearable assistant.

5. RESULT

Since this is a portable assistive device, testing it under various conditions is very important.

1. When the system works in person identification mode: The device detects the face of the person, crops it and applies the deep learning model to identify the person in front of camera. The Voice message is also transmitted via microphone.
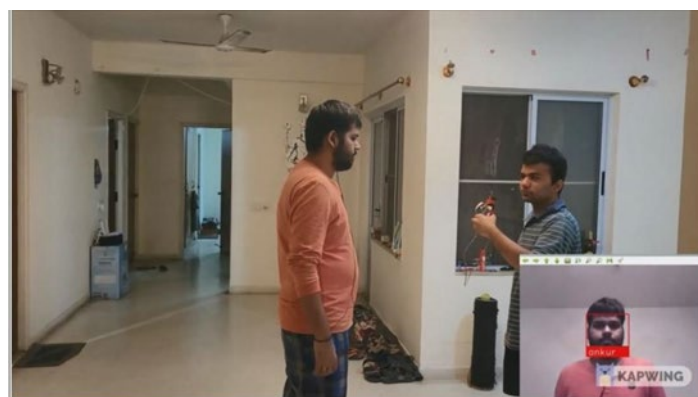
Fig 4.3 & Fig 4.4 Identification of a person

2.   When the system works in object identification mode:

Fig 4.5 Identification of objects

3.   In case an unknown person is detected, the voice assistant gives the user an option to save the person's image in the dataset. The camera clicks 100 images and stores it in database with the name mentioned by the user.

6. FUTURE WORK

In future it can be used for character detection as well i.e. Optical character recognition (OCR), which can recognize italics and handwriting, such as palm leaves. Google Assistant is used to scan QR codes, alarm clock reminders, and date and time notifications. In the future, in a word, it will have a wide range of applications related to artificial intelligence and control. The system can also be used to create a robot that acts as a road-guided robot with a built-in stick that converts images into voice to notify people of obstacles around them. Moreover the accuracy of the system can be improved by training it in multiple environment. With the help of this guide, we can create routes for robots and driverless vehicles and collaborate with the voice commands.
Thus there is great scope for future work.

Acknowledgments

REFERENCES

[1]   International Research Journal of Engineering and Technology (IRJET) e-ISSN: 2395-0056 Volume: 05 Issue: 03 | Mar-2018 www.irjet.net p-ISSN: 2395-0072 Blind Navigation System Using Artificial Intelligence Ashwani Kumar, Ankush Chourasi, Dept. of Electronics and Communication Engineering, IMS Engineering College, INDIA

[2]   Mobile assistant app for visually impaired people, with face detection and sound representation of image. Xiaofei Fu, Electrical Engineering department, Stanford University

[3]   Real Time implementation of Face Recognition System - Neel Ramakant Borkar and Sonia Kuwelkar (Dept. Of ECE, Goa College Of Engineering)

[4]   Outdoor Navigation for Visually Impaired based on Deep Learning Vol – 2514/paper102

[5]   Eyevan: A Smart Artificial Vision for Blind People Renupriya A, Kamalanathan C, Valarmathy S, Yasikka M, International Journal of Advanced Research in Electronics and Communication Engineering (IJARECE) Volume 3, Issue 11, November 2014

[6]   An Assistive System for Visually Impaired using Raspberry Pi, International Journal of Engineering   Research & Technology (IJERT) ISSN: 2278-0181 Published by : www.ijert.org Vol. 8 Issue 05, May-2019

Railway Track Crack Detection and Collision Avoidance using ARM Cortex

[1]Shivarudraiah B, [1]Suryakanth B, [1]Ravindra V Asundi
[2]Abraham Samson Nadar, [2]Nagaraj, [2]Manjunath
[1]Assistant Professors, Department of ECE, BMS Institute of Technology and Management, Bangalore.
[2]Students, Department of ECE, BMS Institute of Technology and Management, Bangalore.

Abstract-**Most of the commercial transport is being carried out by the railway network and therefore, any problems in the same has the capacity to induce major damage to the economy-notwithstanding the societal impact of loss of life or limb. This project is developed to overcome the problem of railway crack, Fog in the environment, Avoiding collision between two stations the cracks is detected through IR Obstacle, the current location is sent through Android GPS and to receive the information GSM is utilized, 4 wheel robot is used in the place of train, as soon as the station unit receives the SMS then E mail has been generated with the android Application and sent to the higher authority Automatically for further proceedings. It is also been planned in this project to implement ultrasound radar sensors to avoid the collision.**

Keywords— GPS, GSM, IR Obstacles, four wheel robot, radar.

I. INTRODUCTION

In today's world, transport is a key necessity because in its absence it would be impossible for products to be consumed in areas which are not in the immediate vicinity of the production centers. Throughout history, transport has been a necessity for the expansion of trade. Economic prosperity can be achieved by increasing the rationality and capacity of transport systems. The proper operation and maintenance of transport infrastructure has a great impact on the economy [1]. Transport, being one of the biggest drainers of energy, its sustainability and safety are issues of paramount importance.

The railway transport occupies a prominent position in quenching the ever burgeoning needs of a rapidly growing economy. The railway transport is growing at a rapid pace, the associated safety infrastructure facilities have not kept up with the aforementioned proliferation. Our facilities are inadequate compared to the international standards and as a result, there have been frequent derailments that have resulted in severe loss of valuable human lives and property as well. On further analysis of the factors that cause these rail accidents, recent statistics reveal that approximately 60% of all the rail accidents have derailments as their cause, of which about 90% are due to cracks on the rails either due to natural causes (like

excessive expansion due to heat) or due to antisocial elements. Hence these cracks in railway lines have been a perennial problem, which has to be addressed with utmost attention due to the frequency of rail usage [2]. These cracks and other problems with the rails generally go unnoticed due to improper maintenance and the currently irregular and manual track line monitoring that is being carried out. The high frequency of trains and the unreliability of manual labor have put forth a need for an automated system to monitor the presence of crack on the railway lines.

To overcome the crack problem we propose robust solution to the problem of railway crack detection using IR LED- IR photo diode which is cost effective. This project is implemented using simple components microcontroller, IR LED, IR photo diode, GSM, GPS module, 4 wheel robot. The IR LED- IR photo diode monitors the cracks in the rails, GPS detects current location, to receive information GSM is utilized, the robot is driven by dc motors [3].

II. DESCRIPTION

The proposed crack detection scheme consists of a Light Emitting Diode (IR LED)-Light Dependent Resistor (IR photo diode) assembly that functions as the rail crack detector. The principle involved in crack detection is the concept of IR photo diode. In the proposed design, the IR LED will be attached to one side of the rails and the IR photo diode to the opposite side. During normal operation, when there are no cracks, the IR LED light does not fall on the IR photo diode and hence the IR photo diode resistance is high. Subsequently, when the IR LED light falls on the IR photo diode, the resistance of the IR photo diode gets reduced and the amount of reduction will be approximately proportional to the intensity of the incident light. As a consequence, when light from the IR LED deviates from its path due to the presence of a crack or a break, a sudden decrease in the resistance value of the IR photo diode ensues. This change in resistance indicates the presence of a crack or some other similar structural defect in the rails. In order to detect the current location of the device

in case of detection of a crack, a GPS receiver whose function is to receive the current latitude and longitude data is used. To communicate the received information, a GSM modem has been utilized. The function of the GSM module being used is to send the current latitude and longitude data to the relevant authority as an SMS. The functionality has been achieved by interfacing the GSM module, GPS module and IR LED -IR photo diode arrangement with a microcontroller. The robot is driven by four DC motors.

Demo points: The ARM is used, IR LED and IR photo diode is used to detect the cracks in the track ,to receive the information GSM module is used, to locate the location Android (APP) GPS is used and 4 wheel robot is used as train.

III.   BLOCK DIAGRAM
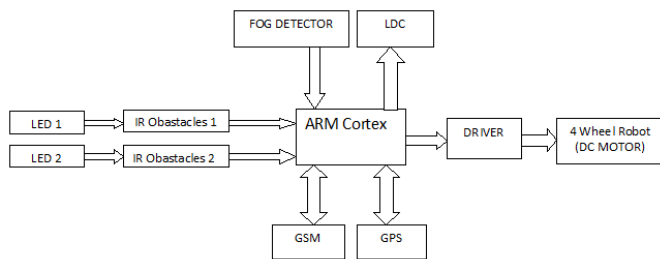
Fig.1 Block diagram for crack detection
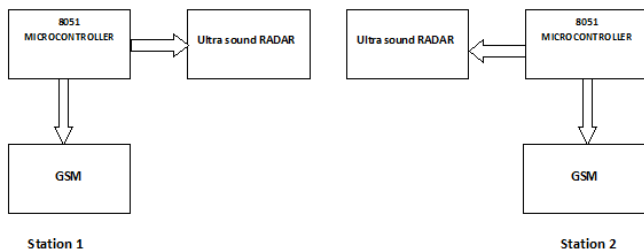
Fig.2 Block diagram of collision avoidance system

WORKING OF THE VEHICLE

1)   When the vehicle is Powered On, it moves along the model track. The IR Obstacle sensors monitors the condition of the tracks .

2) When a crack is detected by the IR sensor the vehicle stops at once, and the GPS in android using android app. The receiver triangulates the position of the vehicle to receive the Latitude and Longitude coordinates of the vehicle position, from satellites.

3)   The Latitude and Longitude coordinates received by GPS are converted into a text message which is done by the microcontroller.

4)   The GSM module sends the text message to the pre-defined number with the help of SIM card inserted in the module.

5) Once the message has been successfully sent to the Number, the vehicle resumes its movement forward depending on the type of crack.
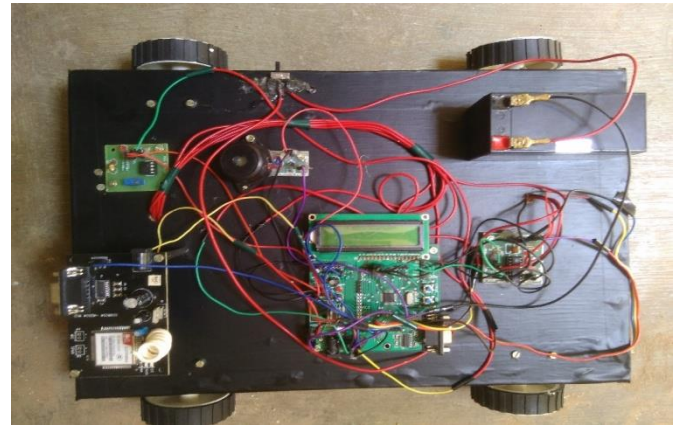
Fig.3 Project model

IV.   FLOW CHART
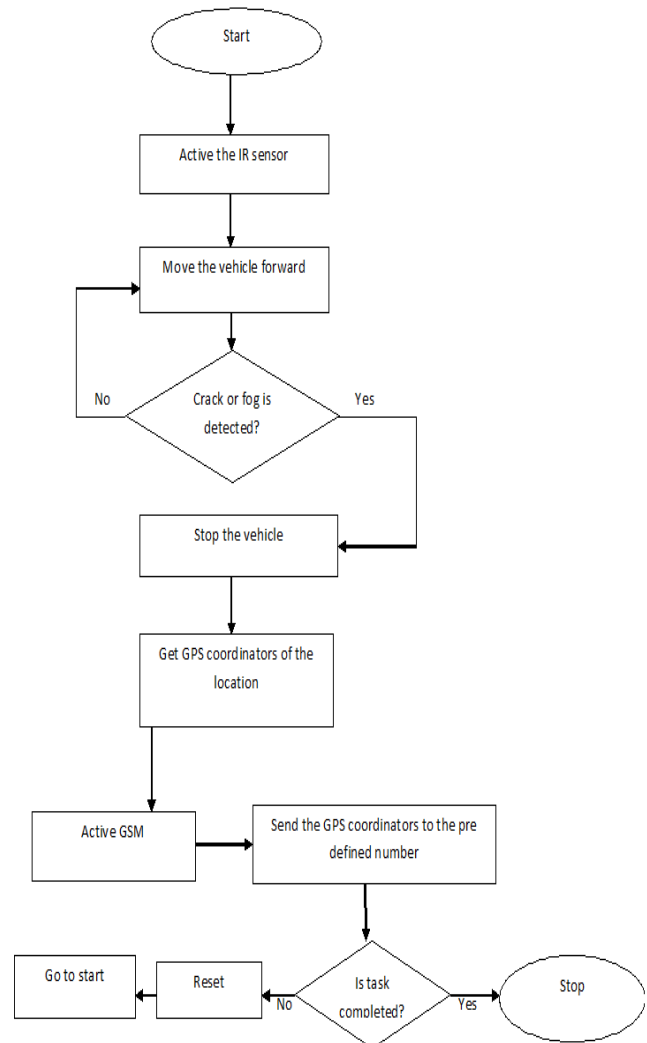
Fig.4 Flow chart for train trolley operation
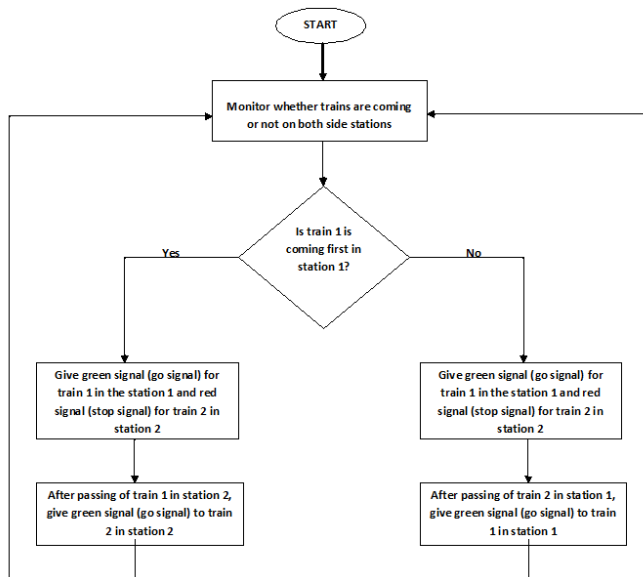
Fig.5 Flow chart of collision avoidance
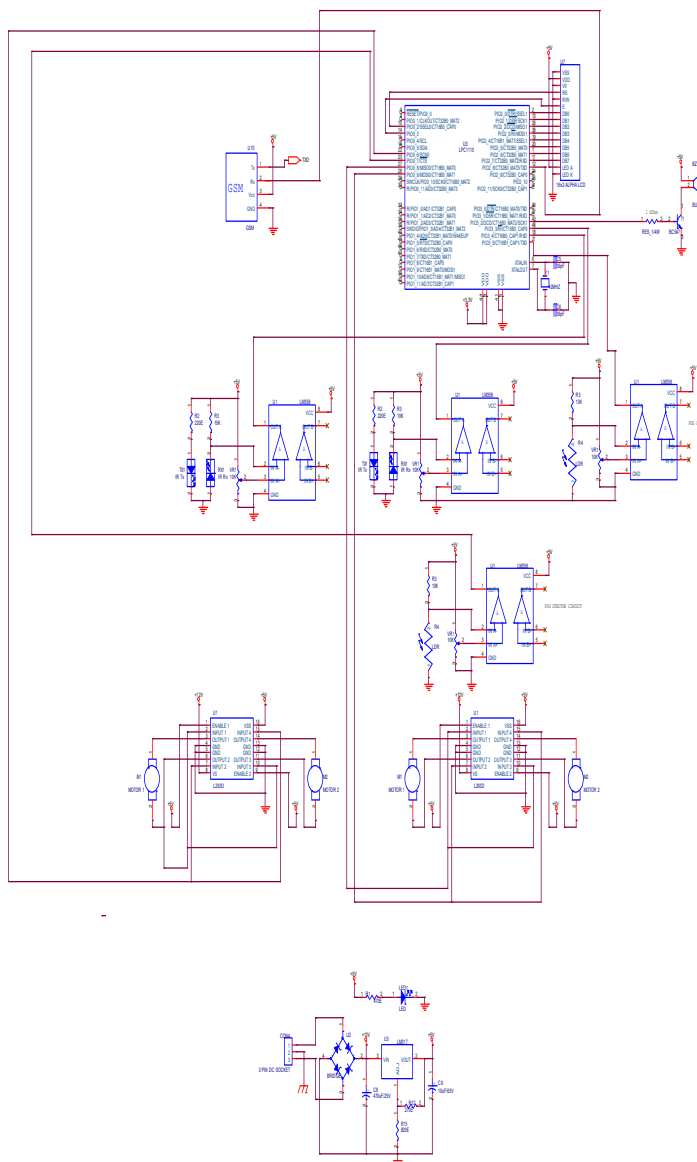
Fig.6 Schematic diagram for trolley module of proposed system

V.   HARDAWARE REQUIREMENTS

A. ARM CORTEX M0(Microcontroller)

The main component of this proposed system is ARM Cortex M0.The LPC111x/LPC11Cxx are a ARM Cortex-M0 based, low-cost 32-bit MCU family. Operate at CPU frequencies of up to 50 MHz. On-chip flash programming memory 64KB. It has 1 KB on-chip static RAM memory. Its operating voltage range from 1.8 to 3.6 volt. Controller contains four ports. Each port contains 12 IO pins. High-current output driver (20 mA) on one pin. It has four general purpose counter/timers with up to eight capture inputs and up to 13match outputs. For converting of signal, it has 10-bitresolution ADC with eight channels. Unique device serial number for identification.

B. Serial interface

UART with fractional baud rate generation, internal FIFO, and RS-485 support. Two SPI controllers with SSP features and with FIFO and multi-protocol Two SPI controllers with SSP features and with FIFO and multi-protocol only. I2C-bus interface supporting full I2C-bus specification and Fast-mode plus with a data rate of 1 Mbit/s with multiple address recognition and monitor mode. Extended temperature (40 °C to +105 °C) for selected parts.

C. Liquid-Crystal Display (LCD)

It is a flat panel display, electronic visual display that uses the light modulation properties of liquid crystals. Liquid crystals do not emit light directly. LCDs are available to display arbitrary images or fixed images which can be displayed or hidden, such as preset words, digits, and 7-segment displays as in a digital clock. They use the same basic technology, except that arbitrary images are made up of a large number of small pixels, while other displays have larger elements

Fig.7 LCD display

D. IR Obstacle sensor

This sensor is a short range obstacle detector with no dead zone. It has a reasonably narrow detection area which can be increased using the dual version. Range can also be increased by increasing the power to the IR LEDs or adding more IR

LEDs. The photo below shows my test setup with some IR LED's (dark blue) as a light source and two phototransistors in parallel for the receiver. You could use one of each but I wanted to spread them out to cover a wider area. This setup works like a Frits LDR but with IR. It has a range of about 10-15cm (4-6 inches) with my hand as the object being detected.
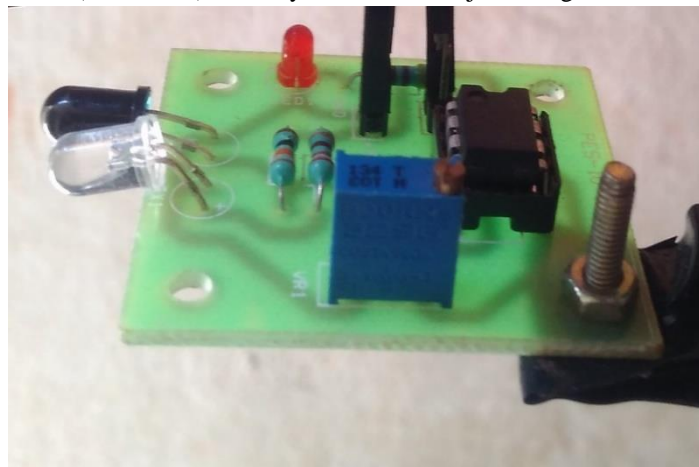
Fig.8 IR Obstacle Sensor used in the project
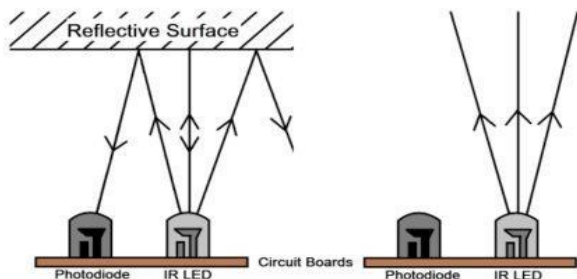
Fig.9 IR obstacle sensor

E.GSM

The GSM standard initially was used originally to describe switched circuit network for full duplex voice telephony to replace first generation analog cellular networks. The standard was expanded over time to include first circuit switched data transport, then packet data transport via GPRS (General Packet Radio Service). Packet data transmission speeds were later increased via EDGE. The GSM standard is succeeded by the third generation (or "3G") UMTS standard developed by the 3GPP. GSM networks will evolve further as they begin to incorporate fourth generation (or "4G") LTE Advanced standards. "GSM" is a trademark owned by the GSM Association.

GSM networks operate in a number of different carrier frequency ranges (separated into GSM frequency ranges for 2G and UMTS frequency bands for 3G), with most 2G GSM networks operating in the 900 MHz or 1800 MHz bands. Where these bands were already allocated, the 850 MHz and 1900 MHz bands were used instead (for example in Canada and the United States). In rare cases the 400 and 450 MHz frequency bands are assigned in some

countries because they were previously used for first-generation systems. Regardless of the frequency selected by an operator, it is divided into timeslots for individual phones to use. This allows eight full-rate or sixteen half-rate speech channels per radio frequency.

India	GSM 900	5 carriers
	GSM 900/1800	3 carriers

Table 1. Types of GSM in India

These eight radio timeslots (or eight burst periods) are grouped into a TDMA frame. Half rate channels use alternate frames in the same timeslot. The channel data rate for all 8 channels is 270.833 Kbit/s, and the frame duration is 4.615 ms. The transmission power in the handset is limited to a maximum of 2 watts in GSM850/900 and 1 watt in GSM1800/1900

Interfacing the GSM Module with Microcontroller.

A SIM card is inserted to the GSM module. After checking as shown in the steps above, RS232 cable is directly connected between DB9 of the module and the DB9 on the controller port. This establishes the serial communication between them. The GSM commands embedded in the controller takes care to communicate with the GSM Module via RS232 cable-MAX232-Serial I/O pins of the Controller.
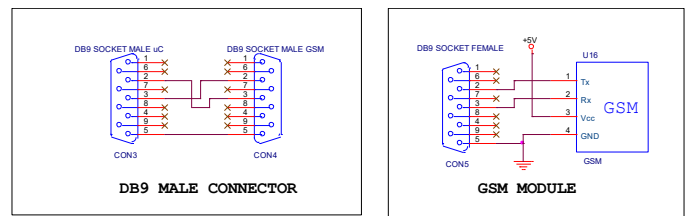
Fig.10 Connection configurations of DB9 male connector and GSM module
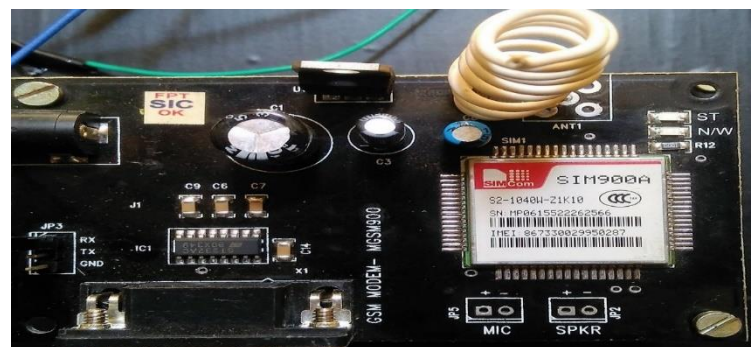
Fig.11 GSM module used in project

F. L293 Motor Driver

The L293 is an integrated circuit motor driver that can be used for simultaneous, bidirectional control of two small motors. Small means small. The L293 is limited to 600 mA,. The L293 comes in a standard 16-pin, dual-in line integrated circuit package. For complete information, consult the Unitrode L293 data sheet (PDF file, 626Kb). The pin out for the L293 in the 16-pin package is shown below in top view. Pin 1 is at the top left when the notch in the package faces up. Note that the names for pin functions may be slightly different than what is shown in the following diagrams. The following schematic shows how to connect the L293 to your motor and the PIC. Each motor takes 2 PIC pins. This table assumes only one motor is connected to the PIC. Here is a table describing the control pin functions. Note that the enable pin is always high when the motor runs so we can simply pull it high on the CCT board and then it will not take up 2 extra PIC pins

CONNECTION DIAGRAMS

Fig.12 Motor driver IC pin diagram
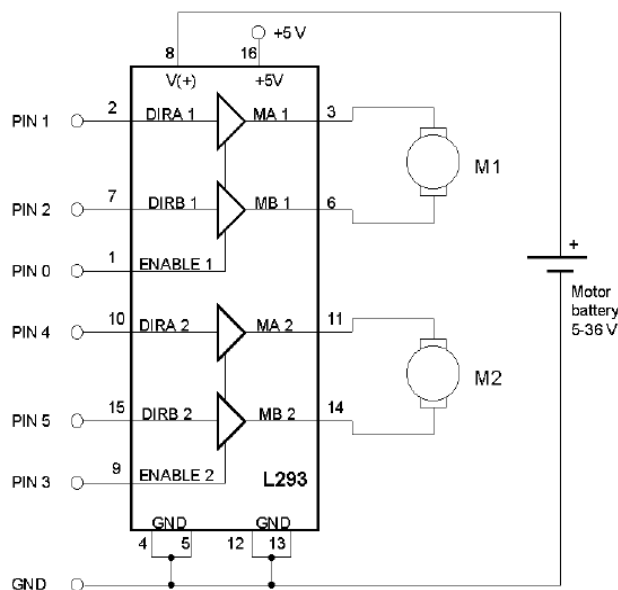
Fig.13 Inaternal diagram of IC L293

ENABLE	DIRA	DIRB	Function
H	H	L	Turn right
H	L	H	Turn left
H	L/H	H/L	Fast stop
L	either	either	Slow stop

Table 2.Truth table of IC L293

Notes

1. You can save on some Stamp pins by connecting the Enable pin to +5V and just using the direction pins to change directions and turn the motor on and off. That means you only need two PIC pins per motor. Put one pin high and the other low for one direction, reverse the state of the pins for the other direction and put both pins low to turn the motor off.

2. Remember to put your finger on top of the L293 when running the motor to see if it is getting too hot.

3. The L293 ground goes to both the battery minus and to the PIC GND.

4. The L293 has an automatic thermal shutdown which means the chip will stop working if it gets too hot.

5. You can also use the L293 to drive relays and solenoids. Just connect the relay coil or solenoid between one of the driver outputs (pins 3, 6, 11, or 14) and ground and turn it on or off with the control pin (pins 2, 7, 10, 15). This is handy because you could run one bidirectional motor and two relays using just 4 Stamp pins and a single L293.

6. The L293 is made by several companies. Use these links to get more information and other versions of the L293 data sheet: Texas Instruments (who bought Unitrode) and STS Thomson .

G. Ultrasound motion sensor

Introduction

Used to detect the move of human or object suitable for indoor and outdoor burglar-proof application, vehicle burglar-proof application, ATM surveillance camera, warehouse surveillance camera, and safety warning application in dangerous site where voltage and temperature exist.

Fig.14 Ultra sound motion sensor
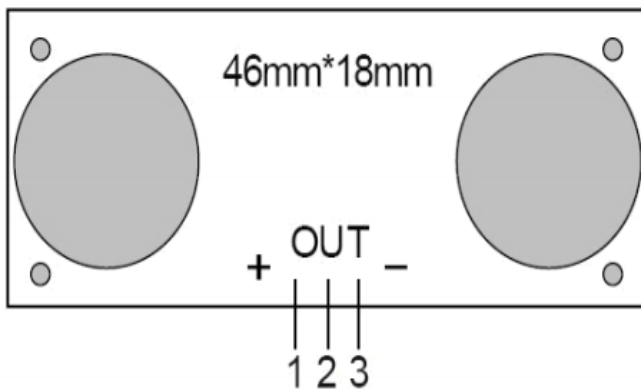
Fig.15 Ultra sound motion sensor connection schematic

1 Power In: positive+

2 Out: signal output

3 Power In : negative –
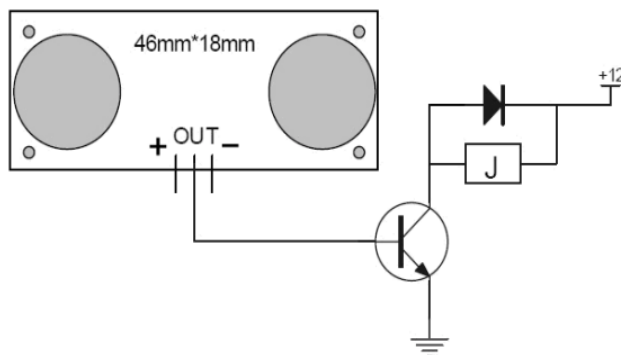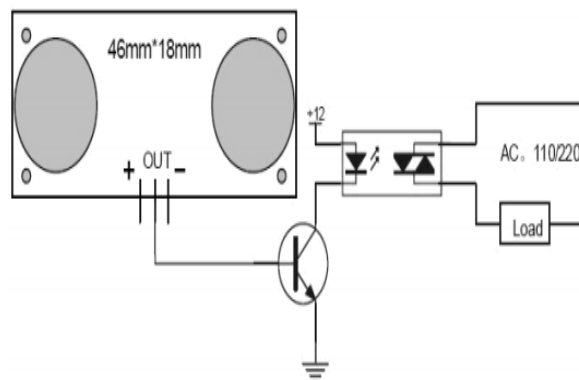

Fig.16 DC Load circuit schematic

Fig.17 AC Load circuit schematic

Main Technical Specifications

1. Power Voltage: DC 6-12V

2. Quiescent current: Less than 2mA

3. Output Level: Low 0V

4. Sensing Angle: no greater than 15 degree

5. Sensing Distance: 2mm-3m

H. Results and discussion

Whenever the crack is detected using IR Obstacles, the signal is sent to controller then the process the signal and send the commands to the appropriate peripherals, such as LCD display and GSM module. The LCD display displays the correct info and GSM sent the information about particular location of the track's crack to the predefined mobile. The android app automatically generates the mail and sent to the predefined mail.

The different problems of the train are solved by this proposed system and are illustrated in the below figures
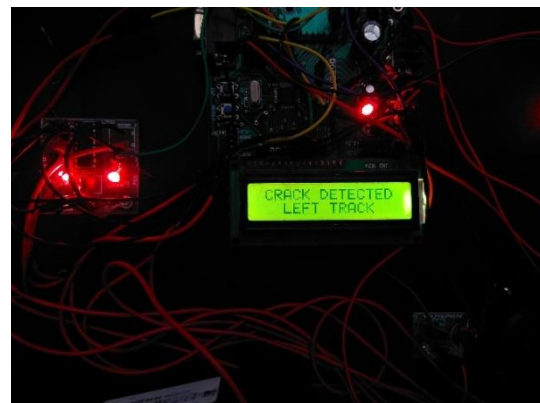
Fig.18 Left track's crack detected and displayed
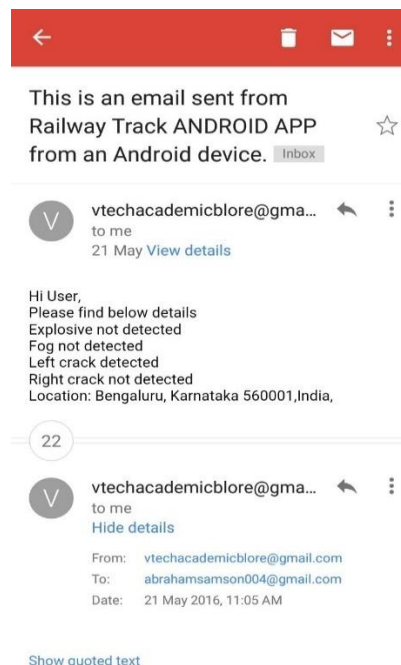
Fig.19 Left track's crack detected and information had sent to the predefined mail
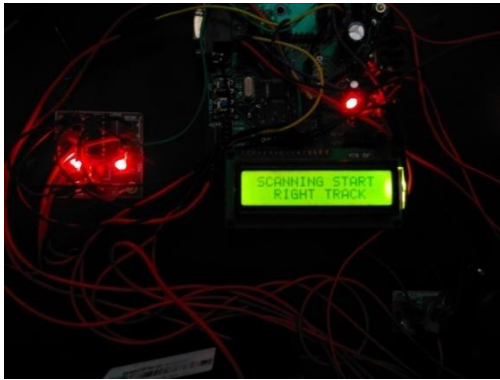
Fig.20 Right track's crack detected and displayed
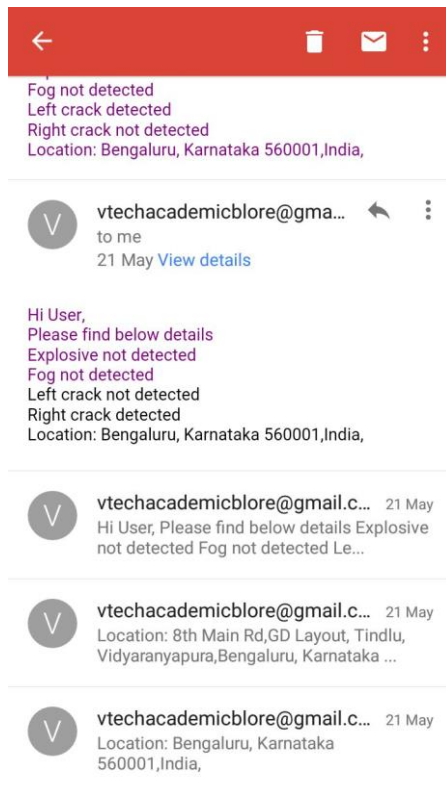
Fig.21 Right track's crack detected and information had sent to the predefined mail

VI. CONCLUSION & FUTURE SCOPE

In this project, we have presented the rationale, design of our robust LED-LDR based railway crack detection scheme. This idea can be implemented in large scale in the long run to facilitate better safety standards for rail tracks and provide effective testing infrastructure for achieving better results in the future. It proposes a cost effective yet robust solution to the problem of railway crack detection utilizing a method that is unique in the sense that while it is simple, the idea is completely novel untested.

The other future scope of this project is automatic welding of cracked railway track. Whenever the cracks are detected, immediately those have to be detected and weld those cracks automatically with appropriate piece of metal. This saves money and time.

REFERENCES

1. Qiao Jian-hua; Li Lin-sheng; Zhang Jing-gang; "Design of Rail Surface Crack detecting System Based on Linear CCD Sensor", IEEE Int. Conf. on Networking, Sensing and Control, 2008

2. K. Vijayakumar, S.R. Wylie, J. D. Cullen, C.C. Wright, A.I. AIShamma'a, "Non invasive rail track detection system using Microwave sensor", Journal of App. Phy., 2009.

3. Wojnarowski, Robert John Welles, II, Kenneth Brakeley Kornrumpf, William Paul, "Electromagnetic system for railroad track crack detection and traction enhancement", Patent US6262573, www.patentstorm.us/patents/6262573/description.html.

4. Tranverse crack detection in rail head using low frequency eddy currents, Patent US6768298, www.google.com/patents/US6768298.

5. M. Cacciola, G. Megali, D. Pellicanµo, S. Calcagno, M. Versaci, and F. C. Morabito, "Rotating Electromagnetic Field for Crack Detection in Railway Tracks", PIERS ONLINE, Vol. 6, NO. 3, 2010.

6. Stuart B Palmer, Steve Dixon, Rachel S Edwards and Xiaoming Jian, "Transverse and longitudinal crack detection in the head of rail tracks using Rayleigh wave-like wideband guided ultrasonic wave", Centre for Materials Science and Engineering The University of Edinburgh,www.cmse.ed.ac.uk/AdvMat45/Rail-crack-detection.pdf.

7. Thomas Heckel, Hans-Martin Thomas, Marc Kreutzbruck and Sven Ruhe, "High Speed Non-destructive Rail Testing with Advanced Ultrasound and Eddy-Current Testing Techniques", NDTIP Proceedings, Prague, 2009.

8. Lanza di Scalea, F., Rizzo, P., Coccia, S., Bartoli, I., Fateh, M., Viola, E. and Pascale, G., "Non-contact ultrasonic inspection of rails.

Influence of Cu Doping in Si–Te-Based Chalcogenide Glasses and Thin Films: Electrical Switching, Morphological and Raman Studies

Diptoshi Roy, B. Tanujit, K. B. Jagannatha, S. Asokan, and Chandasree Das, *Senior Member, IEEE*

Abstract— To understand the electrical switching behavior of $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) series, *I–V* characterization has been performed on bulk as well as amorphous thin films of the as-prepared samples. Both the bulk glasses and amorphous thin films are found to manifest memory-type switching behavior. The threshold voltages of thin-film devices are found to be much lower than the bulk counterparts and hence could find application for phase change memory (PCM). The composition analyses of both have divulged the existence of intermediate phase (IP) in the composition range of $2 \leq x \leq 6$. To examine the probability of the given glass for PCM application, Set–Reset studies have been performed on the glasses with a triangular pulse of 6 mA for set operation and rectangular pulse of 12 mA for the reset operation. The study has revealed a continuous repetition of few Set–Reset cycle by the Si–Te–Cu series. Raman studies carried out on the bulk glasses report the occurrence of blue shift over the composition in a regular manner. Further, SEM studies have been carried out on Si–Te–Cu samples to understand the morphological changes that would have occurred during switching. Additionally, thickness dependence of threshold voltage of representative $Si_{15}Te_{80}Cu_5$ and $Si_{15}Te_{76}Cu_9$ glasses has been carried out to reveal the relationship between threshold voltage and thickness.

Index Terms— Chalcogenide glasses, electrical switching, intermediate phase (IP).

I. INTRODUCTION

PHASE change memory (PCM) is enduring moderate renascence of interest as the faults it encounters can be easily deciphered through engineering solutions. The success of PCM depends on the development of engineering solutions and the corresponding sufficient reduction in costs. The amorphous and crystalline states have a substantial resistance variance which leads to the empowerment of PCM [1], [2]. The noncrystalline state possesses high resistivity whereas the crystalline state manifests resistivity of magnitude which is three to four orders of magnitude lower.

In the last couple of years, a wide variety of improvements like density improvement, aggressive scaling of devices, and reduction in programming power have been seen in the field of PCM. The main reason for reduction in switching power which stands as the rudimentary strength of PCM technology is the pattern of device scaling. The promising attributes of PCM technology are effective scalability (up to few nanometers), faster programming time (in the order of few nanoseconds), ameliorated endurance (up to 10^9 programming cycles), and bit alterability (the capability of PCM to directly write in the memory, without the requirement of prior erasing step). For the features like nonvolatility, reduced usage of power, and direct write, PCM gains to envelope an extensive area of applications like addressing wireless systems, embedded applications, solid-state subsystems, and computing platforms [3].

Metal-doped chalcogenide glasses possess electronic and optical properties and are dissimilar from the properties of the ordinary ones. Glasses containing transition metal like Cu, Ag, and Sn have been employed in optical memory [4] and laser materials [5]. Additionally it has been seen that glasses doped with Cu manifests photoconductivity and shrinkage in bandgap, illustrating their probable usage in economical mid-IR detection which in turn leads to the investigation of electrical properties [6]. The metallic impurities such as copper when added to Si–Te bring interesting variations in their properties. They enter the structural network of chalcogenide glasses in a special way and increase the network connectivity, crystallizing ability, and the electrical conductivity [7], [8]. That is the reason an attempt has been made to study systematically the properties of the Si–Te glassy system with added Cu impurity. The study on the switching behavior of a material is an important part to find the possibility of the material in PCM application.

II. EXPERIMENTAL TECHNIQUES

The bulk sample that is required for making thin-film devices are concocted by melt quenching technique. Pure elements of the order of 99.999% purity have been taken in proportional quantity in clean quartz ampoule which has been sealed maintaining a vacuum of 10^{-6} mbar. The furnace is

Fig. 1. 3-D schematic of sandwich geometry utilized for making amorphous Si–Te–Cu thin-film switching devices.

Fig. 2. $I–V$ characteristic of a representative $Si_{15}Te_{84}Cu_1$ glass showing MS.

programmed to revolve at 10 r/min uninterruptedly with an increment of 100 °C per hour till it reaches 1100 °C. The homogeneity of the melt can be maintained by keeping the ampoule at a sustained temperature of 1100 °C for 24 h and finally quenching it in ice water mixed with NaOH. The studies on the bulk switching have been undertaken to understand whether the material undergoes switching or not. Once it has been confirmed that electrical switching occurs in the bulk samples further switching analysis has been carried out on thin films which are indeed a very important aspect for memory device research.

A clean biological glass slide of 25 mm × 75 mm dimension is used as a substrate on which the source material has been coated using the flash evaporation approach. To execute the switching studies, the devices have been fabricated in sandwich style. The glass substrate has been used as a base for coating and the top and bottom electrodes have been coated with aluminum. The 3-D schematic of sandwich geometry that has been adopted for making amorphous Si–Te–Cu thin-film switching devices is shown in Fig. 1.

The switching studies have been carried out using Keithley (2410^c) source measure unit (SMU) with Lab View 7 (National Instruments) which has a voltage and current limit of 1100 V and 1 A, respectively. To carry out the switching studies on bulk samples, the samples have been polished to obtain 0.3-mm thickness and placed between the contact points of the Keithley SMU. The purpose of testing the bulk sample is to know the property of the material and its feasibility in the application in the form of thin film. 0.3 mm has been maintained only for bulk samples whereas in the films the approximate thickness obtained is 10 μm. A probe station with an applicability of x-, y-, z-direction movement is employed to place the thin film for switching studies. Threshold switching (TS) is volatile which means once the applied field is removed the material reverts to its original state. As memory switching (MS) draws maximal importance in PCM application, in this work MS has been emphasized specifically for both bulk and film.

SEM studies are carried out using VEGA3 TESCAN. Horiba Jobin Yvon (LabRAM HR) in backscattered mode is used to perform Raman studies on the glassy samples. To analyze the backscattering light, the triple monochromator is used and discerned by a charge-coupled device (CCD) cooled at −70 °C. An argon ion laser of 514.5 nm line is
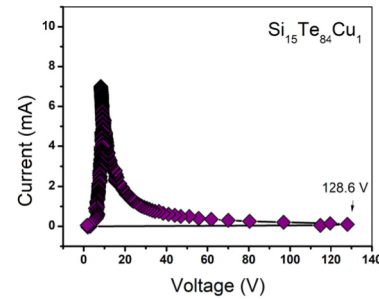
utilized to illuminate the glassy sample wherein the laser is focused with a 50X objective. A superior signal-to-noise ratio is obtained by using an acquisition time of around 120 s.

III. RESULTS AND DISCUSSION

A. Switching Analysis in $Si_{15}Te_{85−x}Cu_x (1 \leq x \leq 10)$ Bulk Glasses

The $I–V$ characteristic of an exemplary bulk $Si_{15}Te_{85−x}Cu_x$ glass is shown in Fig. 2 where an ohmic behavior exhibited by the sample is observed initially followed by a negative resistance behavior. Specified glasses manifest a current-controlled negative resistance (CCNR) behavior at threshold voltage, leading to a high conducting "ON" state. Factors on which the switching by a glassy semiconductor depends are ON-state current, thermal stability of the glass, thermal diffusivity of the material, and the network connectivity. It has been seen here that upon restriction of ON-state current to low values, the glass exhibits TS. The glass is seen to display MS at a higher ON-state current and is being shown in Fig. 2. An analogous trend is also seen in other tellurides [9].

The MS is associated with structural change and it becomes difficult for the system to restructure itself with the increase in cross-linkages and rigidity. Consequently, a memory to TS is noticed in systems having more cross-linking elements [10]. The phase change attribute of the glasses helps in understanding the MS in chalcogenides.

The transition from amorphous to crystalline phase occurs during switching and the later phase is retained in MS. The temperature of the glassy material between the electrodes augments because of the joule heating and creates a filamentary path [11], [12].

The compositional dependence of switching voltage of $Si_{15}Te_{85−x}Cu_x$ glasses is shown in Fig. 3. The two governing features that determine the compositional variation of switching voltage of chalcogenides are dopant resistivity [13] and the network connectedness [14]. Typically the inclusion of metallic dopants is known to reduce the resistivity and consequently the switching voltage in chalcogenides [15]. Conversely, the strengthening of connectivity and rigidity of network is seen to increase the switching voltage. In many cases, high resistivity and network connectivity make the structural changes difficult during switching process resulting in TS as seen in Al–As–Te glasses [16].

Further, the increase or decrease in resistivity and network connectivity can be judged from the variation of threshold (or switching) voltage with composition, and the same has been interpreted in this work. The Si–Te–Cu system has exhibited good MS behavior and also the switching data are reproducible. In the present system, with the intensification of Cu, the threshold voltage (V_T) is seen to increase initially followed by the exhibition of a plateau in the region of $2 \leq x \leq 6$. A precipitous decline is exhibited by V_T above 6% of Cu. The consequence of both the network connectivity because of the incorporation of high coordinated Cu atoms and metallic nature of the dopant are seen to influence the contrast of V_T of bulk $Si_{15}Te_{85-x}Cu_x$ glasses with composition. The analysis of the extended X-ray absorption fine structure (EXAFS) spectrum in literature reveals the existence of coordination number in copper as four [17], a number which has also been observed in the other systems of copper like CuFeS [18], Cu_3AsS_4 [18], and CuAsI [19].

The movement of Cu^+ ions also has a critical contribution in the switching behavior of Si–Te–Cu glasses. The OFF-state highlights the subsequent recombination processes: hole seizure by C_1^- centers ($C_1^- + e^+ \rightarrow C_1^0$) and electron seizure by C_3^+ centers ($C_3^+ + e^- \rightarrow C_3^0$). A passivation of C^- centers can be carried out by Cu^+ ions that exist in the sample by diffusion process [20], consequently minimizing the voltage needed to instigate the switching process, also additionally reducing the switching time of the sample. The initial increment of V_T with composition of $Si_{15}Te_{85-x}Cu_x$ glasses is due to the increase in network connectivity and the subsequent decrement is owing to the incorporation of higher metallic dopant and rapid passivation of charged defect states. Thermal studies on $Si_{15}Te_{85-x}Cu_x$ glasses [21] have shown the presence of a wide trough in the range $2 \leq x \leq 6$ in the compositional dependence of nonreversing enthalpy at glass transition (ΔH_{NR}), revealing the existence of Boolchand intermediate phase (IP). The present study also shows the presence of a wide plateau in the range $2 \leq x \leq 6$ in the compositional dependence of V_T which could be linked with the IP as observed in the electrical switching behavior of Ge–Te and Si–Te glasses [22]. The composition of the glass which resides in the IP is presumed to have self-organization and is independent of stress and so the width of the IP stands as a measurement of self-organization of the glassy system [23]. Additionally, the latent heat of melting of glasses in the IP is nearly absent and this can be a possible reason of low threshold field in the IP [24]. Fig. 3 shows the variation of V_T with respect to composition wherein the thickness of various samples has been kept constant as 0.3 mm.

B. Electrical Switching in $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) Amorphous Thin Films

The $I–V$ characteristic of a typical thin-film device is depicted in Fig. 4. The thickness of the material is approximately 10 μm and the dimension of the top electrode is circular in shape with 6 mm dia.

The comparison of electrical switching response of bulk $Si_{15}Te_{85-x}Cu_x$ glasses with their thin-film counterparts is important for the applicability of the material in PCM.
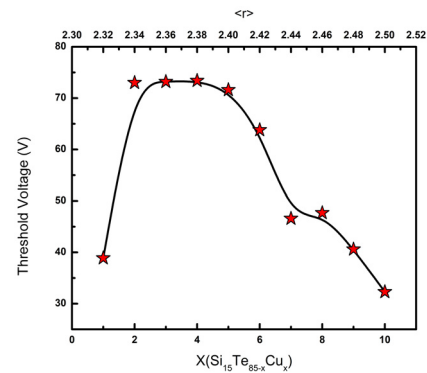

Fig. 3. Compositional dependence of switching voltage of $Si_{15}Te_{85-x}Cu_x$ glasses ($1 \leq x \leq 10$); $< r >$ represents the average coordination number of specific composition.
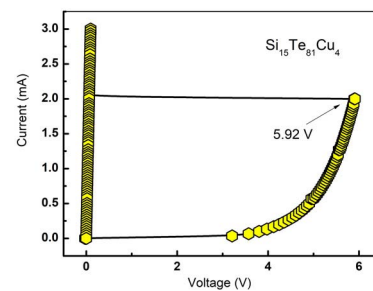

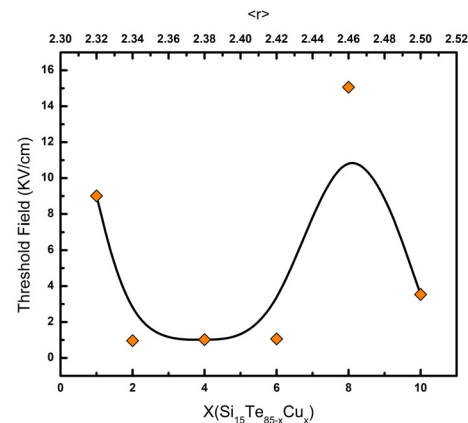Fig. 4. $I–V$ characteristic of a representative $Si_{15}Te_{81}Cu_4$ thin film showing MS.

Fig. 5. Compositional dependence of switching fields of $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) thin films.

The studies on the bulk $Si_{15}Te_{85-x}Cu_x$ sample indicate that both MS and TS [25] have been exhibited by the glassy samples which depend on the ON-state current. Howbeit, the studies on amorphous thin films of $Si_{15}Te_{85-x}Cu_x$ samples indicate the exhibition of only MS. The memory form of the switching responses has been manifested by other binary, ternary, and quaternary amorphous thin films such as GeS_2 [26], Ge–Te–Si [27], and Ge–Te–In–Ag [28].

The compositional dependence of threshold field of amorphous $Si_{15}Te_{85-x}Cu_x$ thin films is shown in Fig. 5. To study the effect of incorporation of dopant on the switching

phenomena, the compositional dependence of threshold field has been investigated for the thin-film devices. In addition, as the usage of flash evaporation technique does not guarantee the uniformity in the thickness of deposition among the devices, the threshold voltage has been normalized as threshold field and used for studying the composition dependence. Further, the coordination number in glasses helps to define the rigidity percolation threshold (RPT) in glasses which manifests the percolative transition of glass from floppy polymeric glass to rigid amorphous solid and is included in Fig. 5. Interestingly, amorphous thin films do not show one to one correspondence of percolation effect when compared with their bulk counterparts. The change observed in the threshold field of thin-film samples as compared to bulk at the onset of percolation could be because of several factors like high field conductions, methods of preparations of thin films, and difference of electrodes in bulk and thin films. The semiconducting thin films which are developed in sandwich geometry with electrodes of metallic nature may have an occurrence of some specific high field conduction process. The several conduction processes are categorized as bulk-limited or electrode-limited.

The bulk-limited is further categorized into space-charge-limited conductivity (SCLC), Poole–Frenkel effect, and hopping. Electrode-limited is classified as tunneling and Schottky effect. These nonlinear conductions may lead to high electric field across the film [29].

The substantial types of variables experienced in thin-film deposition mechanism (e.g., substrate temperature, deposition rate, angle between vapor stream and substrate) enable deposition of different structured films under nominally homogenous environment, which can sequentially influence the optical and electrical properties [29]. Further in bulk the electrode used is brass whereas for thin-film aluminum is used. Even the shape of the top and bottom electrodes is different in thin film and bulk experiments. This difference could also be one of the reasons for change in the threshold field at the onset of percolation. As seen in Fig. 5, the thin films also display an IP in the region $2 \leq x \leq 6$, where it is seen that the samples have a low threshold field which could make them potential phase change materials.

C. Set–Reset Studies in $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) Bulk Glassy System

A study on Set–Reset operation has been carried out on bulk samples of 0.3 mm thickness. To carry out the operation, triangular pulse of 6 mA for Set and rectangular pulse of 12 mA for Reset has been used for bulk samples. The bulk samples of $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) glasses have been found to manifest 5–10 Set–Reset cycles repeatedly. The Set–Reset operation on the representative $Si_{15}Te_{79}Cu_6$ glass is shown in Fig. 6. In the SET operation the sample behaves as a semipermanent storage of information as the sample transits from amorphous to crystalline phase with the introduction of applied current or voltage pulse. The phase is reversed to amorphous by the Reset operation wherein on the application of a sharp current pulse of higher magnitude, local melting
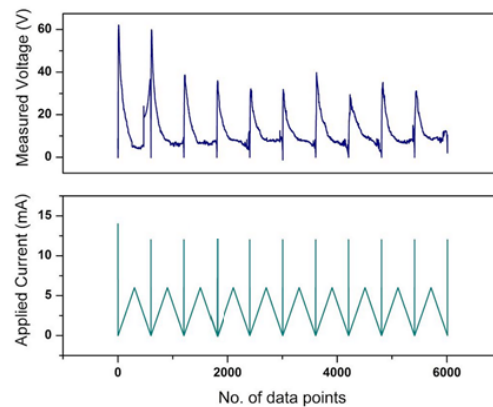

Fig. 6. Electrical switching response of $Si_{15}Te_{79}Cu_6$ glass for a triangular Set pulse of 6-mA amplitude and rectangular Reset pulse of 12-mA amplitude.
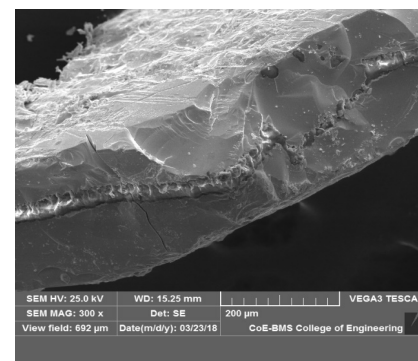

Fig. 7. SEM image of $Si_{15}Te_{77}Cu_8$ showing formation of channel-like structure between the electrodes.

of conducting crystal and amorphization occurs. In Fig. 6, the width of set pulse is 500 ms and that of reset pulse is 1 ms.

D. Morphological Studies on $Si_{15}Te_{85-x}Cu_x$ System Edge Switching

In order to understand the phenomena associated with the switching of $Si_{15}Te_{85-x}Cu_x$ ($1 \leq x \leq 10$) bulk glasses, switching on the edge of a representative bulk sample is carried out. There have been several models (both electronic and thermal) proposed to understand the mechanism of TS and MS in glassy chalcogenides.

For any qualitative analysis of switching behavior, both thermal and electronic effects must be considered, and the two can produce a coupled response called "electrothermal effect." Both the initiation of switching and the maintenance of the filamentary ON-state in chalcogenide threshold switches are fundamentally electronic in nature [30]. The most important role played by the thermal effects is the latching of the sample to the ON-state (MS) which involves a thermally induced amorphous-to-crystalline phase transition in the conducting channel [31]. Fig. 7 is the SEM image of $Si_{15}Te_{77}Cu_8$ bulk sample with a thickness of 0.3 mm, after conducting a switching operation at the edge of the sample. From Fig. 7, a channel-like structure is evident which supports the concept of
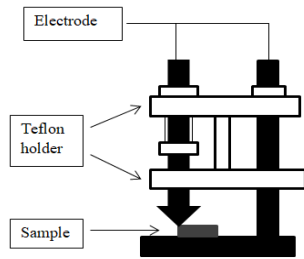
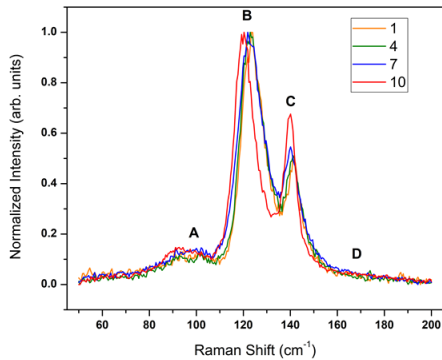Fig. 8. Schematic of edge switching test device and the sample.

Fig. 9. Intensity normalized Raman spectra of selected as prepared Si–Te–Cu glasses. Band A $\sim$ 97.9–92.3 cm^{-1}, B $\sim$ 123.9–120.7 cm^{-1}, C $\sim$ 140.9–139.5 cm^{-1}, and D $\sim$ 165–175 cm^{-1}.

conducting channel like formation during switching. However, further studies on this channel-like structure are not carried out at this stage of research. The schematic of edge switching test device and the sample has been shown in Fig. 8.

E. Raman Studies on $Si_{15}Te_{85−x}Cu_x$ Bulk Glasses

Raman studies are performed on the glassy samples in the backscattering mode. The presented Raman spectra (Fig. 9) are intensity normalized. There are two dominating high-intensity modes (B and C) and two very low intense humps (A and D). The band positions have the following ranges of the wavenumber: band A $\sim$ 97.9–92.3 cm^{-1}, B $\sim$ 123.9–120.7 cm^{-1}, C $\sim$ 140.9–139.5 cm^{-1}, and D $\sim$ 165–175 cm^{-1}. The A and C bands are attributed to the E_{TO} modes of crystalline Te–Te chain. Band B is attributed to A_1 mode of crystalline Te–Te chain [32]. The ordered chains (Te$_n$) of c−Te are recognized to have analogous phonon frequencies at 123 and 143 cm^{-1}.

Band A features chains of Te which are distorted and can be either phase separated by photo-induced bond cleavage or can be attached at the end of each chain of structural elements existing in the glass. Band D can be expected from the stretching mode localized in edge-sharing tetrahedral SiTe$_{4/2}$ units [32]. Band D is the least intense among all. Shift in band A and D does not show much regularity. Band B and C exhibit a significant blue shift over composition, in a regular manner. Within the composition range Si$_{15}$Te$_{84}$Cu$_1$ and Si$_{15}$Te$_{75}$Cu$_{10}$, band B shifts from 123.95 to 120.69 cm^{-1}. The peak intensity remains almost same, within error, for all samples.

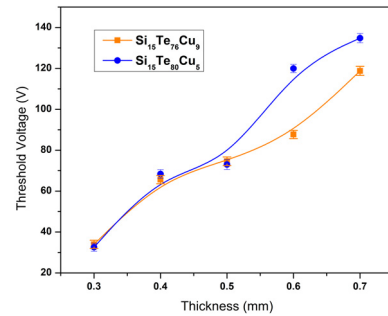Fig. 10. Thickness dependence of threshold voltage of representative Si$_{15}$Te$_{80}$Cu$_5$ and Si$_{15}$Te$_{76}$Cu$_9$ glasses.

Within the composition range Si$_{15}$Te$_{84}$Cu$_1$ and Si$_{15}$Te$_{75}$Cu$_{10}$, band C shifts from 140.89 to 139.53 cm^{-1}. The peak intensity increases significantly with the inclusion of Cu. This blue shift with increasing Cu content in the sample suggests: 1) frequency of phonons interacting with the incident photon is increased; 2) tensile strain in the material lattice is increased; and 3) phonon in the material gains more energy from photons [33]. The blue shift can be anticipated to occur at RPT. The gradual addition of Cu atoms in the range of composition causes cleavage of Te$_n$ chains which in turn blue-shifts the optical-absorption edge and leaving a bulk glass rich in Si inside which prevails metastable clusters of SiTe$_{4/2}$ units tetrahedral in shape [32].

It is worthy to mention here that, in this work, Raman is performed on as-prepared bulk glassy samples; these samples are neither switched nor annealed. The main purpose to perform Raman on as-prepared samples is to identify the signature of RPT. This does not involve crystalline phase formation but a distribution of atomic coordination, which in turn, modifies the mechanical, optical, and electrical behavior of these types of glasses. Thus, by performing Raman the samples are being classified, but how the electric pulse (switching) or temperature (annealing)-induced crystallization takes place in the microstructure of these glasses is not investigated.

F. Thickness Dependence of Threshold Voltage

The V_T of glassy samples manifesting MS is seen to exhibit proportionality to t or $t^{1/2}$, where t denotes the corresponding thickness of the glasses. Fig. 10 shows the trend of V_T with the increase in thickness for two representative Si$_{15}$Te$_{80}$Cu$_5$ and Si$_{15}$Te$_{76}$Cu$_9$ glasses.

Composition like Ge–Te–Al [34] and Ge$_{18}$Te$_{82−x}$Bi$_x$ (1 $\leq x \leq$ 4) [35] exhibits a linear dependence of V_T against t whereas Ge–As–Te [12] is found to exhibit $t^{1/2}$ relation with thickness. The representative Si$_{15}$Te$_{80}$Cu$_5$ and Si$_{15}$Te$_{76}$Cu$_9$ is found to manifest a linear variation of V_T with t.

IV. CONCLUSION

The chalcogenide glasses prepared by the melt quenching technique have been found to exhibit MS behavior followed by the manifestation of same type of behavior in amorphous thin films which are designed in sandwich geometry. The threshold voltages of thin films are found to be much lower than the bulk counterparts and hence could be suitable for

PCM applications. Both the bulk and thin films have been seen to display IP in the composition range $2 \leq x \leq 6$. Amorphous thin films do not show one to one correspondence of percolation effect when compared with bulk, and the various factors responsible for such behavior are assumed to be high field conductions, methods of deposition of thin films, difference of electrode materials in bulk and thin films. Electrical switching operation conducted at the edge of bulk samples has shown the exhibition of MS behavior and also the formation of channel-like structure, which has helped to correlate the phenomena associated with the switching effect in $Si_{15}Te_{85-x}Cu_x$ $(1 \leq x \leq 10)$ glasses. Set–Reset operation on a representative bulk samples of 0.3 mm thickness using a triangular pulse of 6 mA for Set and a rectangular pulse of 12 mA for Reset have shown a manifestation of 5–10 Set–Reset cycles repeatedly. Raman studies performed on the glassy samples have shown an exhibition of significant blue shift by Band B and C. The blue shift is exhibited by the glass at RPT. Linear variation of V_T with thickness is evident in these glasses.

REFERENCES

[1] G. W. Burr *et al.*, "Phase change memory technology," *J. Vac. Sci. Technol. B, Nanotechnol. Microelectron., Mater., Process., Meas., Phenomena*, vol. 28, no. 2, pp. 223–262, Mar. 2010, doi: 10.1116/1.3301579.

[2] H. S. P. Wong *et al.*, "Phase change memory," *Proc. IEEE*, vol. 98, no. 12, pp. 2201–2227, Dec. 2010, doi: 10.1109/JPROC.2010.2070050.

[3] M. H. R. Lankhorst, B. W. Ketelaars, and R. A. M. Wolters, "Low-cost and nanoscale non-volatile memory concept for future silicon chips," *Nature Mater.*, vol. 4, no. 4, pp. 347–352, Mar. 2005, doi: 10.1038/nmat1350.

[4] T. J. Moore, T. L. Gilton, and K. A. Campbell "Methods to form a memory cell with metal-rich metal chalcogenide," U.S. Patent 7 056 762, B2, Jun. 6, 2006.

[5] T. J. Carrig, "Transition-metal-doped chalcogenide lasers," *J. Electron. Mater.*, vol. 31, no. 7, pp. 759–769, Jul. 2002, doi: 10.1007/s11664-002-0233-1.

[6] E. Mytilineou, "Chalcogenide amorphous semiconductors: Chemical modification or doping," *J. Optoelectron. Adv. Mater*, vol. 4, pp. 705–710, Sep. 2002.

[7] Z. U. Borisova, *Glassy Semiconductors*. New York, NY, USA: Springer, 1981, pp. 317–343, doi: 10.1007/978-1-4757-0851-6.

[8] J. Z. Liu and P. C. Taylor, "The formal valence shell model for structure of amorphous semiconductors," *J. Non-Crystalline Solids*, vol. 114, pp. 25–30, Dec. 1989, doi: 10.1016/0022-3093(89)90057-4.

[9] K. B. Jagannatha, D. Roy, and C. Das, "Electrical switching and crystalline peak studies on $Si_{20}Te_{80-x}Sn_x$ $(1 \leq X \geq 7)$ chalcogenide bulk glasses," *J. Non-Cryst. Solids*, vol. 544, May 2020, Art. no. 120196, doi: 10.1006/j.jnoncrysol.2020.120196.

[10] G. Sreevidya Varma, D. V. S. Muthu, A. K. Sood, and S. Asokan, "Electrical switching, SET-RESET, and Raman scattering studies on $Ge_{15}Te_{80-x}In_5Ag_x$ glasses," *J. Appl. Phys.*, vol. 115, no. 16, Apr. 2014, Art. no. 164505, doi: 10.1063/1.4873237.

[11] A. C. Warren, "Thermal switching in semiconducting glasses," *J. Non-Crystalline Solids*, vol. 4, pp. 613–616, Apr. 1970, doi: 10.1016/0022-3093(70)90098-0.

[12] H. J. Stocker, C. A. Barlow, and D. F. Weirauch, "Mechanism of threshold switching in semiconducting glasses," *J. Non-Crystalline Solids*, vol. 4, pp. 523–535, Apr. 1970, doi: doi.org/10.1016/0022-3093(70)90088-8.

[13] B. J. Fernandes, K. Sridharan, P. Munga, K. Ramesh, and N. K. Udayashankar, "Memory type switching behavior of ternary $Ge_{20}Te_{80-x}Sn_x$ $(0 \leq X \leq 4)$ chalcogenide compounds," *J. Phys. D, Appl. Phys.*, vol. 49, no. 29, Jun. 2016, Art. no. 295104, doi: 10.1088/0022-3727/49/29/295104.

[14] S. Asokan and K. P. Lakshmi, "Electrical switching and other properties of chalcogenide glasses," *J. Indian Inst. Sci.*, vol. 91, pp. 319–330, Apr./Jun. 2011.

[15] B. J. Fernandes, K. Ramesh, and N. K. Udayashankar, "Electrical switching in $Si_{20}Te_{80-x}Bi_X$ $(0 \leq X \leq 3)$ chalcogenide glassy alloys," *J. Non-Crystalline Solids*, vol. 483, pp. 86–93, Mar. 2018, doi: 10.1016/j.jnoncrysol.2018.01.001.

[16] S. Prakash, S. Asokan, and D. B. Ghare, "Electrical switching behaviour of semiconducting aluminium telluride glasses," *Semicond. Sci. Technol.*, vol. 9, no. 8, pp. 1484–1488, Aug. 1994, doi: 10.1088/0268-1242/9/8/007.

[17] S. H. Hunter, A. Biennenstack, and T. M. Hayes, *The Structure of Non-Crystalline Materials*, P. H. Gaskell, Ed. London, U.K.: Taylor and Francis, 1977.

[18] L. Pauling, *Uniones Quimicas*. Kapelusz, Argentina: Buenos Aires, 1969.

[19] R.B. Heslop and K. Jones, *Inorganic Chemistry*. Amsterdam, The Netherlands: Elsevier, 1976.

[20] S. R. Gunti and S. Asokan, "Thermal and electrical switching studies on $Ge_{20}Se_{80-X}Bi_X$ $(1 \leq X \leq 13)$ ternary chalcogenide glassy system," *J. Non-Crystalline Solids*, vol. 356, nos. 33–34, pp. 1637–1643, Jul. 2010, doi: 10.1016/j.jnoncrysol.2010.06.028.

[21] D. Roy, B. Tanujit, G. S. Varma, S. Asokan, and C. Das, "Manifestation of intermediate phase in Cu doped Si-Te glasses," *J. Non-Crystalline Solids*, vol. 531, Mar. 2020, Art. no. 119863, doi: 10.1016/j.jnoncrysol.2019.119863.

[22] C. N. Murthy, V. Ganesan, and S. Asokan, "Electrical switching and topological thresholds in Ge-Te and Si-Te glasses," *Appl. Phys. A, Solids Surf.*, vol. 81, no. 5, pp. 939–942, Oct. 2005, doi: 10.1007/s00339-005-3221-5.

[23] P. Boolchand, D. G. Georgiev, and B. Goodman, "Discovery of the intermediate phase in chalcogenide glasses," *J. Optoelectron. Adv. Mater.*, vol. 3, no. 3, pp. 703–720, 2001.

[24] X. Feng, W. J. Bresser, and P. Boolchand, "Direct evidence for stiffness threshold in chalcogenide glasses," *Phys. Rev. Lett.*, vol. 78, no. 23, pp. 4422–4425, Jun. 1997, doi: 10.1103/PhysRevLett.78.4422.

[25] D. Roy *et al.*, "Electrical switching studies on $Si_{15}Te_{85-X}Cu_X$ bulk $(1 \leq X \leq 5)$ glasses," *Proc. AIP Conf.*, 1966, Art. no. 020033, doi: 10.1063/1.5038712.

[26] R. T. A. Kumar, C. Das, S. Asokan, C. Sanjeeviraja, and D. P. Padiyan, "Optical, photo-acoustic and electrical switching studies of amorphous GeS_2 thin films," *Appl. Phys. A, Solids Surf.*, vol. 115, pp. 1151–1158, Jun. 2014, doi: doi.org/10.1007/s00339-013-7976-9.

[27] C. Das, G. M. Rao, and S. Asokan, "Electrical switching behavior of amorphous $Ge_{15}Te_{85-X}Si_X$ thin films with phase change memory applications," *Mater. Res. Bull.*, vol. 49, pp. 388–392, Jan. 2014, doi: 10.1016/j.materresbull.2013.09.020.

[28] D. Roy, G. S. Varma, S. Asokan, and C. Das, "Electrical switching and optical bandgap studies on quaternary ag-doped Ge–Te–In thin films," *IEEE Trans. Electron Devices*, vol. 66, no. 4, pp. 1881–1886, Apr. 2019, doi: 10.1109/TED.2019.2898935.

[29] R. D. Gould, S. Kasap, and A. K. Ray, "Thin films," in *Springer Handbook of Electronic and Photonic Materials*, S. Kasap and P. Capper, Eds. Cham, Switzerland: Springer, 2017, pp. 645–702, doi: doi.org/10.1007/978-3-319-48933-9_28.

[30] D. M. Kroll, "Theory of electrical instabilities of mixed electronic and thermal origin. II. Switching as a nucleation process," *Phys. Rev. B, Condens. Matter*, vol. 11, no. 10, pp. 3814–3821, May 1975, doi: 10.1103/PhysRevB.11.3814.

[31] M. P. Shaw and A. Madan, *The Physics and Applications of Amorphous Semiconductors*. London, U.K.: Academic, 1988.

[32] S. Rao Gunti and S. Asokan, "Thermodynamic, Raman and electrical switching studies on $Si_{15}Te_{85-X}Ag_X$ $(4 \leq X \leq 20)$ glasses," *J. Appl. Phys.*, vol. 111, no. 3, Feb. 2012, Art. no. 033518, doi: 10.1063/1.3682759.

[33] C. N. Banwell and E. M. McCash, "Raman spectroscopy," in *Fundamentals of Molecular Spectroscopy*, 4th ed. New York, NY, USA: McGraw-Hill, 2016, pp. 100–126.

[34] J. R. Bosnell and C. B. Thomas, "Preswitching electrical properties,'forming', and switching in amorphous chalcogenide alloy threshold and memory devices," *Solid-State Electron.*, vol. 15, pp. 1261–1271, Nov. 1972, doi: doi.org/10.1016/0038-1101(72)90047-0.

[35] C. Das, G. M. Rao, and S. Asokan, "Electrical switching and thermal studies on bulk Ge–Te–Bi glasses," *J. Non-Crystalline Solids*, vol. 357, no. 1, pp. 165–169, Jan. 2011, doi: 10.1016/j.jnoncrysol.2010.09.046.

Radar maneuvering target classification using Deep learning

*Rashmi N[1],Premalatha[2],Pushpalatha[2],Pragathi[2],Vikas[2]

Department of Electronics and Communication Engineering, BMS Institute of Technology and Management[1,2,3]

*corresponding author

Abstract: *Classification of warhead and decoys is one of the challenging processing ballistic missile defense system, proper classification of target from the decoys determines the anti-ballistic missile is successful or not. Deep learning is the new technique used for classification of radar target. Convolution neural network as a deep learning technique used to classify the target in the pool of decoys. Classification results suggests that the deep learning technique is efficient and assist in swift decision making in tactical environment*

Keywords: Convolution neural organization (CNN), SqueezNet, Radar Cross Section, target classification

Many countries have spent significant resources to countermeasure research and development in order to condense the efficacy of defense mechanism of Ballistic Missile systems. [1]. Using decoys in large number, or fake targets, to ambiguous defense systems is one of the most prevalent tactics. Various decoy tactics, such as replica decoys, signature diversity decoys, and anti-simulation decoys, are currently accessible [2]. Because the weight of payload determines missile's warhead size and range, lightweight decoys are a very interesting choice in contrast to exo-atmospheric defenses. As a result, missiles can carry lightweight decoys in larger number without compromising the maximum range of the payload. [1]. Long-range BM travel on sub-orbital trajectories, and their

ranges are primarily determined by the height attained with the use of one or more boosters. In exo-atmospheric in which mid-course phase of a BM flight occurs is the longest part of the journey. Since warhead and decoys travel on comparable trajectories due to same physics a large number of decoys are deployed in midcourse phase. [3] Missiles can release accompanying debris and hardware, such as missile launch boosters, which can cause further radar interference. Since anti-ballistic missile systems are equipped with few number of interceptors, In the absence of accurate target identification, it is challenging for anti ballistic missile system to avoid the warhead from reaching its intended target .The classification challenge of Targets , which involves recognizing the warhead among a larger number of decoys and debris, is critical. So the need of the hour is to have the desired classification algorithm with high efficiency, low computational complexity and swift decision times for ground based and sea-based defense station. Furthermore, once the target has been classified and located, the seeker on the interceptor must regulate the target location on the Re-entry Vehicle (RV) for successful lethal guidance and control during the engagement [4]. It is crucial to highlight that decoy can have a significant impact on the efficiency of a defense system in two ways.

If a decoy is mistakenly categorized as a warhead (false alarm), the defense may run out of interceptor ammo too soon. Misclassification (leakage) of a weapon, on the other hand, may result in its destruction.

The ability to categorize between warheads and decoys is an issue that has received a lot of attention in the literature, thanks to the established target credentials system based on the different micro-motions exhibited by Ballistic targets. The stability of warheads is maintained normally to ensure that they organize their intended ballistic trajectories, while also displaying precession and nutation motion due to the gravitational impact of the Earth [4]. On the other hand, due to lack of spinning engine and also due to gravity decoys flip when launched by missiles [5][6]. In order to obtain the target's micromotion, Doppler and range analysis of the radar signals are used. V. Chen initially characterized the phenomenon in the Doppler domain in [7], and it is now known as the micro-Doppler effect. Neural network algorithms are extensively used for classification, in [4] Convolution Neural Network (CNN) for classifying aircraft target using LeNet was presented and compared with SVM.

The author[5] describes the classification of target using recurrent neural network (RNN).

The paper is organized as follows: Section II. Radar cross section concept is described, Section III: describes CNN for classification of target, Section IV presents simulated results, finally conclusion is drawn in section V.

Section II:

Radar Cross Section

The radar cross section (RCS) of a target is the area that the radar signal intercept. Mathematically, it is written as:

$$\sigma = \lim_{R \to \infty} 4\pi \frac{|E_s|^2}{|E_i|^2} R^2 \qquad (1)$$

Where: R=Distance travelled by the radar signal to target

E_s=Electric field strength scattered at target

E_i=Incident electric field strength at target

RCS of cylinder and cone:

The aspect angle, wavelength of operation and polarization are few parameters determines RCS of a target . When the dimension of the target exceeds the wavelength, the RCS of a conducting plate is approximated by σ.

The product is given by:

$$\sigma = G_e. Ap = \frac{4\pi Ap}{\lambda^2}.Ap = \frac{4\pi A^2}{\lambda^2} \qquad (2)$$

The approximate formula for the RCS of simple cylinder and cone is given by:

Cone RCS:

$$\sigma = \frac{\lambda^2}{16\pi} tan^4\theta;$$

where θcone half angle   (3)

Cylinder RCS:

$\sigma = \frac{2\pi a L^2}{\lambda}$ where   a: Radius; L:Length   (4)

Section III: Convolution Neural Network

CNN consists of large number of layers with various parameters that are adjusted as per design requirement. CNN is a deep learning model which is fully connected, layered structure, possessing non-linearity and pooling feature and has many degrees of repetitions inside the model. CNN comprises various architecture models such as AlexNet, Inception, Xception, SqueezeNet, MobileNet etc. Training of CNN could be a challenging with
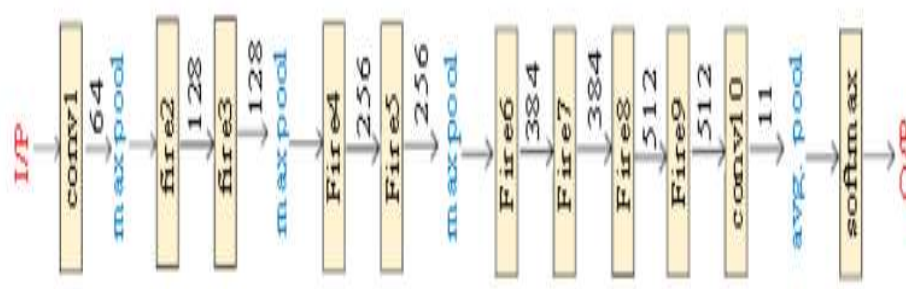

Fig 1: SqueezNet framework.

respect to time as well as resources required. So to solve this problem, we have used transfer learning with fine tuning technique through pre-trained SqueezeNet model for classification of shape of target.

Different activation functions could be used within non-linearity layer like ReLU, tanh and sigmoid. Here, ReLU has been employed after each

convolutional layer to boost the performance (Krizhevsky, A., 2012).In this work we have chosen SqueezeNet model to train for RCS return target classification which results in less computational time.

SqueezeNet CNN Architecture which is designed to classify images, therefore radar returns are converted into images by using time -frequency representation. In order to utilize the squeeze net, we transform the radar return time series results   into image

using wavelet transform. Wavel et transform captures the slowly varying signal structure and enhance the performance of the classifier. Then converted images are fed to squeezeNet's input layer and we control Squeeze Net to classify the resulting images.

Section IV: Simulation results:

In our work we created RCS synthesized data for cylinder and cone to train the learning algorithms. For cylinder with radius 1m and height of 10m with operating frequency of radar is 850MHz is considered. For cone synthesized data is generated similar to cylinder. In order to generate the training set, synthesized data is repeated for 5 times and randomly selected cylinder radii. By varying incident angle 10 motion profiles followed by 10 sinusoid curve around boresight are randomly generated to create training data set. Each motion profile consists of 701-by -50 samples. The process is repeated to generate 701 -by 100 matrix of training data set. For testing 25 profiles of each target profiles are used. The training has been performed on GPU. Its immediate result and entire training process is exhibited in the table 1. The figure 2 shows RCS of a cylinder-shaped target with dimension one meter radius, ten meter height with operating frequency of the radar is 850 MHz. Figure 3 shows radar return for one of the motion profiles of target. The plots shows variation the RCS values over time for both the incident azimuth angles and the target returns.   Table 1 exhibits the training accuracy and training loss. From table 1, we observe that loss is decreased and accuracy increased.
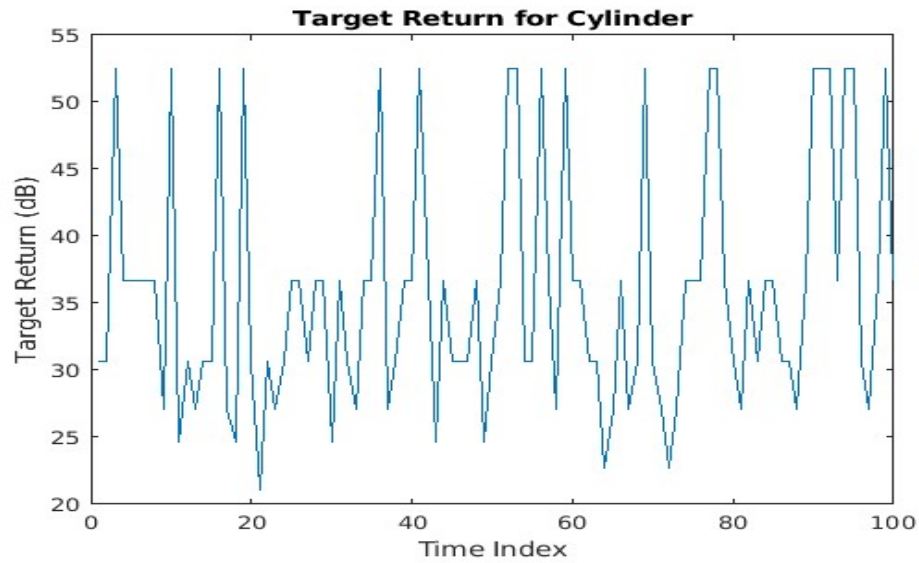
Fig2: Radar return of Cylinder.

Table 1: Result of accuracy and loss at every epoch with learning rate 1.0e^{-04}.

Epoch	Iteration	Accuracy(in %)	Loss
1	1	60	2.6639
5	50	100	0.0001
10	100	100	0.0002
15	150	100	2.2264e^{-05}
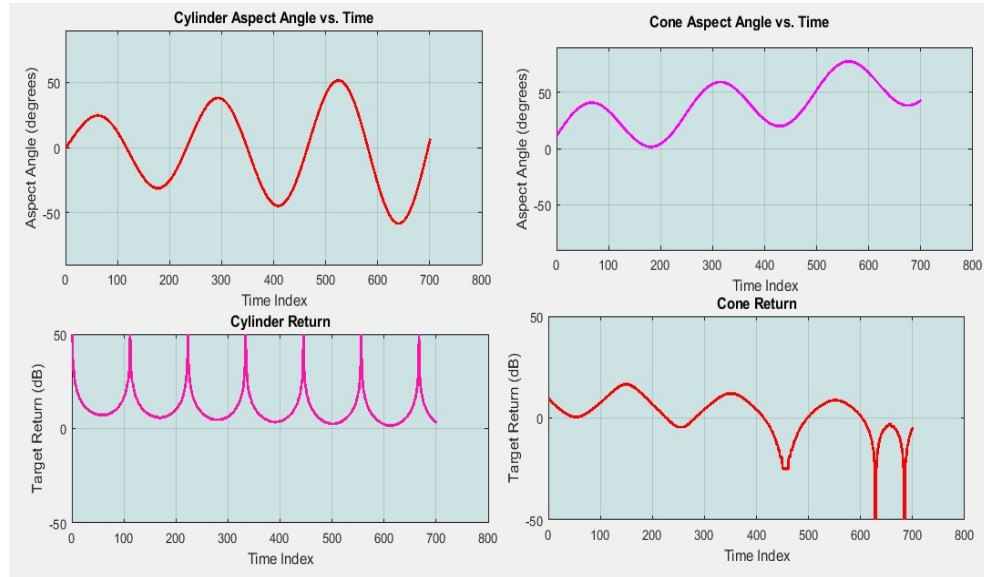
Fig 3: Shows radar return for one motion profile of cone and cylinder target
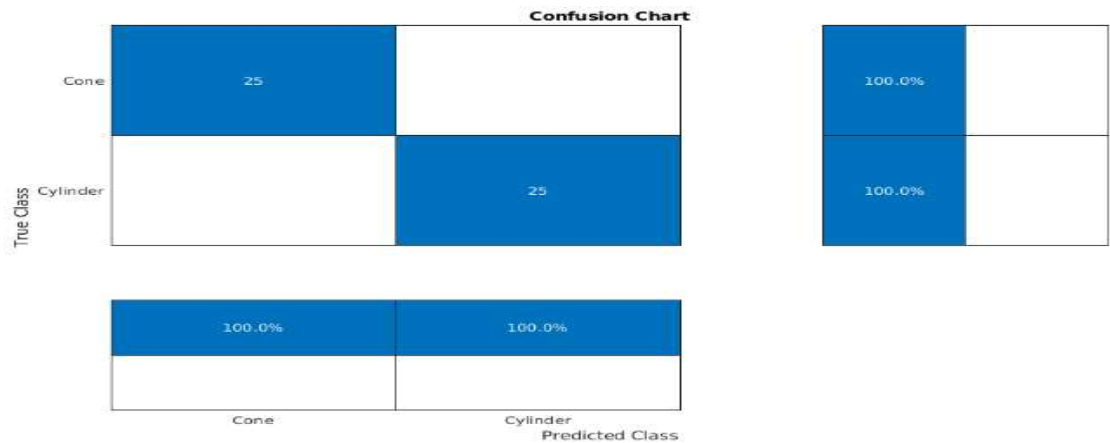
The fig 4: Shows the confusion matrix of test samples for proposed radar target classification model using the CNN/Squeeze Net

The figure 4 shows the confusion matrix of test samples for proposed radar target classification model using the CNN/Squeeze Net. All the classes contribute the best recognition accuracies i.e 100%.

Section V: Conclusion

In this paper, we synthesized the radar returns data for cylinder and cone shaped target. On synthesized data we perform classification of target using CNN. As the epoch increase the accuracy increases and

loss decreases. The results are obtained by propagating radar RCS through wavelet transform to convert into one dimensional time frequency representation and fed to SqueezeNet to classify the target. This work can be extended to accommodate real radar returns.

REFERENCES

[1] Fetter, Steve, Andrew M. Sessler, John M. Cornwall, Bob Dietz, Sherman Frankel, Richard L. Garwin, Kurt Gottfried et al. "Countermeasures: a technical evaluation of the operational effectiveness of the planned US national missile defense system." (2000).

[2] Persico, Adriano Rosario, Christos V. Ilioudis, Carmine Clemente, and John J. Soraghan. "Novel classification algorithm for ballistic target based on HRRP frame." IEEE Transactions on Aerospace and Electronic Systems 55, no. 6 (2019): 3168-3189.

[3] R. Persico, C. Clemente, D. Gaglione, C. V. Ilioudis, J. Cao, L. Pallotta, A. De Maio, I. Proudler, and J. J. Soraghan, "On model, algorithms, and experiment for micro-dopplerbased recognition of ballistic targets," IEEE Transactions on Aerospace and Electronic Systems, vol. 53, no. 3, pp. 1088– 1108, June 2

[4] Isaac Bankman, Eric Rogala, and Richard Pavek, "Laser radar in ballistic missile defense," Johns Hopkins APL Technical Digest, vol. 22, no. 3, pp. 379– 393, 200

[5] H. Gao, L. Xie, S. Wen, and Y. Kuang, "Micro-doppler signature extraction from ballistic target with micro-motions," IEEE Transactions on Aerospace and Electronic Systems, vol. 46, no. 4, pp. 1969–1982, Oct 2010.

[6] P. Lei, K. l. Li, and Y. x. Liu, "Feature extraction and target recognition of missile targets based on micro-motion," in Signal Processing (ICSP), 2012 IEEE 11th International Conference on, Oct 2012, vol. 3, pp. 1914–1919.

[7] V.C. Chen, F. Li, S.S. Ho, and H. Wechsler, "Micro-Doppler effect in radar: Phenomenon, model, and simulation study," IEEE Transactions on Aerospace and Electronic Systems, vol. 42, no. 1, pp. 2–21, Jan 2

[8] Jian Chen1, Shiyou Xu1, Zengping Chen Convolutional neural network for classifying space target of the same shape by using RCS time series IET Radar, Sonar & Navigation E-First on 4th September 2018

[9] Jun ding, bo chen, convolutional neural network with data augmentation for sar target recognition IEEE geoscience and remote sensing letters, vol. 13, no. 3, march 2016

[10]   S. Chen, H. Wang, F. Xu, and Y. Jin, "Target classification using the deep convolutional networks for sar images," IEEE Transactions on Geoscience and Remote Sensing, vol. 54, no. 8, pp. 4806–4817, Aug 2016.

[11]   G. L. Silberman, "Parametric classification techniques for theater ballistic missile defense," In Proc. Johns Hopkins Appl. Tech. Dig., 1998,vol. 19, no. 3, pp. 323–339

[12]   J. X. Zhou, Z. G. Shi, X. Cheng, and Q. Fu, "Automatic target recognition of SAR images based on global scattering center model," IEEE Trans. Geosci. Remote Sens., vol. 49, no. 10, pp. 3713–3729, Oct. 2011.

[13]   N. George and A. Theodore, "Future challenges to ballistic missile defense," IEEE Spectrum, vol. 34, no. 9, pp. 60–68, Sept. 1997.

[14]   H. Gao, L. Xie, S. Wen, and Y. Kuang, "Micro-Doppler signature extraction from ballistic target with micro-motions," IEEE Trans. Aerosp. Electron. Syst., vol. 46, no. 4, pp. 1969–1982, Oct. 2010

[15]   D. H. Hubel and T. N. Wiesel, "Receptive fields, binocular interaction and functional architecture in the cat's visual cortex," The Journal of physiology, vol. 160, pp. 106-154, 1962

[16]   K. Fukushima and S. Miyake, "Neocognitron: A self-organizing neural network model for a mechanism of visual pattern recognition," in Competition and cooperation in neural nets, ed: Springer, 1982, pp. 267-285.

[17]   Shadman Sakib1, Nazib Ahmed1#, Ahmed Jawad Kabir2@, and Hridon Ahmed1 An Overview of Convolutional Neural Network: Its Architecture and Applications.doi:10.20944/preprints201811.0546.v1

[18]   R. B. Arif, et al., "Study and Observation of the Variations of Accuracies for Handwritten Digits Recognition with Various Hidden Layers and Epochs using Convolutional Neural Network," arXiv preprint arXiv:1809.06187, 2018.

[19]   Jian Chen1✉, Shiyou Xu1, Zengping Chen convalutional neural network for classifying space tharget of the same shape by using rcs time series IET Radar Sonar Navig., 2018, Vol. 12 Iss. 11, pp. 1268-1275© The Institution of Engineering and Technology 2018

[20]   Jeffrey P. Tate, Ph.D. Santa Ana, target recognition and classification using neural networks ca 92707 october 8,2002 0-7803-7625-0/02 ww2 ieee.

[21]   Guy Ardon1,2, Or Simko1,2 and Akiva Novoselsky2 Aerial Radar Target Classification using Artificial Neural Networks January 2020

[22]   Karel Pitaš ,Tan N. Nguyen FMICW Radar Target Classification By Neural Network 978-1-7281-6468-7/20 ©2020 IEEE

[23]   Kanil Patel1,2, Kilian Rambach1 Deep Learning-based Object Classification on Automotive Radar Spectra

[24]   Bharat Sehgal, Hanumant Singh Shekhawat Automatic Target Recognition Using RecurrentNeural Networks 174102015@iitg.ac.in

[25]   Y Sun, T Fei, S Gao, N Pohl - Automatic radar-based gesture detection and classification via a region-based deep convolutional neural network ICASSP 2019-2019 IEEE …, 2019 - ieeexplore.ieee.org

[26]   Qingyuan Zhao ,Xin Du,Yaobing Lu,Aircraft Target Classification Based on CNN University of Massachusetts Amherst July 04,2020

[27]   Ding J, Chen B, Liu H W, et al. Convolutional neural network with data augmentation for SAR target recognition [J]. IEEE Geoscience and Remote Sensing Letters, 2016, 13(3): 364–368.

[28]   N.K. Ibrahim, R.S.A. Raja Abdullah and M.I. Saripan Artificial Neural Network Approach in Radar Target Classification Journal of Computer Science 5 (1): 23-32, 2009 ISSN 1549-3636

[29]   Li Wang , Jun Tang, and Qingmin Liao A Study on Radar Target Detection Based on Deep Neural Networks Department of Electronic Engineering, Tsinghua University, Beijing 100084, China VOL. 3, NO. 3, MARCH 2019

[30]   Swathi Nambari, Gottapu Sasibhushana Rao and Kolluri Sri Ranga Rao Development of Algorithm for RCS Estimation of a Perfectly Conducting Sphere using Spherical Polar Scattering Geometry International Journal of Applied Engineering Research ISSN 0973-4562 Volume 12, Number 11 (2017) pp. 2656-2661

Comparative Analysis on Fall Detection of Elderly People

Rashmi N[1]*, Nayana G[2], Ramyashree S[2],Sameeksha S[2],Supriya Singh[2]

Department of Electronics and Communication Engineering,
BMS Institute of Technology and Management

ABSTRACT:*Accidental falls have been one of the most serious issues concerned with the elderly and an important research area in the life signs domain. The injuries can have a serious effect on their lives. Furthermore, if the elderly remain lying down for a long period post-fall, then there is a chance of suffering from serious complications which increases. Such complications should be avoided beforehand. Thus, in the field of life signs, it is important to achieve a methodology to study fall detection and prevent it. This paper reviews the work as to how to detect movement of the human beings which causes a change in radar signals and then produce corresponding Doppler signatures. These Doppler estimations acquired are utilized to recognize the diverse human movement utilizingmicro-Dopplersignal. Themeasurements which are obtained will yield the results of the various human motions at different ranges and simulate thesame.*

KEYWORDS:*Doppler estimation, fall detection, Life signs, micro-Doppler signals. radar echo signal.*

1.   INTRODUCTION:

Maturing is an incredible association which fuses natural and mental changes. India which is viewed as the second-most populous country on earth and has seen a gigantic extension in the number of people in senior age and it has been projected that it would rise to around 324 million by 2050. The quantity of people over the age of 60 years is quickly developing, particularly in a crowded nation like India. The significant worry with this age gathering of individuals is the unplanned falls which happen offers ascend to various direct wounds which may cause hurt as of now or sooner rather than later. Following the fall which happens clinical mediation ought to be followed up to decrease the drawn-out results. Albeit many fall discovery calculations and gadgets have been presented, it is in its beginning phaseof improvement, it conveys extraordinary potential to be one of the main advancements sooner rather thanlater.

2.   EXISTINGFALL DETECTIONTECHNIQUES:

2.1   Micro-Dopplersignature

Four key highlights removed from the micro-Doppler signature short-time Fourier transform analysis change the investigation coming about that include vectors which are utilized as individual, pairs, trios and all togetherbeforecontributing to various sorts of classifiers dependent on the discriminant examination strategy. The

presentation of various classifiers and distinctive element combinations examined targeting distinguishing the most fitting highlights for the unarmed against equippedpersonnel grouping, just as the advantage of joining multi static information as opposed to utilizing monostatic information as itwere [1].

2.2 Linear Predictive Coding(LPC): Linear predictive coding is one of the methodswhich is used widely in the areaof signal source modeling in signal processing. LPC is put forward to bring out the features of Micro Doppler that are the combination of different frequencies. This method not only represses the time frame that is essential to capture the Doppler signatures of human motion but also retrench computing time for extracting its features for real-time processing. Some of the human activities like running, boxing, bending, walking, crawling, sitting, and more, are the data extracted from 12 humans performing different activities using Doppler radar. A supporting machine is trained using these outputs of LPC to classify the activities. Extraction of the Micro Doppler signatures includes the effect of several LPC coefficients, size of sliding time window, and decision time framewindow [2].

2.3 Convolutional NeuralNetwork:

To build up a fall identification framework, two significant issues must be addressed the first is a classification of exercises and the other is analyzing the attributes of a fall addressed by a consecutive change in the present. Based on these two clues along with speed fall event can be identified the block of input frames represent input images and label block contains information about the pose representing a different type of activitiesand the block of subtraction extracts human silhouette.The block of the convolutional neural network takes input as the result occurring from the background and their corresponding label.An event of fall can be considered as series of change of poses and therefore recognition of all these poses in consecutive frames which indicates that the fall event might have occurred within those frames.With the help of background subtraction detection of human silhouette can be done, the mean values of pixels of both RGB and Depth images are used as a threshold [3].

2.4 Wireless Sensor Network:

It is a planned and actualized care framework that can recognize and confine the fall of a human. It is coordinated into existing low-tech homes, to empower Ambient Assisted Living conditions, where programming and equipment gadgets endeavor to encourage a protected and proactive autonomous living. The proposed framework design is comprised of a versatile gadget like a cell phone or a tablet, Zigbee, and an end-client cell phone. The versatile gadget consequently identifies the fall of a subject and sends an alarm message to the nursing place through the Zigbee passageway. The proposed framework accomplished a normal fall identification precision of 99.9%. Moreover, the framework could decide the area of the fall occurrence dependent on the triangulation strategy to help clinical staff locate the patient rapidly. The proposed framework was effective as far as fall precision and force utilization, as the framework could turn out ceaselessly for a multi-month without the requirement for a battery change and could send fall and area data overmultiple times [4].Table 1 list the key focus in pioneering papers.

Table 1: Key Results of Pioneering

| S l no	METHOD	AUTHOR	TECHNIQUE ADOPTED	ADVANTAG E	DISADVANTA GE
1	**Micro-Doppler signature**	Raja syamsul, Azmirraja Abdullah,AsemAhma d Salah,AduwatiSali,Al iAlnaeb	Discriminantanalysis method	100% accuracy achieved when ST-FT features are extracted and implemented	Complex installations
2	**Linear Predictive Coding**	Rios Jesus Javier and Youngwoo kKim	Linear predictive Coding for features extraction of Micro Doppler	The classification results are over 85% accurate	Interference on this method is more
3	**Convolutional Neural Network:**	Kripesh Adhikari, Hamid Bouchachia , Hammadi Nait - Charif	CNN trained with background subtracted RGBD	74% overall accuracy was accomplished.	System should have been generic.
4	**Wireless Sensor Network**	V.Sreenithi Nithya Selvabala , A.BalajiGanesh	Wireless network architectures; PIC18 LF4620 wireless module	Precision up to 99.9%	Replacement cost is high
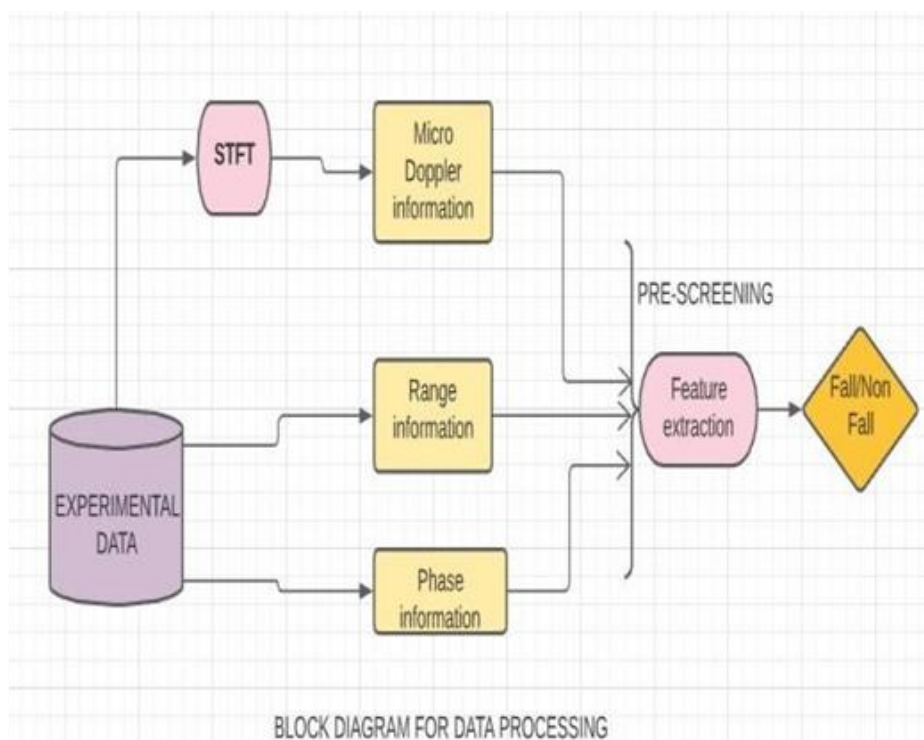
Figure 1: Data processing mechanism

.

3.   BLOCK DIAGRAM

The block diagram implementation of data processing for fall detection is described below. The trial information is first changed to a fitting area, trailed by a prescreening step which decides if a significant occasion may have happened and, provided that this is true, its time area. When an occasion is identified by the pre screener, an arrangement interaction is started to recognize whether such an occasion is a fall or a non-fall.

3.1 Experimental data

The initial stage is the collection of theavailable data set to improve the arrangementexecution.

3.1   Feature extractionmethod

To separate the highlights from the exploratory information it tends to be done in different space for example range area, stage space or Doppler space. On account of miniature Doppler data, an appropriatetime-recurrence change, like Short Time Fourier Transform (STFT), is applied totheinformation before the element extraction step to get the Time-Doppler example of the development of the objective.

3.2   Pre-screening

The element extraction step can bejoined with a pre-screening step to at select the information which can be utilized else dispose of the undesirable information.

4.   Conclusion

In this review we give a holistic point of view on fall detection systems that includes data collection, data management, feature extraction method. In particular we compare methods that rely on individual sensors and micro doppler radar signatures the survey provides a description of the different method on fall detection and it is aimed to give a comprehensive understanding of working principles and techniques, that concern fall detection systems. Sensor fusion seems to be the way forward.It provides more robust solutions in fall detection systems but come with higher computational costs when compared to those that rely on individual sensors. The challenge is therefore to mitigate the computational costs. Using micro doppler signatures method for fall detection using discriminate analysis technique 100% accuracy can be achived. In this paper focus is on the detection of elderly falls since it is one of the many factors of autonomous health monitoring

systems. While the focus here was on elderly people, the same or similar systems can be applicable to people with mobility problems.

References

[1] Alnaeb, Ali, et al. "Human Activities   Detection and Classification Based on Micro- Doppler Signatures near the Baseline of Forward Scattering Radar." 2018 International Conference on Radar (RADAR). IEEE, 2018.

[2] Javier, Rios Jesus, and Youngwook Kim. "Application of linear predictive coding for human activity classification based on micro- Doppler signatures." IEEE Geoscience and Remote Sensing Letters 11.10 (2014): 1831- 1834.

[3] Adhikari, Kripesh, Hamid Bouchachia, and Hammadi Nait-Charif. "Activity recognition for indoor fall detection using convolutional neural network." 2017 Fifteenth IAPR International Conference on Machine Vision Applications (MVA). IEEE, 2017.

[4] Selvabala, V. Sreenithi Nithya, and A. Balaji Ganesh. "Implementation of wireless sensor network based human fall detection system." Procedia Engineering 30 (2012): 767- 773.

[5] Stadelmayer, Thomas, et al. "Human Activity Classification Using mm-Wave FMCW Radar by Improved Representation Learning." Proceedings of the 4th ACM Workshop on Millimeter-Wave Networks and Sensing Systems. 2020.

[6] Jin, Feng, Arindam Sengupta, and Siyang Cao. "mmFall: Fall Detection Using 4-D mmWave Radar and a Hybrid Variational RNN AutoEncoder." IEEE Transactions on Automation Science and Engineering (2020).

[7] Akagündüz, E., Aslan, M., ¸Sengür, A., Wang, H., and ˙Ince, M. C. (2017). Silhouette orientationvolumes for efficient fall detection in depth videos. IEEE J. Biomed. Health Inform. 21, 756–763. doi: 10.1109/JBHI.2016.2570300

[8] REAL ESTATE LOCATION ANALYTICS AND PROTOTYPE DESIGN USING SPATIAL MINING, Prof. Neeraj Bhargava, Vaibhav Khanna, International Journal Of Advance Research In Science And Engineering http://www.ijarse.com IJARSE, Volume No. 09, Issue No. 12, December 2020 ISSN-2319-8354(E).

[9] DESIGN AND WEIGHT OPTIMIZATIONOF A CONNECTING ROD, T. Kalyani , P Raja Naveen, International Journal Of Advance Research In Science And Engineering http://www.ijarse.com IJARSE, Volume No. 09, Issue No. 12, December 2020 ISSN-2319-8354(E).

[10] Auvinet, E., Multon, F., Saint-Arnaud, A., Rousseau, J., and Meunier, J. (2011). Fall detection with multiple cameras: an occlusion- resistant method based on 3-D silhouette vertical distribution. IEEE Trans. Inform. Technol. Biomed. 15, 290–300. doi:10.1109/TITB.2010.2087385

[11] Noury, Norbert, et al." Fall detection- principles and methods." 2007 29th Annual International Conference of the IEEE Engineering in Medicine and Biology Society. IEEE, 2007.

[12] Hyndman, Dorit, Ann Ashburn, and Emma Stack." Fall events among people with stroke living in the community: circumstances of falls and characteristics of fallers." Archives of physics.

Efforts To Overcome The Existing Impairments In Implementing Helicopter Satellite Communication System (HSCS)

Rashmi N[1*], Hamsavahini R[1], Abhinandini H R[2], Sanjana B A[2]

[1]Department of Electronics and Communication Engineering, BMS Institute of Technology andManagement, Bangalore, Karnataka, 560064,India

[2]Department of Electronics and Communication Engineering, BMS Institute of Technology andManagement, Bangalore, Karnataka, 560064,India

rashmiswamy@bmsit.in

Abstract

Communication can be broadly classified as analog and digital. These two types only deal with terrestrial networks. But, when we want communication to be achieved with very high efficiency it is very much essential to deal with satellites. Satellite Communication is becoming more advantageous day by day. Almost in all the areas, the importance of satellite communication is being recognized. The role of satellite communication in the disaster management has played an important role these days. This has gained very high popularity because of the failure or delay of the organization to detect the disaster at the right time. Satellite communication is generally useful for information gathering during the disaster [1]. Helicopters are used in the satellite communication technology and this technique is called as Helicopter Satellite Communication System (HSCS). There are few technical issues existing in this system too. Thus to overcome these issues and make it more efficient many schemes have been implemented. One such scheme is hybrid modulation with Automatic Frequency Control (AFC). [2]

1. Introduction

Communication is the most important part of our daily life. Not only owing to one's life, communication also plays a major role when it comes to avionics. Avionics is the science which deals with aircrafts and rotorcrafts. Helicopter is the most important rotorcraft which is used in Indian Army, Navy and Air force. Hence communication is very much important between the crew members of the helicopter as well as the pilots of other helicopters. Thus to facilitate such communication, Analog Communication System was introduced Analog Intercommunication System provides the direct mode of communication. The problems associated with the analog system are, deformation and the loss appearing in the signal and the proper control over the audio, that is, its volume and frequencies were not taken care of. Thus to overcome these problems, Digital Intercommunication System was introduced. Digital intercommunication system provides reliability and resistance against tightly packed voice signals. It also provides the communication of data from air to air, air to ground and ground to air using high frequency technologies such as Very High Frequency and Ultra High Frequency. The foremost aim of the Digital Intercommunication System is providing two way communication amidst the crew members (pilot and co-pilot) in the flight or helicopter (any aircraft or rotorcraft). In addition to this, the system has the provision to communicate the signals to the crew members at the ground station using a jack externally during the sustentation operations at any time necessary. This system is a chief control for all the communication equipment established on the Advanced Light

Helicopter. The Digital Intercom System connects itself with the various radio channels, warnings and other types of audio outputs. This system also provides the provision to communicate with the pilot and co-pilot of the same or different helicopters, and also with the Air Traffic Controller (ATC). [3] There were some of the problems associated with Digital Communication Systems too. Thus replacement of this old technology communication with Helicopter Satellite Communication is done in the recent years. Helicopter Satellite Communication plays a vital role in today's world because; helicopters are not just limited to one specific application. The Helicopters are used over a large area covering a wide range. They are used to acquire the information and broadcast the same in the cases of incidents such as disasters which take place over a large scale and also in the case of other emergencies. According to the present day scenario, the repeater stations are widely used to transfer the information such as video clippings and the photographs captured by the helicopter. Capturing of the videos using the helicopter is one of the best methods to obtain on scene information for relief and the management of disaster. [4.

There are few technical issues in this system. Many solutions are being provided to overcome these issues and declare the Helicopter Satellite Communication System (HSCS) as the best technique to provide communication using Helicopters.

2. Predicaments Associated With The Digital Communication System

The high requirement of bandwidth is the major problem in these systems. The proportion or the amount of information which can be transmitted per unit time depends on the bandwidth of the media usually used in the respective transmission. The analog systems have less requirement of bandwidth when compared to the digital systems. This is because, the digital signals, for example, the square pulse, delta function and many other digital signals, have many rising and falling pulses. These pulses have a lot of variation in their voltages within a very small amount of time. Thus we can say that, almost within a very few milliseconds, there will be a huge variation in its voltage. Because of this reason, these signals have very high frequency during their rise and fall. As the frequency increases, bandwidth also increases. Thus they occupy very large bandwidth. High power consumption is also a major issue in digital systems. The digital circuits consume a lot of power. This can be proved by the power dissipation in integrated circuits that is, in digital VLSI (Very Large Scale Integration). With regard to the integrated circuits, three major causes for the large power utilization are given by- utilization of the power dynamically, utilization of the power statically and utilization of the power in the short circuit. Improvement or rise in the power density and fall in the hardware measurements of the transistor are the stamp of authentication of the present day computer chips. Both of these technologies are increasing the thermal conductivity management challenge inside the chip and the package surrounding it. Both of these technologies are increasing the thermal conductivity management challenge inside the chip and the package surrounding it, as well as escalating research related progress on the materials which possess very high conductivity. The management of the thermal conductivity is largely acknowledged as an important characteristic of the computer design, due to which, the temperature significantly affects the performance of the device. Adding on to this, the lifetime of the device can be reduced extremely because of the excess thermal stresses which take place at the interfaces. [5]. Hardware complexity in the digital systems is another hurdle for the communication. The logic gates serve as an important factor in the designing of a digital circuit. These circuits put together manifest a digital system. The major problem while designing any kind of digital circuit is the hardware complexity of that circuit. When the concept of conventional gate logic is adopted in the design of the digital circuit, then the utilization of more number of components is the major challenge which it has to face. Thus these systems have more complex circuitry and require a sophisticated device making. [6]. The digital systems can be easily broken or damaged, that is, even if a small amount of digital data is unable to be recovered or if it is not interpreted exactly, the meaning of the large amount of data can change completely.

3. Role of Satellite Communication For Disaster Management

Whenever the earthbound communication resources are inadequate it means that there is a huge damage caused to the infrastructure or it can also be due to the enormous traffic. Hence we cannot depend entirely on the earthbound or the terrestrial communications for the disaster management operations. And it is also essential that for the disaster management operations the communication should be very flexible and reliable as it is the part of emergency communications. Thus everyone who desires to have such type of communications is able to experience perfect working of the communication system every time and everywhere under any circumstances. Furthermore, the broadband communications have been playing an important role in increasing the communications during the emergencies. By considering the radio links as a basis, the systems which are based on the satellites have a lot of benefits which include reliability, availability, robustness, quickness, deploy ability and accessibility from any location. The Satellite Communication System is a core network which is used for disaster management and its recovery which includes the topology of the satellite network, configuration of service and also for the technical applications. The communication provides the favorable path for managing the critical situations such as occurrence of the disaster and also helps to provide relief during these emergency situations. [7]

4. The Satellite Communication System Employing Helicopters

5. How Exactly The System Looks Like?

The continuous prevention of the radio waves being reached to the destination in the Helicopter and Satellite Communication is mainly due to the blades of the Helicopter. These helicopter blades intervene or interrupt the helicopter directly by communicating with the satellite in order to check whether the antenna is situated right under the blade of the helicopter. The time break is about 6.93 milliseconds and the amount of interception taking place is for 39.1 milliseconds. By picturing the case of the helicopter transmission, the system which usually transmits at the exact time of the blade has to be designed and this is considered to be an alternative to solve this complicated issue. [4]

6. The Designing Requirements For Helicopter Satellite Communication System

While designing the system based on Satellite Communication and Helicopter, the most important conditions to be followed were; 1. The system which has to be designed must be very small in size; 2. The system to be designed should not be bulky and heavy, instead light weight is preferred; 3. A very effective link must be established between a helicopter and the satellite for the communication to take place, and it should be irrespective of the helicopter direction; 4. The power due to transmission should not harm the pilot, co-pilot and other crew members; 5. The transmitting power should be maintained such that it should not prevent the radio waves reaching the destination due to the blockages in the rotor blade; 6. The video capturing objectives must be based on the positioning function. [4]

7. Technical Issues Ceasing To Implement The System

- Avoidance of the shadowing due to rotor blades;
- Satellite tracking;
- Polarization tracking;
- Avoidance of interference with other satellites [1]

8. Helicopter Videoing

The technique of capturing the video using helicopter is the most popular method to obtain the information for emergency communications such as disaster management. A new system has been developed which captures the information related to the disaster and immediately accounts for the transmission of this information. Such a system is called as Ku band utilization in the satellite communication with helicopter. As we have already discussed about the major issue in such type of communications, that is, interruption due to rotor blades of the helicopter, there are techniques proposed to overcome these drawbacks. These techniques are called as time diversity transmission technique which is applicable in the forward link, and the other technique namely blade synchronized transmission technique which can be applied at the return link. The estimation of the position is also implemented for on scene videoing pictures. [8]

9. Techniques To Improve The Efficiency of Hscs

In this system, there exists two major issues and they have to addressed properly; 1. Rotor blades of the helicopter is the major part which is leading to the blockage of the received signal; 2. The Doppler Shift in the system which accounts for the fluctuation of the carrier frequency. New techniques have been proposed to overcome these two drawbacks. There were many other techniques which were implemented to overcome these issues; but they were inefficient. The continuous blockage of the signal received can be cancelled by using a hybrid modulation scheme with Automatic Frequency Control (AFC). This technique is better when compared to the use of former diversity technique based on time, as this increases the amount of data transmitted per unit time; whereas AFC transmits the original amount of data per unit time. [2]. Most of the communications in the helicopter are handled by a radio which operates in very high frequency range and it has the limits such as its coverage area and the range at which it can operate. As the amount of area occupied by the relay station which is established on the ground is only around 40km (radius), there is a huge requirement for many numbers of such stations to be established. If we consider an idea of establishing relay stations, it doesn't work out in the areas covered completely with mountains or sea. Thus to overcome all these geographical limitations, the satellite itself is made as a relay station in this system. Whenever the helicopter is outside the terrestrial frequency range, the pilot operations must be undertaken with a very high efficiency because even it includes the satellite communication with the Air Traffic Controller. [2]   This system which is based on Satellite and Helicopter comprises of the two important stations; one at the ground and one at the helicopter. Communication between the two stations is considered to be bidirectionaland it can be achieved using a transponder. The transmission link is established between these two stations, that is forward and reverse link. The location of the Antenna should be such that it should not interfere with the rotor blades of the helicopter. It would be considered the best way if it is situated at the tail of the helicopter. Some of the experiments have proven that, as the antenna is located at a large distance from the main rotor blades, then the amount of interruption of signal with the blades would be avoided. [2]
    Establishing a system which communicates using the satellite technology with a helicopter is a tedious task, because the rotor blades of the helicopter is the major obstruction in such systems, as this leads to the Doppler shift. Thus because of these major impairments, many applications such as Global Positioning System, degradation in the performance of bit error rate and the discontinuity in the received signal over the recovered carrier. [2]

10. Results And Discussions

The transmission from the helicopter can automatically stop when 1. The tracking error angle becomes large 2. The receiving level becomes below the limited value 3. The posture of the helicopter inclines beyond the anticipation 4. The onboard equipment fails. The final flight test included the following ventures.

- Verification of the satellite tracking performance
- Measurement of the communication quality

- Operation check of transmission protection
- Picture transmission and position fixing. [1]

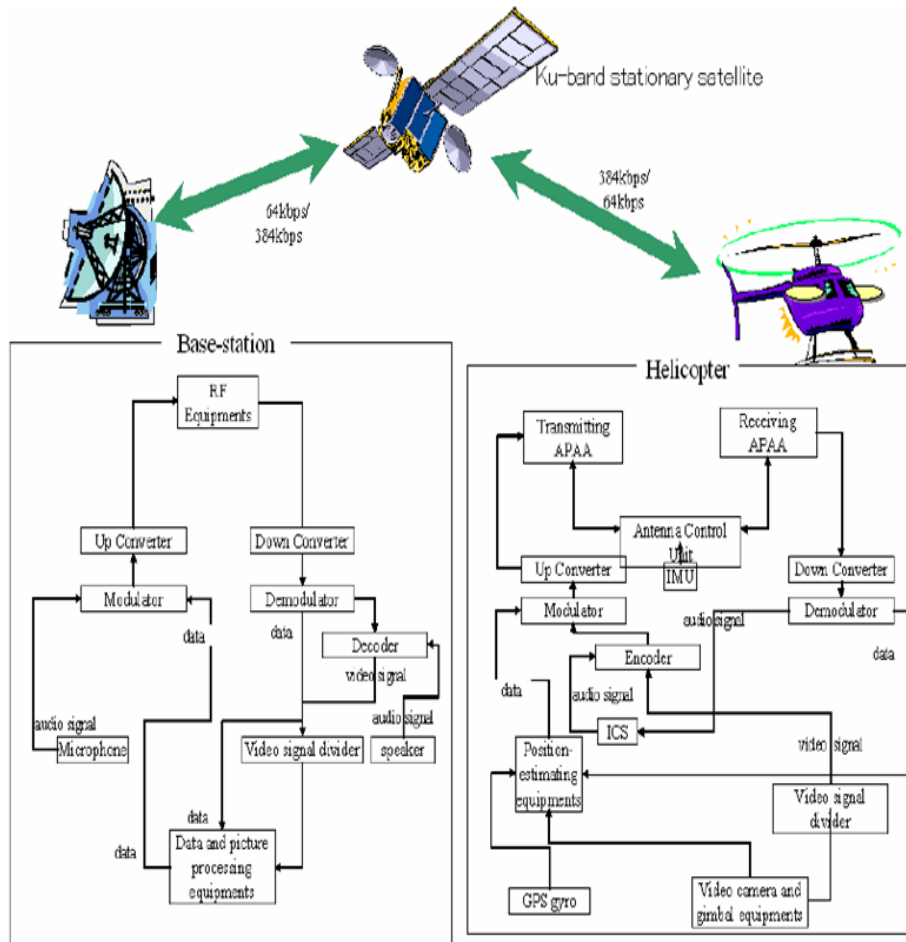The parameters of simulation in order to calculate BER is given below:

Figure1. Diagram showcasing how exactly the communication takes place by adopting the Principles of satellite communication in the Helicopter
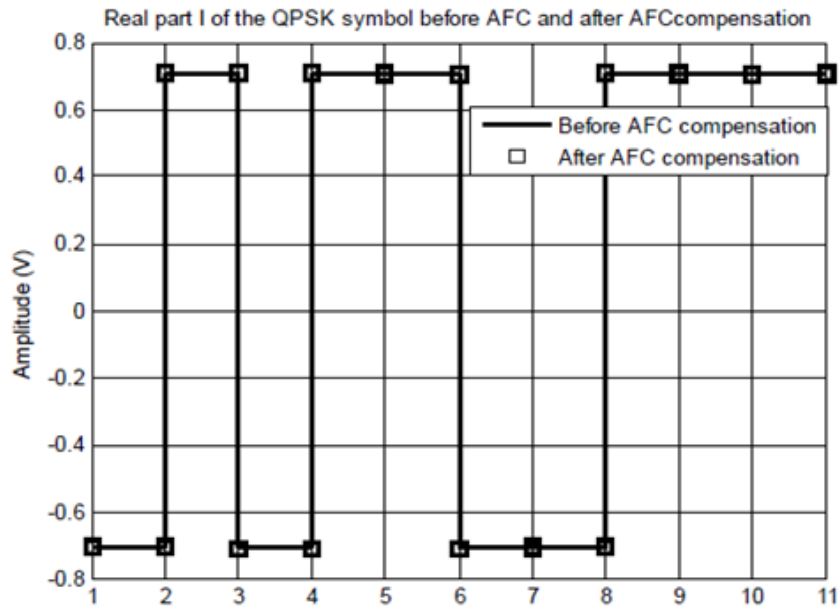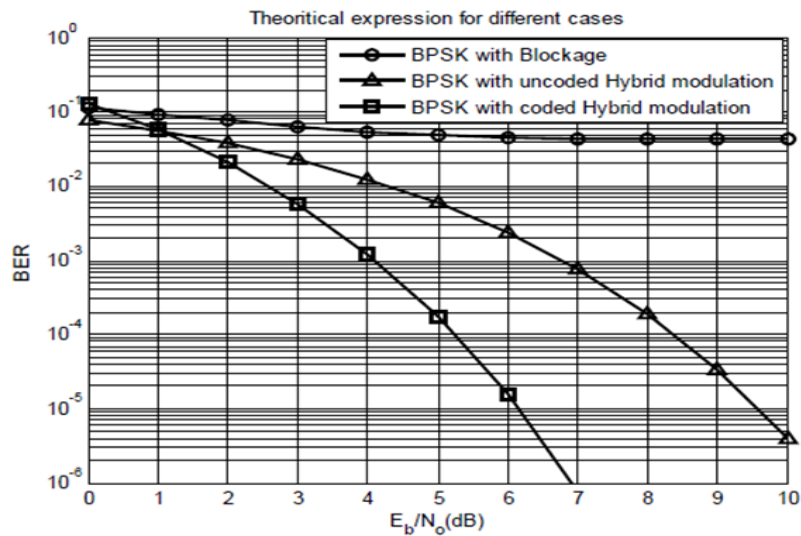
Figure 2 Phase symbol before and after AFC

Figure 3 BER behavior for different schemes

Table   I   .

Helicopter	
Rotor revolutions	$\omega = 6.45$ rev/s (387 rev/min)
Blades	$M = 4$
Maximum speed	$V = 68.4\ m/s$ (246$km/hour$)
Maximum acceleration	$A = 3.09\ m/s2$ (11.1$km/hour/s$)
Modem	
Modulation	BPSK with time diversity
Time diversity factor	$L = 2$
Information data rate	$R_b = 3600\ bit/s$
Carrier frequency	$f_C = 12.5\ GHz$
Demodulation	Coherent detection
Forward error correction	
Code	Convolutional code with
Viterbi decoding	
Rate	$R = 1/2$
Constraint length	$K = 7$

11. CONCLUSION

The difficulties in analog as well as digital communication systems led to the development of new technology in communication era which is called as Helicopter Satellite Communication System (HSCS). Thus the helicopter satellite system was developed successfully and it was also demonstrated for the first time. The key technologies of the HSCS for disaster control operations were experimentally demonstrated. The developed HSCS system provides a real time information collection and transmission directly via a satellite. Research and Development aiming at small-size, light-weight and wide-band HSCS is proceeding. Continuous research and development collaborated with disaster management bodies will be carried out. [1] This paper concludes that the Helicopter Satellite Communication system is one of the best ways to realize communication. A novel modulation scheme based on hybrid techniques for the accurate cancellation of the periodic or continuous blockage in Helicopter Satellite Communication System is implemented instead of conventional time diversity scheme. The proposed AFC scheme incorporates adaptive modulation in an efficient way. The traditional time diversity scheme increases the transmission bit rate to reduce errors, but in the novel AFC scheme, similar type of signal is made to transmit just by changing the number of time slots. [2]

REFERENCES

[1] WataruChujo, "Helicopter Satellite Communication System for Disaster Control Operations", Kashima Space Research Centre, National Institute of Information and Communications Technology, 893-1, Hirai, Kashima, Ibaraki 314-8501, Japan.

[2] Emad. A. Elaty, Sami A El-Doli, Mohammed Abd-Elnaby, Fathi E Abd El-Samie, "An Adaptive Modulation Scheme in Helicoper-Satellite Communication", The Mediterranean Journal of Electronics and Communications, Vol.9, No.3, September 2013.

[3] DankanGowda, MahaveerPenna, SatyasrikanthPalle, SaiPriya K A and Chaithra K N, "Digital Intercommunication System in Advanced Light Helicopter", IEEE, Advances in Computational Sciences and Technology, ISSN 0973-6107 Volume 10, Number 5, pp. 699-708, Research India Publications, 2017

[4] Masaki Satoh et al, Huan-Bang Li, Yoshiyuki Fujino, HiromtsuWakana, Amane Miura, Yutaka Ozaki, Hiroyuki Satou, Eiji Watanabe, Manabu Sawa, "Helicopter-Satellite Communication System Developed for Transmission of Disaster and Emergency Information", Communication Research Laboratory, American Institute of Aeronautics and Astronautics, Inc., AIAA-2003-2319, 2003.

[5] Patrick K. Schelling, Li Shi and Kenneth E. Goodson, "Managing heat for Electronics", Materialstoday, June 2005.

[6] Muhammad Khalid, ShubhankarMajumdar and Mohammad JawaidSiddiqui, "Reduction of Hardware Complexity of Digital Circuits by Threshold Logic Gates Using RTDs", Information and Communication Technology for Intelligent Systems, Smart Innovation, Systems and Technologies 106.

[7] Yang-Min Lee, Bon-lun Ku and Do-SeabAhn, "A Satellite Core Network System for Emergency Management and Disaster Recovery", ICT Standardization program of MKE (The Ministry of Knowledge Economy) [20l0-Pl-02, Standardization on PPDR communication for next generation], IEEE, 2010.

[8] Huan-Bang Li, Masaki Sato, Amane Miura, Shinichi Taira, and HiromitsuWakana, "Ku-band Helicopter Satellite Communications for on scene disaster information transmission", National Lnstitute of Information and Communications Technology 893- 1 Hirai, Kashima-shi, Ibarah 3 14-8501 Japan, 0-7803-8523-3/04/$20.0, IEEE, 2004.

Noise Identification and Removal Technique for Precise Lung Cancer Cell Detection using Active Contour Model

Shilpa Hiremath[(1)], Dr. A Shobha Rani[(2)]

BMS Institute of Technology and Management, Bangalore

Abstract

Image processing is a domain where a lot of research is required, as each other field is associated with image processing directly or indirectly. In our work we considered the medical field, where segmentation of an image is required to extract the information from the given image and illustrate anatomical structures based on the feature of the image known prior to us or based on the expert knowledge. The input image that we consider is the noisy medical image which is applied to various filtering techniques to remove the noise present in an image, after recognizing the nature of noise present in an image. Edge detection technique is applied for the filtered image to obtain the better results. Finally edge segmentation is done to the output of the edge detected output, to identify the region of tumor.

Keywords: Active Contour, Edge detection, Frost Filter, Image noise, Image Segmentation, Region Growing.

1. INTRODUCTION

Medical imaging application plays an important role in the anatomical structures. Understanding the data from the medical image is a very tedious task due to the numerous features that an image has, which leads to the complication of segmentation hence image segmentation of a medical image is an inspiring work that a researcher can take up. Information content of an image can be better analyzed and read by the user by performing the segmentation of an image. Currently lot of work is going on in the medical image processing area namely in the computer tomography image, X-ray images, Magnetic Resonance Images and Position Emission Tomography images etc. [1].

By performing the image segmentation technique on a medical image, we can obtain the regions of interest for visualization which is very necessary for the doctor to take any decision about the patient like for adjuvant therapy, planning for a surgery, model for teaching, prosthetic design, and various others [22] which can be seen in figure 1.
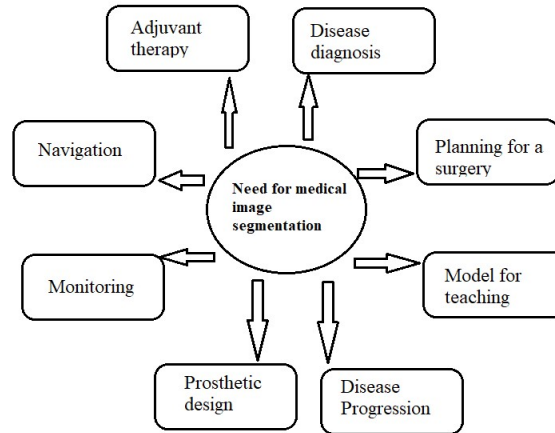
Figure 1: Need for Medical Image Segmentation

Segmentation of an image usually divides the entire image into its constituent components [35]. Image segmentation helps in extracting the selected object in each image hence Segmentation is usually followed by image representation and description and identifying the object in an image.

Before the part of segmentation methods, we apply image to filtering. In-order to retain all the fine details and edges of an image and to remove the noise we apply filtering technique. These techniques differ in their straightforward procedures applied to remove the noise or suppress the noise. Few techniques identify the noise and then apply the filtering technique and few others directly apply the filtering technique without considering the noise parameter.

In our paper we are considering the noise factor, firstly we are recognizing the nature of noise, which is present in an image, the image we have considered here is a medical image and then a suitable smoothening technique is applied which is used to remove the noise from an image and then segmentation of this image is done which helps in medical image diagnosis.

2. LITERATURE SURVEY

Lot of work have been carried out in the field of medical image analysis [35]. Here we will provide brief survey on the work being carried out in this field by various researchers by considering thresholding and region growing value.

(a) **Thresholding:**

AbuBaker, A. A., Qahwaji, R. S., Aqel, M. J., & Saleh, M. H. [23] proposed two image segmentation technique upon considering threshold values by scanning each row in an image and making a small modification to this method to find the threshold value by averaging.

Ng, H. P., Huang, S., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. [24] considered the threshold value by studying the gradient magnitude image also by considering the histogram of the gradient magnitude map the automated threshold value is obtained.

Kotropoulos, C., & Pitas, I. [25] have well thought out the use of support vector machine to segment the image and find the region of interest (ROI).

Khare, A., & Tiwary, U. S. [26] proposed a method to denoise the medical image by considering the threshold value and proposed a method, where the threshold value can adapt to change in the value based on the mean, median and standard deviation of wavelet coefficients.

(b) Region Growing:

Haider, W., Sharif, M., & Raza, M. [28] worked with breast and lung cancer where they used edge detection and segmentation as morphology. The main aim of this work is to detect the tumor region.

Siddique, I., Bajwa, I. S., Naveed, M. S., & Choudhary, M. A. [29] researchers worked on the brain images; they have considered the concept of region growing technique based on initial seed point concern.

Oghli, M. G., Fallahi, A., & Pooyan, M. [30] have considered the same factor as in [29] but proposed an automatic approach for region growing. The seed point is selected based on co-occurrence matrix.

Deng, W., Xiao, W., Deng, H., & Liu, J. [31] have studied the variations happening near the edges and the gradient value. The method turns out to be very efficient near the edges as it helps in identifying the tumor region.

Poonguzhali, S., & Ravindran, G. [32] have thought of the ultrasound images and developed a segmentation technique which is more efficient for speckle noise.

The researchers in [33] Guan, H., Li, D. Y., Lin, J. L., & Wang, T. F. have developed a method comprising of two methods namely region growing and region merging for ultrasound images.

Pan, Z., & Lu, J. [34] used Baye's concept for segmenting a medical image, the researchers developed a multistage approach which adjusted the parameter for region growing and improves the efficiency.

3. METHODOLOGY

Considering the original noisy medical image, the very first step is to apply the filtering technique depending upon the type of noise present in an image. In our work we have considered three types of noise i.e., impulsive noise, Gaussian noise, and multiplicative noise. If the noise present is impulsive then we use median filter to remove the impulsive noise, if the noise present is Gaussian then use Gaussian filter else, we apply frost filter to remove the multiplicative noise. The next step is to detect the edges of an image by performing the edge detection technique and for this edge detected output we

perform image segmentation using snake method. Figure 2 shows the flow chart of the steps performed to identify the lung cancer region in an image.
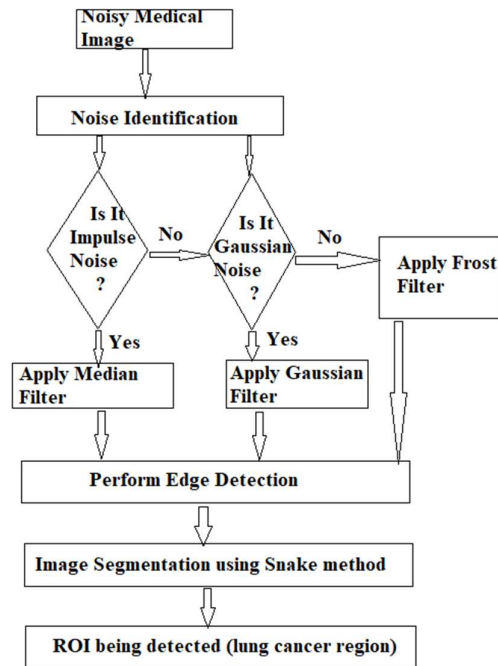

Figure 2: Flow chart of computed tomography imaging segmentation and analysis of lung cancer

4. NOISE IDENTIFICATION

Any unwanted information which is present in the image can be considered as image noise [27]. Image noise makes its presence in an image because of various reasons like atmospheric effects while capturing the image, it can be because of the sensor which is used to capture the image or even during the transmission of an image, noise can be added. Irrespective of the way the noise is being added, it is affecting the image and changing the pixel value or say intensity value. Hence identifying those pixel values which are changed because of this noise is very crucial in image processing.

There are two methods which helps in analyzing the histogram of the image in a step-by-step way namely: The Peak detection method and valley extraction method. Once the analysis of histogram of the medical image is done then this is used in extracting the valleys which are used as threshold value for segmentation process. A distinctive histogram analysis is usually done in three stages:

Stage 1: Identify the leading modes of the histogram.
Stage 2: The valleys among different modes is recognized.
Stage 3: The identified threshold is applied to the medical image for segmentation.

For identification of the noise present in an image first we need to study the behavior of the noise i.e., the variation of the gray level values of the local homogeneous regions which are affecting the image. Depending upon the image size the pixel values considered for the local homogeneous region varies, approximately it must have at least 128 pixels.

Here in our work, we have considered two step which are used to find the noise type which is affecting the medical image.

Step 1: If impulsive noise is present in an image being considered then apply filtering technique to remove it, else if.

Step 2: Additive noise is present apply filtering technique to remove additive noise or else apply filtering technique to remove the multiplicative noise.

(a) Detection of Impulsive Noise:

An input image is said to be affected by impulsive noise if the ratio of the dynamic mean of a homogeneous region to that of the dynamic max of a homogeneous region is greater than the threshold value.

(b)**Detection of Additive or Multiplicative Noise:**

An input image is said to be affected by additive noise if the gray level value of the homogeneous region is closer to the value of zero. Else it can be considered as a multiplicative noise is present in an image.

4. IDEOLOGY OF MEDICAL IMAGE SMOOTHING

The process of removing the noise from an image to retain the original image by applying a suitable filtering technique is said to be as image smoothing. The noise can appear in an image because of various reasons, and it is because of numerous non-idealities which occur in an image process. We usually say noise is present if the intensity value or say the pixel value of an image is changed usually having the values near or equal to the minimum or maximum of the allowable dynamic luminance range.

Almost all applications which deal with images has to be noise free in order for getting the correct results and having a good performance. Hence, this acts as a primary step for any image processing task to be carry forwarded. However, this is a tedious task since removal of noisy pixel should not alter the original pixel of an image which may represent the edge of an image or texture of an image etc. [3].

The image under consideration is denoised in two stage, in first stage if the noise present in a medical image is identified as impulsive noise then a suitable filter is applied usually it is a median filter which is very good for removal of impulsive noise. And, in second stage if the noise present in a medical image is additive then a filter suitable for removal of additive noise is applied i.e Gaussian filter or if the noise present is multiplicative, the image is then applied to Frost filter to remove it [5].

(a) **Impulsive Noise Smoothing:**

For better image restoration we use median filtering which is basically a non-linear filter. This median filter replaces the center pixel by the new pixel value by considering rank order information and then considering the median value of the pixel within the window. This technique of filtering using median value uses a two-dimensional window which is applied to every pixel of the input image.

$$y_{median} = \begin{cases} median \ \{x_k\}; & if \ |median \ \{x_k\} - x| > threshold \\ x & ; \quad otherwise \end{cases}$$

(b) **Additive Noise Smoothing:**

Gaussian filtering technique is used to remove the additive noise present in an image. This linear filtering technique removes the Gaussian noise present in an image by performing convolution method over the window. Digital images are usually represented in the form of a discrete value; hence the Gaussian window will also have a discrete approximation of the two-dimensional Gaussian function. The Gaussian expression can be represented as below:

$$g(x) = \frac{1}{\sqrt{2 \cdot \pi \cdot \sigma}} \cdot e^{-\frac{x^2}{2\sigma^2}}$$

(c) **Multiplicative Noise Smoothing:**

Frost filtering technique is used to remove the multiplicative noise present in an image. The window is moved over the entire image to identify and remove the multiplicative noise. This filtering technique considers the window and substitutes the center pixel with a weighted sum of the pixel values. As the distance of the center pixel value increases the value of the weighting factors decrease [5].

5. IDEOLOGY OF MEDICAL IMAGE EDGE DETECTION

Edge detection is defined as the process of detecting the edges in an image. Edges are usually having a sharp transition when compared to the other regions of an image which means that the pixel value of an edge varies very significantly from its neighboring pixel. The Edge detection technique helps in identifying the boundaries, sharp transitions, corners of an image and remove the smooth transitions which assists in understanding the image. Further processing of an image like image enhancement, registration, segmentation, or identification of an object in an image can be performed based on the results obtained from edge detection. Multidimensional filtering method can be used to perform edge detection.

M-Dimensional filtering method:

A real image which has intrinsic geometrical constructions in particular the object boundaries or edges or corners of an object in an image are more highlighted for the visual information. Hence, performing a filtering technique to visualize these vital information of edges plays a very important step.

A (discrete domain) filter can be characterized in two forms either by its impulse response which is denoted by h[n] or by transfer function denoted by H(z) = Z{h[n]}. This filtering technique performs the convolution operation on the input image and the impulse response and produces the resultant output. In the resultant output, each sample has the inner product with a translated copy of the flipped input signal and complex conjugated filter impulse response. That is, the output signal of the filtering operation is a linear combination of the impulse response h and its translated copies [7].

6. IDEOLOGY OF MEDICAL IMAGE SEGMENTATION

Image segmentation is defined as partitioning of an image into various sections. This is domain dependent where it considers the disjoint regions present in an image which are visually different, homogeneous, and meaningful in their pixel level or texture or some other characteristics and considers it to be as one group. The main objective is to streamline the depiction of an image into a more meaningful image and easy to analyze it.

(a) Region growing method:

In region growing method the very first step is to select some start point popularly known as seed point. Seed point can be selected based on application or based on user criteria like considering some grey scale region and one pixel in that range as a seed point, color criteria or grey level texture etc. The region can grow starting from these seed points and further can extend to the adjacent points based on the criteria being selected. This process is continued till all the allocated pixels are compared using the difference between the intensity value of the pixel and the regions mean value. This factor is used to identify the similarity in the regions. If the difference value is small, then the threshold value under consideration then it is considered as the belonging to the same region else to the other region. This iterative process stops if the resultant value of the difference is larger than the threshold value being considered.

4	5	6	4	5	6	7	7		4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5		5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7		4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5		6	5	4	7	6	5	4	5

0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1
3	2	1	0	1	1	0	2

Before region growing

Threshold < 3, image size = 8X8,

red indicates seed point

0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1
3	2	1	0	1	1	0	2

After region growing

Threshold < 3, image size = 8X8,

 red indicates seed point (i.e pixel value = 6);

white region indicates one segment and gray indicates the other segment

Figure 3: Region Growing Method

(c) **Split & merge image segmentation:**

In this technique, a threshold value is considered and based on this threshold we can decide which all regions in an image can be merged and considered as one segment or object. The entire image is split into blocks and then difference between the maximum and minimum pixel value in each block is calculated. This difference value is checked with the threshold value and if this value is below the threshold value then it is considered as the one region else further the block is spilt into smaller block and the process is repeated until no blocks can further be split. Threshold value plays a very important role in this technique hence choosing the right value of threshold is very significant [8]. In the below figure the image is of size 8X8, considering the entire image as one segment the maximum value is 7 and minimum value is 0 in this segment. The threshold value chosen is less than or equal to 3. Since the difference value is equal to 7 which is greater than threshold value hence the image is divided into four segments indicated by white region, green, yellow and gray region. Now, considering the grey region the max value is 7 and min value is 0 which gives the result of 7 which is more than the threshold value this leads to further splitting of the grey region which are indicated by different colored numbers and the process is repeated. Figure b shows the resultant image pixel values obtained by performing the split and merge segmentation.

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5
0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1

4	5	6	4	5	6	7	7
5	6	7	7	6	5	5	5
4	5	6	6	5	4	7	7
6	5	4	7	6	5	4	5
0	0	1	2	3	7	6	1
1	2	3	1	2	3	5	0
2	3	1	3	2	4	6	1

3	2	1	0	1	1	0	2		3	2	1	0	1	1	0	2	
Threshold<=3																	

Figure 4: Split and Merge Image Segmentation

(C) Active contour (Snake) method:

Snake method is usually implemented after the edge detection technique, where all the local regions of the object are identified in the edge detection method later which snake method is implemented which reduces the energy function. This method is used in various applications where identification of an object in an image is required, there is a need to identify the shape of the object, or there is a need for stereo matching etc. This method is said to be as active method as this reduces the energy required and has dynamic behavior. This method leads to an arbitrary shape which is close to the object contour as much as possible. Basically, we can consider this energy functions as two components: potential component and internal deformation energy.

a)   If the contour which is aligned to the image edge is small, then its potential energy.

b)   If the contour which is smooth is small, then its internal deformation energy.

The termination functional can be implemented with a gradient direction calculus in a slightly smoothed version of the image.

An Active contour can be represented as $v(s) = (x(s), y(s))$ and its energy functional is represented as follows:

$$E = \int_0^1 E_{int}(v(s))ds + \int_0^1 E_{image}(v(s))ds + \int_0^1 E_{ext}(v(s))ds$$

Where E_{int} represents the internal energy of the spline due to bending, E_{image} represents the image forces and E_{ext} represents the external constraint forces, *a(s) and B(s) are* the controlling spline energy values. Hence, the internal spline energy expression can be given as:

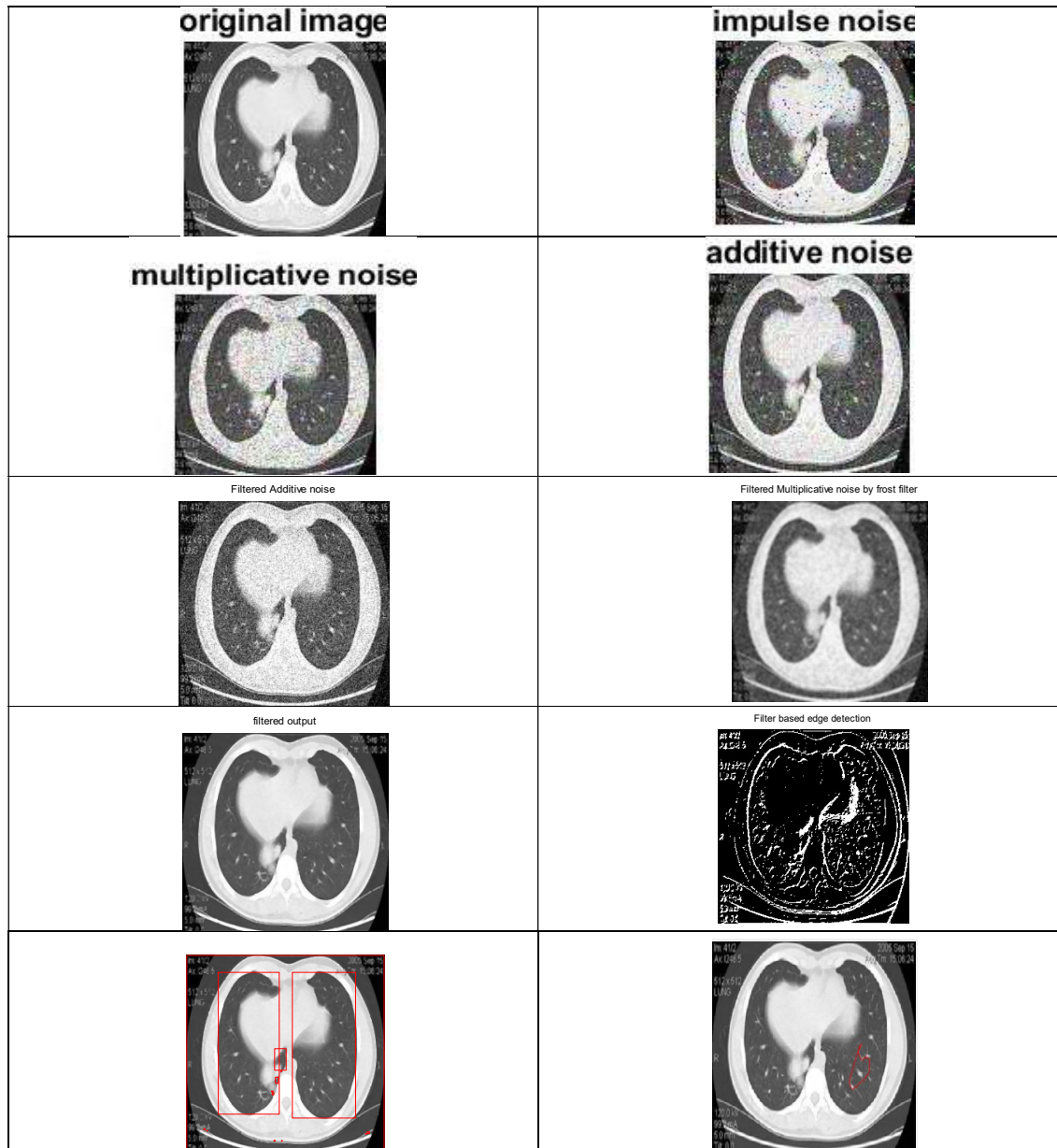$$E_{int} = \frac{(a(s)|v_s(s)|^2 + B(s)|v_{ss}(s)|^2)}{2}$$

The total energy of an image is given by the sum of weighted values of the functions of all the three energies.

$$E_{image} = w_{line}E_{line} + w_{edge}E_{edge} + w_{term}E_{term}$$

The advantageous factor of this model is that without extra processing, the object shape in an image is being identified in the piece-wise linear manner at the time of convergence.

7. RESULTS AND DISCUSSION

The objects in an image which is of our interest and the boundaries of an image namely lines and curves can be identified by the image processing technique called image segmentation. The proposed algorithm is implemented using MATLAB platform and the test results after performing various operations are as shown in figure 5.

a	b
c	d

e	f
g	h
i	j

Figure 5: (a) Original image, (b) Image with impulse noise, (c) Image with multiplicative noise, (d) Image with additive noise, (e) Filtered additive noise, (f) Filtered multiplicative noise, (g) Filtered output, (h) M-Dimensional filtered output, (I) Region growing segmentation, (j) Active counter (Snake) segmentation.

8. CONCLUSION

In this paper we have successfully identified the type of the noise and filtered it by using suitable noise removal filters. Multidimensional filter provides a powerful computational tool including natural images, medical diagnostic images. The specialty of merge-split algorithm which is to find the variation among the maximum and minimum values of a pixel within the window is like the edge detection technique. In images such as an image which has a smooth texture or an image which has less gradient, the edges play a very important role since this is the only region in the image where there is a large difference in the pixel values as compared to other regions in an image. Hence, merge and split technique is very helpful in such kind of images. This technique splits the blocks into individual pixels in which larger merged blocks appear in the interiors. This technique plays an important role and is carried out in the beginning of the image segmentation which leads to quicker region growing. From the initial contour, the object boundaries can be automatically located with the help of active contour. The most advantageous benefit of active contours is its ability to give a piece-wise linear description of the object shape at the time of convergence, without extra processing.

REFERENCES

[1] Pujar, J. H., Gurjal, P. S., Shambhavi, D. S., & Kunnur, K. S. (2010). Medical image segmentation based on vigorous smoothing and edge detection ideology. International Journal of Electrical and Computer Engineering, 5(2), 121-127.

[2] Bindu, C. H. (2009). An improved medical image segmentation algorithm using Otsu method. International Journal of Recent Trends in Engineering, 2(3), 88.

[3] Lagendijk, R. L., & Biemond, J. (2009). Basic methods for image restoration and identification. In The essential guide to image processing (pp. 323-348). Academic Press.

[4] Pujar, J. H., & Kunnur, K. S. (2010). A NOVEL APPROACH FOR IMAGE RESTORATION VIA NEAREST NEIGHBOUR METHOD. Journal of Theoretical & Applied Information Technology, 14.

[5] Frost, V. S., Stiles, J. A., Shanmugan, K. S., & Holtzman, J. C. (1982). A model for radar images and its application to adaptive digital filtering of multiplicative noise. IEEE Transactions on pattern analysis and machine intelligence, (2), 157-166.

[6] Liu, K. (2006). An Implementation of the Median Filter and Its Effectiveness on Different Kinds of Images. Thomas Jefferson High School for Science and Technology Computer Systems Lab, 2007.

[7] Do, M. N., & Lu, Y. M. (2011). Multidimensional filter banks and multiscale geometric representations. Foundation and Trends in Signal Processing, 5(3), 157-264.

[8] Kamdi, S., & Krishna, R. K. (2012). Image segmentation and region growing algorithm. International Journal of Computer Technology and Electronics Engineering (IJCTEE), 2(1), 103-107.

[9] Aly, A. A., Deris, S. B., & Zaki, N. (2011). A novel image segmentation enhancement technique based on active contour and topological alignments. arXiv preprint arXiv:1106.0371.

[10] Kass, M., Witkin, A., & Terzopoulos, D. (1988). Snakes: Active contour models. International journal of computer vision, 1(4), 321-331.

[11] Pujar, J. H., & Shambhavi, D. S. (2010, March). A novel digital algorithm for Sobel edge detection. In International Conference on Business Administration and Information Processing (pp. 91-95). Springer, Berlin, Heidelberg.

[12] Shi, J., & Malik, J. (2000). Normalized cuts and image segmentation. IEEE Transactions on pattern analysis and machine intelligence, 22(8), 888-905.

[13] Liu, K. (2006). An Implementation of the Median Filter and Its Effectiveness on Different Kinds of Images. Thomas Jefferson High School for Science and Technology Computer Systems Lab, 2007.

[14] Nafi'iyah, N., Fatichah, C., Astuti, E. R., & Herumurti, D. (2020, December). The Use of Pre and Post Processing to Enhance Mandible Segmentation using Active Contours on Dental Panoramic Radiography Images. In 2020 3rd International Seminar on Research of Information Technology and Intelligent Systems (ISRITI) (pp. 661-666). IEEE.

[15] Hsu, W. Y., Lu, C. C., & Hsu, Y. Y. (2020). Improving segmentation accuracy of CT kidney cancer images using adaptive active contour model. Medicine, 99(47).

[16] Wu, X., Tan, G., Li, K., Li, S., Wen, H., Zhu, X., & Cai, W. (2020). Deep Parametric Active Contour Model for Neurofibromatosis Segmentation. Future Generation Computer Systems, 112, 58-66.

[17] Inayat, A., Jawaid, M., & Talpur, S. (2020). Optimizing Active Contour Evolution for Robust Object Segmentation in Medical Imagery.

[18] Malligeswari, N., & Kavya, G. (2021). A Versatile Segmentation Approach For Diagnosis of Lung Cancer.

[19] Voronin, V., Zhdanova, M., Semenishchev, E., Zelensky, A., & Agaian, S. (2020). An active contour model for medical image segmentation using a quaternion framework. Electronic Imaging, 2020(10), 62-1.

[20] Malathi, M., Kesavan, S., & Praveen, K. (2021). *Segmentation of Spine Tumour Using K-Means and Active Contour and Feature Extraction Using GLCM. In AI Innovation in Medical Imaging Diagnostics (pp. 194-207). IGI Global.*

[21] Ma, J., Wang, D., Wang, X. P., & Yang, X. (2020). *A fast algorithm for geodesic active contours with applications to medical image segmentation. arXiv preprint arXiv:2007.00525.*

[22] Masood, S., Sharif, M., Masood, A., Yasmin, M., & Raza, M. (2015). *A survey on medical image segmentation. Current Medical Imaging, 11(1), 3-14.*

[23] AbuBaker, A. A., Qahwaji, R. S., Aqel, M. J., & Saleh, M. H. (2006, January). *Average row thresholding method for mammogram segmentation. In 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference (pp. 3288-3291). IEEE.*

[24] Ng, H. P., Huang, S., Ong, S. H., Foong, K. W. C., Goh, P. S., & Nowinski, W. L. (2008, August). *Medical image segmentation using watershed segmentation with texture-based region merging. In 2008 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society (pp. 4039-4042). IEEE.*

[25] Kotropoulos, C., & Pitas, I. (2003). *Segmentation of ultrasonic images using support vector machines. Pattern Recognition Letters, 24(4-5), 715-727.*

[26] Khare, A., & Tiwary, U. S. (2005). *Soft-thresholding for denoising of medical images—a multiresolution approach. International Journal of Wavelets, Multiresolution and Information Processing, 3(04), 477-496.*

[27] Hiremath, S., & Rani, A. S. (2008). *A Concise Report on Image Types, Image File Format and Noise Model for Image Preprocessing.*

[28] Haider, W., Sharif, M., & Raza, M. (2011). *Achieving accuracy in early stage tumor identification systems based on image segmentation and 3D structure analysis. Computer Engineering and Intelligent Systems, 2(6), 96-102.*

[29] Siddique, I., Bajwa, I. S., Naveed, M. S., & Choudhary, M. A. (2006, December). *Automatic functional brain MR image segmentation using region growing and seed pixel. In International Conference on Information & Communications Technology, December (Vol. 6).*

[30] Oghli, M. G., Fallahi, A., & Pooyan, M. (2010, May). *Automatic region growing method using GSmap and spatial information on ultrasound images. In 2010 18th Iranian Conference on Electrical Engineering (pp. 35-38). IEEE.*

[31] Deng, W., Xiao, W., Deng, H., & Liu, J. (2010, October). *MRI brain tumor segmentation with region growing method based on the gradients and variances along and inside of the boundary curve. In 2010 3rd International Conference on Biomedical Engineering and Informatics (Vol. 1, pp. 393-396). IEEE.*

[32] Poonguzhali, S., & Ravindran, G. (2006, December). *A complete automatic region growing method for segmentation of masses on ultrasound images. In 2006 International Conference on Biomedical and Pharmaceutical Engineering (pp. 88-92). IEEE.*

[33] Guan, H., Li, D. Y., Lin, J. L., & Wang, T. F. (2007, May). Segmentation of ultrasound medical image using a hybrid method. In 2007 IEEE/ICME International Conference on Complex Medical Engineering (pp. 644-647). IEEE.

[34] Pan, Z., & Lu, J. (2007). A Bayes-based region-growing algorithm for medical image segmentation. Computing in science & Engineering, 9(4), 32-38.

[35] 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 (pp. 1-22). IGI Global.

Digital Design of Optimized DCTQ Processor for Image Compression

Suryakanth B | B M Srinvasulu | Hussain Ghantiwala | Aviral Singh | Mallikarjun J Pathanshetty

DEPARTMENT OF ELECTRONICS AND COMMUNICATION ENGINEERING

BMS INSTITUTE OF TECHNOLOGY

EMAIL:1BY17EC034@bmsit.in

ABSTRACT: In this era of Internet of Things, wherein every 'thing' is integrated within the existing internet architecture, it becomes quite necessary that embedded computing systems process quickly, occupy less area and consume low power. This would enable them to work quickly with real time data and have a large shelf life. As such there is a need for development of optimized algorithms and their efficient implementation in hardware. This paper presents a novel architecture for obtaining, DCTQ coefficients suitable for FPGA Implementation. The design is highly parallel and pipelined so as to exploit the massive parallelism of FPGAs and occupies considerably less area with a very high processing speed.

INTRODUCTION

Digital images have been widely used today in various applications and it is growing rapidly. Video and television transmission is becoming digital. More and more digital image sequences are used in various multimedia applications. Image compression addresses the problem of reducing the amount of data required to represent a digital image. It is a process intended to yield a compact representation of an image, thereby reducing the image storage transmission requirements. Image compression basically stems from two fundamental concepts. They are

1. Redundancies reduction: It aims at removing duplication from the signal source (image/video).
2. Irrelevancy reduction: It erases that part of the signal which will be not noticed by the signal receiver, namely the Human Visual System.

The redundancy in data representation is reduced such a way that it can be subsequently reinserted to recover the original data, which is called decompression of the data. Data compression can be understood as a method that takes an input data, D and generates a shorter representation of the data (D) with a smaller number of bits compared to that of D. The reverse process is

called decompression, which takes the compressed data c (D) and generates or reconstructs the data, D. Most often the compression (coding) and decompression (decoding) systems together are called a "CODEC". The

objective of image compression is to reduce redundancy of the image data in order to be able to store or transmit data in an efficient form. Image compression can be lossy or lossless. Lossless compression is sometimes preferred for artificial images such as technical drawings, icons or comics. This is because lossy compression methods, especially when used at low bit rates, introduce compression artifacts. With the development of imaging multimedia applications (image archiving, network image transmission, document imaging, digital photography, medical imaging, remote sensing etc.) image compression requires higher performance as well as new features. Compression is an active research area in which great effort has been made to deliver new standard by providing features inexistent in previous standards, but also by providing higher efficiency for features that exist in other.

LITERATURE SURVEY

A good algorithm for compression is the hybridization of discrete cosine transform on the discrete wavelet transform coefficients. This algorithm performs much better in terms of peak-signal to noise ratio with a higher compression ratio compared to standalone DCT and DWT algorithms. This scheme is intended to use as the image/video compressor engine in medical imaging and video applications, such as telemedicine and wireless

capsule endoscopy.

A good compression algorithm of higher compression ratio based on SPIHT (Set Partition in Hierarchical Tree). This algorithm shows a higher peak signal to noise ratio (PSNR) with same compression ratio.

A good algorithm which outperforms all other previous Algorithm is the Discrete Cosine Transform which attempts to decorrelate the image and after de correlation each transform coefficient can be encoded independently without losing compression efficiency. Like other transforms, the Discrete Cosine Transform (DCT) attempts to de correlate the image data. After de correlation each transform coefficient can be encoded independently without losing compression efficiency. Discrete Cosine Transform (DCT) has emerged as the de-facto standard for image transformation in most visual systems. DCT has been widely deployed by modern video coding standards, for example, MPEG, JVT etc. This article introduces the DCT, elaborates its important attributes and analyzes its performance using information theoretic measures. Discrete Cosine Transform is widely used in such problems of digital signal processing as feature extraction, filtering, and especially, in image compression. DCT has very good energy compaction properties and near optimal performance which is closest to that of the Karhunen-Loeve transform (KLT). The JPEG and MPEG coding algorithms, based on DCT, have been established as industry standards for image compression.

In Discrete Cosine Transform, computation complexity is quite large. For reducing the computational complexity several fast DCT algorithms has developed such as, DCT based on the Residue Number System (RNS). This algorithm uses the combination of Fast Cosine Transform (FCT) algorithm and the matrix-vector multiplication (MVM). RNS based FCT MVM implementation provides a throughput improvement over the equivalent binary system up to 72%, while its advantage over the binary distributed arithmetic implementation is up to 128%. A custom architecture for Discrete Cosine Transform (DCT) using No-Instruction Set Computer (NISC) technology that is developed for fast processor customization. Using several software transformations and hardware customization, performance achievement is 10 times more, 2 times power reduction, 12.8 times energy reduction, and 3 times area reduction compared to an already optimized soft core MIPS implementation. In DCT computation, complexity increases mainly for

huge amount of multiplication. So, idea arises that, if it is multiplication less, then computation can be faster. A good technique is that which implements direct 2-D DCT approach with regular butterfly structure.

This architecture employs eight 1-D DCT processors and four post-addition stages to calculate two-dimensional DCT coefficients. Each 1-D DCT processor is designed using Algebraic Integer Encoding architecture which requires no multiplier, therefore the entire 2-D DCT design is multiplier less. Critical path is shortened by employing five pipeline stages on each compression and decompression system. A good architecture is one which uses the 2-D DCT separability property, such that the whole architecture is divided into two 1-D DCT calculations by using a transpose buffer. A new alternative of DCT is Integer Cosine Transform (ICT) used for image processing. Parallel-pipeline architecture of an 8 x 8 ICT processor for image compression whose characteristics are high throughput, low latency, reduced internal storage and 100% efficiency in all computational elements. A good architecture of an efficient implementation of a two-dimensional DCT/IDCT transforms processor via a serial-parallel systolic array that does not require transposition. A fast two-dimensional discrete cosine transform optimized for software implementation on a RISC microprocessor with an integer multiplier accumulator (MAC) unit. The number of processor cycles as well as computational error is less than that of the row-column approaches based on fast 1D DCTs.

A new processor that computes the Discrete Cosine Transform (DCT) which overcomes the major drawbacks of the original implementation resulting in a design with considerably less area consumption and higher speed. To achieve these results, a novel architecture, based on the CORDIC Circular Rotation Algorithm, is introduced; it reduces the required area by more than 60% compared to the use of standard CORDIC architectures. Furthermore, bit serial arithmetic is used, resulting in a very compact design. In order to get maximum throughput ,the processor is fully pipelined, achieving a Performance efficient even for signals as fast as HDTV.
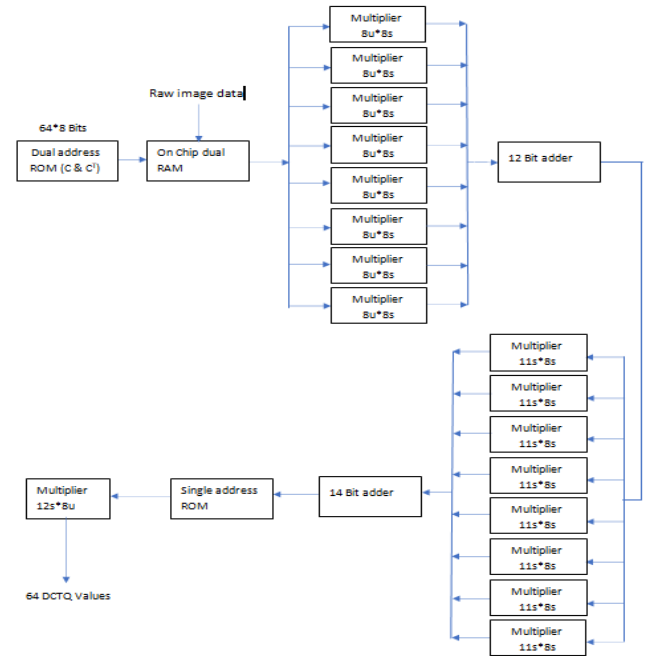
PROPOSED METHODOLOGY

The conventional memories that we use either have single address and single data output for ROM or separate read and write addresses and data bus for a dual port RAM. In real time systems such as video processing systems, these conventional memories may not be of help since the typical applications often demand access to two memory locations simultaneously, write word-wise and read column-wise, etc. These application-specific requirements, arising mainly due to the need for efficient implementation of computationally intensive algorithms.

• The ROM stores the cosine terms 2 × C instead of C in order to improve accuracy.
• Two-stage pipelining for 2 × C matrix of cosine terms to keep pace with dual RAM used in the DCTQ design.
• ROM size is 8 × 64 bits. Two locations, each of size 64 bits, can be accessed and output to the data bus, 'dout1' and 'dout2' simultaneously using the two addresses, 'addr1' and 'addr2' respectively.

This ROM, with single address and single data output, is used to store inverse quantization values. In quantization process, we need to divide the DCT coefficients by the corresponding quantization values. However, division can also be implemented as multiplication if we take the inverse of the quantization values. The dual RAM consists of two RAMs, each of which stores the image information. This information will be written from a host computer such as a PC into one of the RAMs through peripheral connect interface (PCI) bus. Initially, one of the double memory buffers, RAM 1,is filled and once it is full, the image information is written to the second RAM. While the second memory, RAM 2, is being written into, the RAM 1 will be read concurrently to process the DCTQ coefficients. If RAM 1 is in read only mode, then RAM 2 is automatically configured to the write only mode and vice-versa. The RAM is written row-wise and read column-wise. This is due to the complexity of the DCT algorithm (While reading, the RAM is accessed column-wise, since in the computation of C*X of the DCT algorithm we need to multiply a row of C matrix with the column of X (image input) matrix as explained).

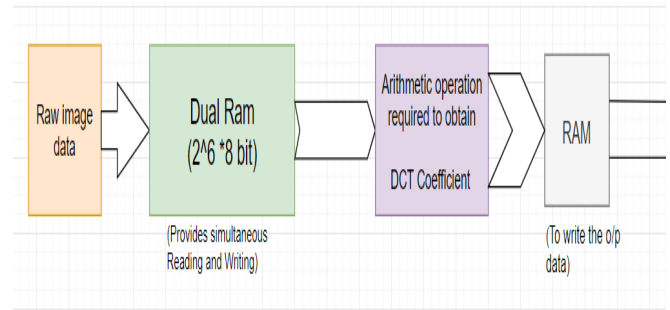Serial and Parallel adders are chosen depending on Application type and chip area. Multipliers are also Pipelined to increase the computation speed.The two-stage matrix multiplication can be implemented by parallel architecture, where in eight partial products,

which are the row vectors of C*X generated in the first stage, are fed to the second stage. Subsequently, multiplying row vector of CX by the CT matrix generate seight DCT coefficients. Quantized outputs can be obtained by dividing each of the 64 DCT coefficients by the corresponding quantization values in the Single address ROM.

Algorithm for Parallel Matrix Multiplication for DCTQ

DCT is an orthogonal transform consisting of a set of vectors that are sampled cosine functions. 2D-DCT of a block of size 8 × 8 pixels of an image is

$$
DCT = \begin{bmatrix} c_{00} & c_{01} & \ldots & c_{07} \\ c_{10} & c_{11} & \ldots & c_{17} \\ \vdots & \vdots & & \vdots \\ c_{70} & c_{71} & \ldots & c_{77} \end{bmatrix} \begin{bmatrix} x_{00} & x_{01} & \ldots & x_{07} \\ x_{10} & x_{11} & \ldots & x_{17} \\ \vdots & \vdots & & \vdots \\ x_{70} & x_{71} & \ldots & x_{77} \end{bmatrix} \begin{bmatrix} c_{00} & c_{10} & \ldots & c_{70} \\ c_{01} & c_{11} & \ldots & c_{71} \\ \vdots & \vdots & & \vdots \\ c_{07} & c_{17} & \ldots & c_{77} \end{bmatrix}
$$

defined as

$$
DCT(u,v) = \frac{1}{4} c(u) c(v) \sum_{x=0}^{7} \sum_{y=0}^{7} f(x, y) \left[\cos \frac{(2x+1)u\pi}{16} \right] \left[\cos \frac{(2y+1)v\pi}{16} \right]
$$

where $f(x, y)$ is the pixel intensity and
$c(u) = c(v) = 1/\sqrt{2}$ for $u = v = 0$ and
$= 1$ for $u, v = 1$ to 7.

The DCT can be expressed conveniently in a matrix form: DCT = C X C^T .where X is the input image matrix, C the cosine coefficient matrix, and C^T, its transpose with constants $(1/2)c(u)$ and $(1/2)c(v)$ absorbed in C and C^T matrices respectively. For a clearer understanding, the DCT may be expressed in an expanded form:

$$p_{jk} = \sum_{i=0}^{7} c_{ji}\, x_{ik}$$

$$= \begin{bmatrix} \sum_{i=0}^{7} p_{0i} c_{0i} & \sum_{i=0}^{7} p_{0i} c_{1i} & \cdots & \sum_{i=0}^{7} p_{0i} c_{7i} \\ \sum_{i=0}^{7} p_{1i} c_{0i} & \sum_{i=0}^{7} p_{1i} c_{1i} & \cdots & \sum_{i=0}^{7} p_{1i} c_{7i} \\ \vdots & \vdots & & \vdots \\ \sum_{i=0}^{7} p_{7i} c_{0i} & \sum_{i=0}^{7} p_{7i} c_{1i} & & \sum_{i=0}^{7} p_{7i} c_{7i} \end{bmatrix}$$

$$= \begin{bmatrix} p_{00} & p_{01} & \cdots & p_{07} \\ p_{10} & p_{11} & \cdots & p_{17} \\ \vdots & \vdots & & \vdots \\ p_{70} & p_{71} & \cdots & p_{77} \end{bmatrix} \begin{bmatrix} c_{00} & c_{10} & \cdots & c_{70} \\ c_{01} & c_{11} & \cdots & c_{71} \\ \vdots & \vdots & & \vdots \\ c_{07} & c_{17} & \cdots & c_{77} \end{bmatrix}$$

Where

The two-stage matrix multiplication shown above can be implemented by parallel architecture, wherein eight partial products, which are the row vectors of CX generated in the first stage, are fed to the second stage. Subsequently, multiplying row vector of CX by the C^T matrix generates eight DCT coefficients, corresponding to a row of C X C^T. While computing the $(i+1)^{th}$ partial products of CX, the ith row DCT coefficients can also be computed simultaneously since the ith partial products of CX are already available. Application of DCT on an 8 × 8 pixel block, thus, generates 64 coefficients in a raster scan order.

Quantized outputs can be obtained by dividing each of the 64 DCT coefficients by the corresponding quantization table values given in the standards [27] as per the expression:

DCTQ (u,v) = DCT (u,v) / q (u,v) ; u, v = 0 to 7

These stages can be pipelined in such a way that one DCTQ output can be generated every clock cycle.

BLOCK DIAGRAM

Raw image data → Dual Ram (2^6 *8 bit) (Provides simultaneous Reading and Writing) → Arithmetic operation required to obtain DCT Coefficient → RAM (To write the o/p data)

ADVANTAGES

The main advantages of this system are reductions in storage hardware, data transmission time, and communication bandwidth. This can result in significant cost savings. Compressed files require significantly less storage capacity than uncompressed files, meaning a significant decrease in expenses for storage. A compressed file also requires less time for transfer while consuming less network bandwidth. This can also help with costs, and also increases productivity.

CONCLUSION

In this paper we have designed a system whose inputs are pixel values of an image and the output is the computed DCTQ of those pixels. The system computes DCTQ of 64-pixel values and outputs each value at the end of each clock cycle. The core of the project is to design and verify the most efficient Arithmetic circuits which we have obtained by using pipeline concept. Pipelining increases the throughput as well as chip area. As the cost of silicon is pretty less and it is abundantly available on earth our focus in mainly on increasing the overall systems performance.

REFERENCES

[1] Kapil Ram Gavali, Poonam Kadam, Dwarkadas J "A High Throughput Architecture of DCTQ Processor suitable for FPGA Implementation utilizing Fox Algorithm".

[2] Yogesh M. Jain, Aviraj R. Jadhav and Harish V. Dixit, Akshay S. Hindole and Jithin R. Vadakoott, Devendra S. Bilaye "A Novel VLSI Design of DCTQ Processor for FPGAImplementation".

[3] Digital VLSI Systems Design by Dr. S. Ramachandran.Saurabh Chaudhury, Debashish Dash" Design and Implementation of a DCTQ Processor for Low Area, Power and High Performance"

[4] K.R. Rao and P. Yip, Discrete Cosine Transform: Algorithms, Advantages, Applications, Academic Press, New York, NY, 1990.

[5] B.G. Lee, A new algorithm to compute the discrete cosine transform, IEEE Trans. Acoust., Speech and Signal Proc., ASSP-32, pp. 1243–1245, 1984.

[6] B.G. Lee, Input and output index mappings for a prime-factor-decomposed computation of discrete cosine transform, IEEE Trans. Acoust., Speech and Signal Proc., ASSP-37, pp. 237–244, 1989.

[7] A.N. Skodras, and A.G. Constantinides, Efficient input-reordering algorithms for fast DCT, Electron. Lett., 27, pp. 1973–1975, 1991.

[8] P. Lee and F.Y. Huang, an efficient prime-factor algorithm for the discrete cosine transform and its hardware implementations, IEEE Trans. Signal Process., 42, pp. 1996– 2005, 1994.

[9] C.L. Wang and C.Y. Chen, High throughput VLSI architectures for the 1-D and 2-D discrete cosine transforms, IEEE Trans. Circuits Syst. Video Technol., 5, pp. 31–40, 1995

[10] Yung-Pin Lee, Thou-Ho Chen, Liang-Gee Chen, Mei-Juan Chen and Chung-Wei Ku, A cost-effective architecture for 8 × 8 2-D DCT/IDCT using direct method, IEEE Trans. Circuits Syst. Video Technol., 7, 1997.

[11] Yukihiro ARAI, Takeshi AGUI and Masayuki NAKAJIMA, A fast DCT-SQ scheme For images, Trans. IEICE, E71, pp. 1095–1097, 1997.

[11] Yi-Shin Tung, Chia-Chiang Ho and Ja-Lung Wu, MMX-based DCT and MC Algorithms for real-time pure software MPEG decoding, IEEE Computer Society Circuits and Systems, Signal Processing, 1, Florence, Italy, pp. 357–362, 1999.

[12] H.S. Hou, A fast recursive algorithm for computing the discrete cosine transform, IEEE Trans. Acoustics, Speech: Signal Proc., ASSP-35, pp. 1455–1461, 1987.

[13] C. Loeffler, A. Ligtenberg and G.S. Moschytz, Practical fast 1-D DCT algorithms with 11 multiplications, Proceedings of IEEE ICASSP, 2, pp. 988–991, 1989.

[14] N.I. Cho and S.U. Lee, DCT algorithms and VLSI implementations, IEEE Trans.Acoust., Speech and Signal Process., ASSP-38, pp. 121–127, 1990.

[15] N.I. Cho and S.U. Lee, Fast algorithm and implementation of 2-D discrete cosine transform, IEEE Trans. Circuits Syst., 38, pp. 297–305, 1991.

[16] Y.P. Lee, T.H. Chen, L.G. Chen, M.J. Chen and C.W. Ku, A cost-effective architecture for 8 × 8 2D-DCT/IDCT using direct method, IEEE Trans. Circuits Syst. Video Technol., 7, pp. 459–467, 1997.

[17] M. Yoshida, H. Ohtomo and I. Kuroda, A new generation 16-bit general purpose programmable DSP and its video rate application, IEEE Workshop on VLSI Signal Processing, pp. 93–101, 1993.

[18] I.Kuroda, Processor architecture driven algorithm optimization for fast 2-D DCT, IEEE Workshop on VLSI Signal Processing, VIII, pp. 481–490, 1995.

PLACEMENT MANAGEMENT SYSTEM

Akash P. Galagali
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: akashgalagali4@gmail.com

Chethan
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: chethanjkulkarni@gmail.com

Mrs. Mamatha K.R
(Assistant professor)
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: mamathakr@bmsit.in

Debanshu Biswas
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: debanshubiswas111@gmail.com

Siddesh S
Dept. of Electronics and Communication
BMS Institute of Technology & Management
(Affiliated to VTU)
Bengaluru, India
Email: 1by17ec164@bmsit.in

Abstract— As we are moving forward in life, we are facing challenges everyday and till date we are trying to have a better solution to it. With the increasing number of human population, the crave for jobs and recruitment is also increasing in heavy demand day by day and this won't stop ever. Rather it would get worse. So, looking into this factor the placement offices are facing problems to conduct the recruitment process smoothly. They are looking forward to some automated and online process so that the officers can take a break and cut down their stress of work duty. In the view of this we bring you "Placement Management System."

Keywords— ◎ Django ◎ HTML ◎ CSS ◎ JS

◎ GIT ◎ Docker ◎ Bash Scripting ◎ Nginx

I. INTRODUCTION

The project "PLACEMENT MANAGEMENT SYSTEM" is a web-based application developed for the placement department of the college in order to provide the details of its students in a database for the companies to their process of recruitment. The main purpose of the System is to avoid manual process because the manual work makes the process slow and other problems such as inconsistency & ambiguity on operations. In order to avoid this web-based placement managed system is proposed, where the student information in the college with regard to placement is managed efficiently. It intends to help fast in fast access procedures in placement related activities and ensures to maintain the details of the student. Students logging should be able to upload their personal and educational information. The key feature of this project is that it is onetime registration enabled. The placement cell calls the companies to select their students for jobs via the campus interview. The placement cell allows the companies to view the student resumes in selective manner. They can filter the student profile as per their requirement. The job details of the placed students will be provided by the administrator. The administrator plays an important role in our project. Our project provides the facility of maintaining the details of the students and gets the requested list of candidates for the company who would like to recruit the students based on given criteria.
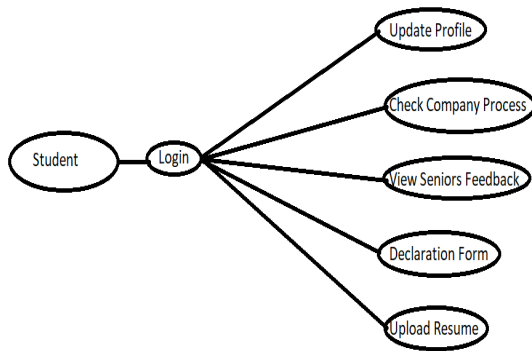
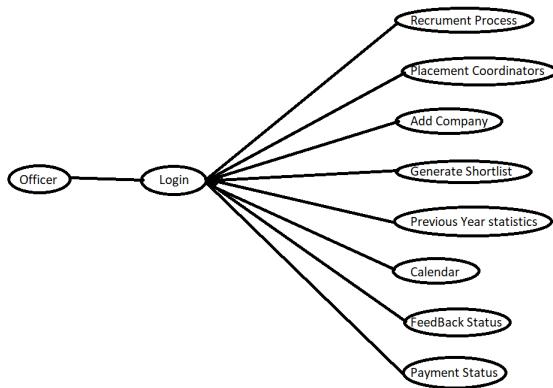**Figure 1: Student Registration Process**

Figure 2: Working at the Admin Panel

II.   PROBLEM ANALYSIS

Nowadays campus placements are conducted in all colleges. Various software and other sector companies are conducting campus selections for selecting merit candidates. When campus selections are conducted the students should provide their curriculum vitae to the concern officer for attending the campus interviews. This routine process is maintained manually, like maintenance of their resumes in papers. This can be automated by designing software.

The old manual system was suffering from many drawbacks. The PMS System is to maintain the data and make the process smooth, maintaining the data and retrieving information is very tedious and lengthy. there used to be lots of difficulties in associating any particular transaction with particular context. It is very difficult to find errors while entering the records. Once the records were entered it was very difficult to update records.

Main reason behind it is a lot of information to be maintained and managed. For this reason we provide features in PMS that can potentially automate the whole system and reduce the errors and manual work.

In Colleges, Training and Placement Officers(TPO) have to manage the student profile and data for training and placement process manually. Placement officers have to collect the information from companies and notify the students who are clearing the company criteria and this manual process is tedious and erroneous. Any modification in the details of a student or updates required in the profile of the student has to be searched and get done manually.

III.   PROPOSED SYSTEM

Proposed system is inexpensive and no manual workload is needed. Everything, however smaller the details for the registration process are, they are going to get into the system through this Online Placement Management System.

The software used in this regards are as listed below:
- Django
- HTML
- CSS
- JS
- GIT
- Docker
- Bash Scripting
- Nginx

All these softwares are easily available in the market. If need be any college can develop it for the better conduction of placement process.

Below given are the basic functionalities of the softwares used in this project:

- Django is a high-level Python Web framework that encourages rapid development and clean, pragmatic design.

- Hypertext Markup Language (HTML) is the standard markup language for documents designed to be displayed in a web browser.

- Cascading Style Sheets(CSS) is designed to enable the separation of presentation and content, including layout, colors, and fonts.

- JavaScript(JS) conforms to the ECMAScript specification and is one of the core technologies of the World Wide Web.

- GIT is a distributed version-control system for tracking changes in any set of files, originally designed for coordinating work among programmers cooperating on source code during software development.

- Docker is a set of platform as a service (PaaS) products that use OS-level virtualization to deliver software in packages called containers.

- Bash Scripting helps in executing a shell command, running multiple commands together, customizing administrative tasks, performing task automation etc.

- Nginx is used as a web server designed for maximum performance and stability.

IV. FLOW OF THE PROCESS

In this project we will be using Django and our main framework backing up most of the authentication, communication with the frontend and the database, and effectively handling traffic, sending emails, sending SMS and much more. Front end will be handled by HTML, CSS and JS. JavaScript will be handling all the robust frontend actions like making XHR requests, visibility of elements, interactions with the users, etc. MySQL will be the database
holding all the information about the students and PTO.

EXISTING SYSTEM	PROPOSED SYSTEM
1. Manual SGPA CGPA calculation	1. Automatic CGPA SGPA calculation
2. Manual shortlisting	2. Can generate Shortlist with Criteria
3.Notifying students individually	3. Notifying students on one click
4. No option for students to update Resume	4. Can Update resume at any time
5. Collect Feedback Manually	5. Can write Feedback on site
6. Manual Reminding about Training and Placement Fees	6. Automatic Reminding about Training and Placement Fees

Figure 4: Allover Statistical Display of Registrations

Figure 5: Notification box

Applications: -

1. It can be used in colleges to ease the way of the placement campus drive.
2. On a large scale it will help in the recruitment process more efficiently.

Advantages: -

1. Easy to Install
2. No Manual workload
3. Easy to use
4. More accuracy

V. RESULTS

This paper depicts how the system will work and how smooth the conduction process is going to take place from the very beginning of registration till the recruitment of the individuals and how stress-free the process is this for the placement officers.

VI. CONCLUSION

According to various surveys done, this turns out really well for both students as well as the officers and the companies' HRs too. We hope to look forward to implementation of this system not only in our college but in all the colleges of this region as well as the country. This way the process of recruitment is going to be hassle-free.The main advantage of this paper is that it suggests the solution to a very existing

problem at a very low cost. The accuracy of the system is no doubt of high rate for good possible outputs.

VII. ACKNOWLEDGMENT

The authors of this paper would like to thank the management of BMS Institute Of Technology for providing support to execute this project.

VIII. REFERENCES

- https://www.slideshare.net/
- https://acadpubl.eu/
- http://ijcsit.com/
- http://data.conferenceworld.in/

IMPLEMENTATION OF 3-DIMENSIONAL DISCRETE WAVELET TRANSFORMS (3D-DWT) USING LIFTING BASED ALGORITHM

Hamsavahini R[1]*,
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Mamatha K R[2]
Department of Electronics and Communication Engineering
BMS Institute of Technology and Management
Bangalore, INDIA

Abstract— **In this paper we implemented 3D discrete wavelet Transform(DWT) using lifting based algorithm .The lifting based DWT architecture has the advantage of low power computational complexities transforming the signals with extension and regular data flow. We also provide a survey on these architectures for both 1-D and 2-D DWT. 3-D DWT uses a cascade combination of three 1-D wavelet transforms.**

KEYWORDS: DWT, LIFTING, IMAGE COMPRESSION

I.INTRODUCTION

Wavelet transforms is one of the most modern areas of mathematics. In the last few decades, there has been an enormous increase in the applications of wavelets, in which it is used as an analytical tool in various areas of technical research such as electronics, computers and telecommunication systems that require efficient speed, resolution & real time memory and optimization with maximum hardware utilization. A wavelet, as the name suggests, is a little piece of a wave. They are used in representing data and other functions that satisfy certain mathematical requirements. The fundamental idea behind wavelets is to analyze according to scale. Wavelet algorithms process data at different scales or resolution. Wavelet transforms offer a wide variety of useful features in contrast to other transforms such as Fourier or cosine transform. Fourier transform is a powerful tool for analyzing the components of stationary signal, but it failed for analyzing non-stationary signal, whereas wavelet transforms allows the components of a non-stationary signal to be analyzed by using a set of damped oscillatory functions known as wavelet basis. Wavelet transforms in its continuous form is represented as CWT. A discrete and fast implementation of CWT is known as standard DWT. It decomposes the signal into different sub-bands with both time and frequency and facilitates to arrive at high compression ratio.

Recent advances in implementation of image coding and progressive image transmission through DWT demand both a large number of computations and large storage features that are not desirable for either high speed or low power applications. A methodology for implementation of the above complex applications have been proposed known as lifting based DWT that often has far fewer computations than DWT. In lifting scheme, the signal is divided like a zipper and a series of convolution-accumulate operations across the divided signals is applied, since this technique applies to each of this individual divided signals. Finally, brief summaries are given in below sections to conclude the paper.

II.WAVELETS

2.1. Wavelet Definition: A 'wavelet' is a small wave which has its energy concentrated in time. It has an oscillating wavelike characteristic and is a suitable tool for transient, non-stationary or time-varying phenomena.

2.2 Wavelet Characteristics: The difference between wave (sinusoids) and wavelet is shown in figure 1.1. Waves are smooth, predictable and everlasting, whereas wavelets are of limited duration, irregular and may be asymmetric. Waves are used as deterministic basis functions in Fourier analysis for the expansion of functions (signals), which are time-invariant, or stationary. The important characteristic of wavelets is that they can serve as deterministic or non-deterministic basis for generation and analysis of the most natural signals to provide better time-frequency representation, which is not possible with waves using conventional Fourier analysis
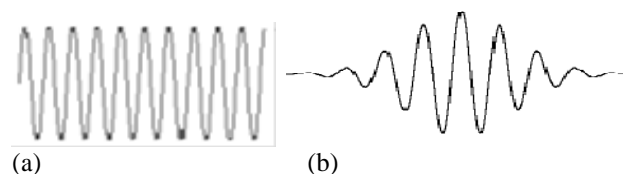

(a)   (b)

Fig.2. Representation of a (a) wave (b) wavelet

2.3 Wavelet Analysis: The wavelet analysis procedure is to adopt a wavelet prototype function, called an 'analyzing wavelet' or 'mother wavelet'. Temporal analysis is performed with a contracted, high frequency version of the prototype wavelet, while frequency analysis is performed with a dilated, low frequency version of the same wavelet. Mathematical formulation of signal expansion using wavelets gives Wavelet Transform (WT) pair, which is analogous to the Fourier Transform (FT) pair. Discrete-time and discrete-parameter version of WT is termed as Discrete Wavelet Transform (DWT).

2.4. Advantages of Wavelets:
*(a).*Wavelet analysis is an exciting new method for solving difficult problems in mathematics, physics, and engineering, with modern applications as diverse as wave propagation, data compression, signal processing, image processing, pattern recognition, computer graphics, the detection of aircraft and submarines and other medical image technology.
*(b).*Wavelets allow complex information such as music, speech, images and patterns to be decomposed into elementary forms at different positions and scales and subsequently reconstructed with high precision.
*(c).*Wavelets are a powerful statistical tool which can be used for a wide range of *applications*, namely: Signal processing, Data compression, Fingerprint verification, Blood-pressure, heart-rate and ECG analyses, DNA analysis, and protein analysis.

III.DISCRETE WAVELET TRANSFORM

3.1. Wavelet Transforms: Wavelet transform has gained widespread acceptance in speech, image and video processing, and in signal processing in general because of its attractive characteristics to represent non-stationary signals in both frequency and time domains. The wavelet transform is computed separately for different segments of the time-domain signal at different frequencies. *Multi-resolution analysis*: analyzes the signal at different frequencies giving different resolutions. Multi-resolution analysis is designed to give good time resolution and poor frequency resolution at high frequencies and good frequency resolution and poor time resolution at low frequencies. Good for signal having high frequency components for short durations and low frequency components for long duration, e.g. Images and video frames.
3.2. Discrete Wavelet Transforms: It is a wavelet transform for which the wavelets are discretely sampled. DWT of a signal x(n) is calculated by passing it through a series of filters. First the samples are passed through a low-pass filters with impulse response g(n) resulting in a convolution of the two.

$$\mathbf{Y[n]=(x*g)[n]=} \sum_{k=-\infty}^{\infty} x[k]g[n-k] \quad (1)$$

The signal is also decomposed simultaneously using a high-pass filter h(n). The outputs giving the detail coefficients(from the high pass filter)and approximate coefficients(from

low pass filter) as shown in fig.3.These 2 filters are related to each other and they are known as a quadrature mirror filters. The frequency of the signal have now been removed, half the samples can be discarded according to Nyquist'srule. The filter output's are then sub sampled by 2(g-high pass and h-low pass).

$$\mathbf{Y_{low}[n]=}\sum_{k=-\infty}^{\infty} x[k]g[2n-k] \quad (2)$$

$$\mathbf{Y_{high}[n]=}\sum_{k=-\infty}^{\infty} x[k]h[2n-k] \quad (3)$$

This decomposition has halved the time resolution since only half of each filter output characterizes the signal.Each output has the frequency band of the input so the frequency resolution has been doubled.
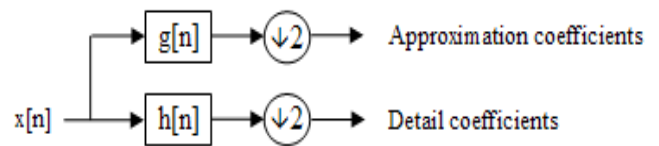

Fig.3. One-Dimensional DWT

IV.TWO-DIMENSIONAL DWT

One-Dimensional DWT can be easily extended to two dimensions which can be used for the transformation of two dimensional images. A two dimensional digital image which can be represented by a 2-D array x[n1,n2] with n1 rows and n2 columns, where n1 and n2 are positive integers.First, a one-dimensional DWT is performed on rows to get low frequency 'L' and high frequency 'H' components of the image.The fundamental block diagram of a 2-D DWT is as shown in fig 4a.
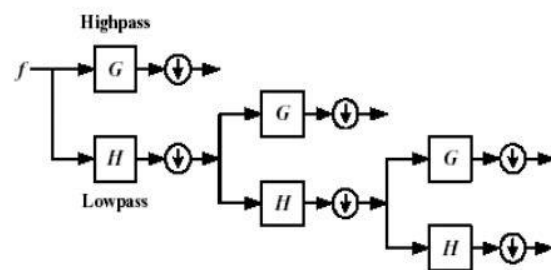

Fig.4 Two-Dimensional DWT

At each scale, an image f(x,y) is decomposed into an approximation image a_j of a low-pass band, and three detail images d_j^x, d_j^y, d_j^{xy} corresponding to a horizontal high-pass

band d_j^x, a vertical high-pass band d_j^y, and a diagonal high-pass band d_j^{xy} as given below:

$$a_j(x,y) = \left\langle f(x,y), \phi_j(x)\phi_j(y) \right\rangle$$
$$d_j^x(x,y) = \left\langle f(x,y), \psi_j(x)\phi_j(y) \right\rangle$$
$$d_j^y(x,y) = \left\langle f(x,y), \phi_j(x)\psi_j(y) \right\rangle$$
$$d_j^{xy}(x,y) = \left\langle f(x,y), \psi_j(x)\psi_j(y) \right\rangle$$

Compare with single dimension case

$$Wf(b,s) = \int f(t)\frac{1}{\sqrt{s}}\psi\left(\frac{t-b}{s}\right)dt = f*\overline{\psi}_s(b)$$

$$= <f(t), \psi(t)> \qquad (6)$$

Due to down-sampling of the coefficients, the size of the image at lower level is half of that of the higher level.In two-dimensional image analysis, the three wavelets are given by:

$$\psi^1(x,y) = \varphi(x)\psi(y),$$
$$\psi^2(x,y) = \psi(x)\varphi(y)$$
$$\psi^3(x,y) = \psi(x)\psi(y) \qquad (7)$$

Extract image details at different scales and orientations. At each scale, we end up with three "detail" images:(a) Low-pass filtered in the x-direction and high-pass filtered in the y direction, $\psi^1(x,y) = \varphi(x)\psi(y)$, yielding detail D^x. (b)Low-pass filtered in the y direction and high-pass filtered in the x direction, $\psi^2(x,y) = \psi(x)\varphi(y)$ yielding detail D^y, and, (c)finally, high-pass filtered in both x and y directions yielding detail D^{xy}

I$_2$	D$_2^x$	D$_1^x$
D$_2^y$	D$_2^{xy}$	
D$_1^y$		D$_1^{xy}$

Fig.5 Components of decomposed image, three details and one approximation at each scale.

So, we have three orientations for details: Horizontal, Vertical and Diagonal. To cover the entire image using one dimensional wavelets, we consider image composed of rows or columns which are one dimensional in rows or columns. We apply wavelet transform for each row (i.e. keeping x constant but changing y) starting from the top row. Then apply wavelet transform on the results of row operations but now move column-wise starting from the left column where we use wavelet or scaling function depending on whether we want LL, LH, HL or HH signal components. Since we are using one dimensional wavelet transform at each scan of x or y direction, we have only two filters low pass 'L' and high pass 'H' where we use them in sequence in x and y directions. To obtain low frequency band in x and y direction, we need to use to low pass filters as shown in Fig.4c Similarly we use LH, HL and HH for other frequency bands.
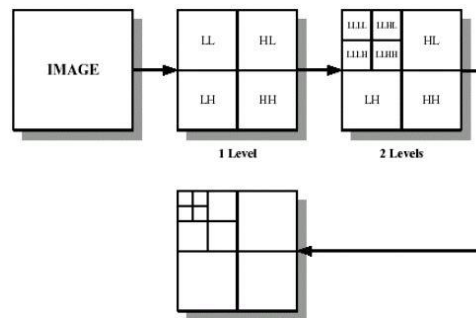

Fig.6. Different levels of decomposition of an image

V.3-DIMENSIONAL DWT

The 3D DWT can be considered as a combination of three 1D DWT in the x, y and z directions as shown in the fig. 5. The preliminary work in the DWT processor design is to build 1D DWT modules, which are composed of high-pass and low-pass filters that perform a convolution of filter coefficients and input pixels. After a one-level of 3D discrete wavelet transform, the volume of image is decomposed into HHH,HHL,HLH,LHL,LHH,LLH and LLL signals as shown in fig.5.The process results in 8-data stream. The approximate signal resulting in scalar operations only goes to the next octave of the 3D transform. It has roughly 90 percent of the total energy. Meanwhile, the seven other streams contain the detail signals. Note that though conceptual drawing of the 3D DWT for one octave has 7 filter phases, this doesn't mean that the process needs 7 physical pairs. For example, a folded architecture maps multiple filters onto one filter pair.
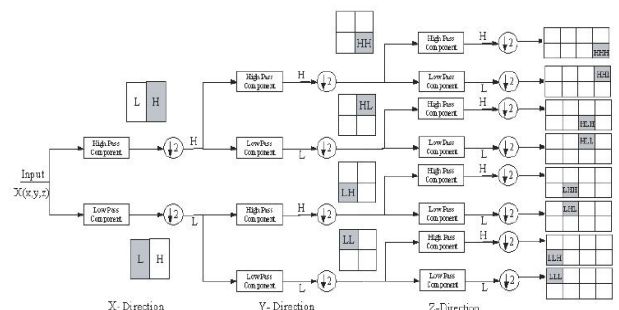

Fig.7. 3-Dimensional DWT

VI. LIFTING BASED DWT

The development of the lifting scheme was proposed for the construction of bi-orthogonal wavelets. The main feature of the lifting scheme is that all constructions are derived in the spatial domain. It does not require complex mathematical calculations that are required in traditional methods. Lifting scheme is simplest and efficient algorithm to calculate wavelet transforms. It does not depend on Fourier transforms. Lifting scheme is used to generate second - generation wavelets, which are not necessarily translation and dilation of one particular function.

The memory modules are divided into multiple banks to accommodate high computational bandwidth requirements .a number of architectures have been proposed for calculation of the convolution based DWT. The architectures are mostly folded and can be broadly classified into series (where the inputs are supplied to the filters in a serial manner) and parallel architectures (where the inputs are supplied to the filters in a parallel manner). The basic principle of lifting scheme is to factorize the poly-phase matrix of a wavelet filter into a sequence of alternating upper and lower triangular matrices and diagonal matrix. This leads to the wavelet implementation by means of a banded-matrix multiplications.

Constructing wavelets using lifting scheme consists of three steps: (a)Split step: where the data is split-up into odd and even sets. (b)Predict step: in which odd set is predicted from even set. Predict phase ensures polynomial cancellation in high pass. (c)Update phase: that will update even set using wavelet coefficient to calculate scaling function. Update stage ensures preservation of moments in low pass. Block diagram of forward Lifting scheme is as shown in fig.6a.
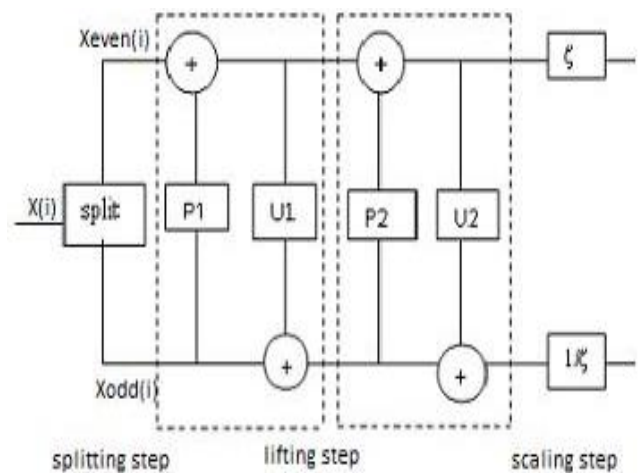
6.1. Working of lifting scheme:

The basic idea behind lifting scheme is that, it tries to use the correlation in the data to remove redundancy. First split the data into 2 sets(split phase) i.e., odd sample and even sample. Because of the assumed smoothness of the data, we predict that odd samples have a value that is closely related to their neighboring even samples. We use N even samples to predict the value of a neighboring odd value (predict phase).With a good prediction method, the chance is high that the original odd sample is in the same range as its prediction. We calculate the difference between the odd sample and its prediction and replace the odd sample with this difference.

6.2. For Image Compression:
In each lifting stage, the predicting or updating operations are carried out in the direction of image edges and textures in a local window, and are not necessarily horizontal or vertical. This adaptation can significantly reduce the signal energy of high-pass sub-bands. High angular resolution in prediction is achieved by the use of fractional pixels in prediction and update operations. The fractional pixels can be calculated by any existing interpolation method. In order to guarantee perfect reconstruction, the predicted and updated samples are always in integer pixel positions.

Fig.8. Block diagram of forward Lifting scheme

Fig.9. Architecture of DWT based on lifting scheme

6.3. Reasons for the choice of Lifting scheme

We have used lifting scheme of wavelet transform for the digital speech compression, because lifting scheme is having following advantages over conventional wavelet transform technique. (1)It allows a faster implementation of the wavelet transform. It requires half number of computations as compare to traditional convolution based discrete wavelet transform. This is very attractive for real time low power applications. (2)The lifting scheme allows a fully in-place calculation of the wavelet transform. In other words, no auxiliary memory is needed and the original signal can be replaced with its wavelet transform. (3)Lifting scheme allows us to implement reversible integer wavelet transforms. In conventional scheme it involves floating point operations, which introduces rounding errors due to floating point arithmetic. While in case of lifting scheme perfect reconstruction is possible for loss-less compression.

VII. RESULTS

Consider an color RGB image of size 225x225 and convert it into a gray scale image which is then resized into 256x256 as shown in fig 10.
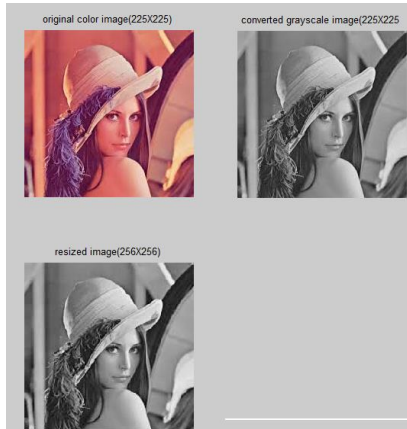

Fig 10. Resized Input Image

We extract the required wavelet co-efficients and combine it with corresponding filter coefficients. Divide the image coefficients into odd and even indexed values. Convolute filter coefficients with image coefficients taking two adjacent pixel values at a time. After performing predict and update step column wise first and row wise next we get 128x128 pixel image which is the output of 1D DWT as shown in fig b.

fig 11. Output Of 1D DWT

The output of 1D DWT is given as an input to the cascaded filters of 2D DWT where the image is compressed further that leads to 64x64 pixel image as shown in fig c.

Fig 12. Output of 2D DWT

The output of 2D DWT is given as an input to the cascaded filters of 3D DWT. The final image is the compressed image(32x32) of the original input image due to normalization/down sampling which is shown in fig d.

Fig 13. Output of 3D DWT

VII.COMPARISIONS

At first, the comparision between different stages of DWT implemented in this paper using lifting based algorithm is tabulated below:

Levels of dwt	Extent of relative compression
1st level DWT	2:1
2nd level DWT	4:1
3rd level DWT	8:1

Next, the time difference in computation speed between normal dwt and discrete wavelet transforms implemented using lifting based algorithm is shown below in fig 14 and fig 15.

As we have already seen the advantages of using lifting based algorithm that it takes less time to compute than normal DWT. As we see in fig 14. that normal dwt takes a total time of 1.108s to process an image. Where as lifting based algorithm shown in fig 15. Takes the total time of 1.044s to process same image ,there is a time difference of 0.064s for just one image but in real world we deal with millions of images at a time where time constraint is more effective and hence lifting based occupies high priority.

Fig 14. Time simulation results of normal DWT

Fig 15. Time simulation results of lifting based DWT

CONCLUSION

In this paper, we implemented the lifting based implementations of 1-dimensional, 2-dimensional and 3-dimensional Discrete Wavelet Transform. We briefly described the principles behind the lifting scheme in-order to better understand the different implementation styles and structures.

The result of this implementation is the compressed image whose relative compression is 8 times the input image as shown in simulation figures.

REFERENCES

[1] Senthilkumar.M ,Uma .S " High speed 3d DWT VLSI architecture for image processing using lifting based wavelet transform".International journal of communication and computer technologies. Vol 01-no.36.

[2] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecturefor Image Processing".IMACST: vol 3.

[3] Kishore Andra, Chaitali Chakrabarti, member IEEE, and Tinku Acharya, sentor member IEEE" A VLSIfor lifting based forward and inverse wavelet transform".IEEE transactions on signal processing. Vol 50,no. 4.

[4] Malay Ranjan Tripathy, Kapil Sachdeva and Rachid Talhi "3D Discrete wavelet transform VLSI architecture for Image processing" progress in Electromagnetics research symposium proceedings Moscow, Russia.

[5]Patrick Longa " An Optimized architecture for 2D discrete wavelet transform on FPGA's usingdistributed arithmetics 2006.

[6] O. Rioul and M. Vetterli, "Wavelets and Signal Processing," IEEE Signal Processing, vol. 8, issue: 4,pp. 14-38, Oct. 1991.

[7] ISO/IEC. International Standard, 15444-1: 2000(E), JPEG2000 Image Coding System – Part I Core coding system.

[8] C. Chakrabarti, M. Vishwanath, and R. M. Owens, "Architectures for wavelet transforms: A survey," J. VLSI Signal Process., vol. 14, pp.171–192, 1996.

[9] W. Sweldens, "The lifting scheme: A construction of second-generation wavelets," SIAM J. Mathematical Analysis, vol. 29, no.2, pp. 511–46, 1997.

[10] P. Chen, "VLSI implementation for one-dimensional multilevel lifting-based wavelet transform," IEEE Trans. on Computers, vol.53, no.4, pp.386-398, April 2004.

[11] Niladri Ghosh and K. Pradeep Vinaik "A Lifting based wavelet transform VLSI Architecture for Image Processing".IMACST: vol 3.

[12] N.D. Zervas, G.P. Anagnostopoulos, V. Spiliotopoulos,Y. Andreopoulos, and C.E. Goutis, "Evaluation of Design Alternatives for the 2-D Discrete Wavelet Transform," IEEE Trans. Circuits and Syst. Video Tech., Vol. 11, No. 12, pp. 1246-1262, December 2001.

[13] C. Chrysafis and A. Ortega, "Line-Based, Reduced Memory, Wavelet Image Compression," IEEE Trans.Circuits and Syst. Video Tech., Vol. 9, No. 3, pp.378-389, March 2000.