Design and Analysis of an Algorithm for Breast Tumor Segmentation in Mammogram and Ultrasound Images

Shwetha S. V., Dharmanna L., Basavaraj S. Anami, and Mohamed Rafi

***Abstract*—The breast cancer is having high mortality rate among ladies. The current trend of identifying the cancerous tumor is by medical image processing such as mammogram and ultrasound. The heart of medical image processing lies in the segmentation of the tumor in the mammogram. Still the conventional method of segmentation faces many dynamic challenges due to various noises such as Gaussian, Pepper & Salt, and Speckle noise hence eliminating such noise and segmenting tumor with high precision from the ultrasound and mammogram images is the goal.**

The task of finding the suitable segmentation algorithm for the segmentation of different medical images with a high accuracy plays a vital role and is another challenge. Also the current segmentation algorithm misguides the actual feature extraction of the tumor and also leads to high mortality rate in ladies. In this work the medical images are enhanced to avoid the various noises using modified Gabor filter and estimated the quality of the mammograms for the segmentation with the metrics MSE (Mean Square Error) and PSNR (Poisson Signal to Noise Ratio) of the image. Several segmentation algorithms like Otsu, SRM (Statistical Region Merging), Region growing& merging and FCM (Fuzzy C means clustering) are applied on images. Along with that five edges based segmentation algorithms like Canny, Sobel, LoG(Laplacian of Gaussian), Prewitt and Roberts are also applied and their performance has been measured with respect to gold standard images of the Berkeley Database.

In this research work region growing and merging and FCM and Otsu had been adopted for tumor segmentation and region growing and merging has performed better for breast cancer tissue segmentation in the medical images. The performance of the Region growing and merging, FCM and Otsu segmentation has been measured by the metrics like F-score with the value 0.9673, 0.9573 and 0.9489 respectively. Hence these three algorithms can be adopted for the better segmentation of the breast image.

***Index Terms*—Segmentation, region growing and merging, Otsu, FCM.**

I. Introduction

This section deals about the breast and its tumors such as malignant and benign with their shape size and their characteristics had been explained. For the radiologist [1] the analysis of the only tumor is very essential in order to diagnose breast cancer and assist the pathologist to extract the tumor candidate. The conventional procedure suffers to fetch the actual required portion of mass (tumor) and that will lead for the wrong diagnosis of the disease [2]. For the detection of type of the tumor required to analyze logical geometry of the tumor and other features. The high accuracy segmentation is the heart of the diagnosis of cancer [3]. In this work we presented several region based and edge based segmentation presented and also tested with benchmark segmented results that is extracted from several medical instruments such as mammogram and ultrasound. The brief description about breast and its tumor are presented in section 1.1, in section 1.2 about cancer and in section 1.3 regarding segmentation algorithms and its importance.

A. Breast Tumors

The deposition of uncontrolled dead cells in the breast is called breast tumor. The below Fig. 1 describes the various types of tumors [4]. Fig. 1 (a) Control Mammogram (b) Benign mammogram with circular shape (c) Benign Mammogram with lobular shape (d)Malignant Mammogram with irregular shape [5] (e) Malignant Mammogram with nodular shape [6]. By observing the Fig. 1 benign and malignant tumor will have entirely different shape and sized tumor. So this is the clue for the assessment of the disease for the diagnosis. Hence better segmentation algorithms are required.

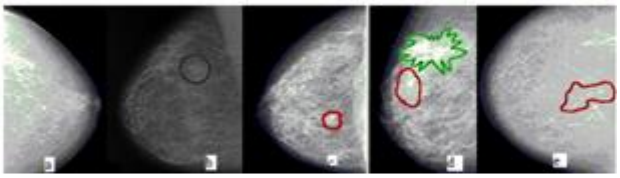

Fig. 1. a) Normal Mammogram b) Benign Mammogram with Circular tumor c) Benign Mammogram with lobular tumor d)Malignant Mammogram with irregular tumor e) Malignant Mammogram with nodular tumor.

B. Cancer

The cancer tumor initially is the formation of normal tissues in the breast. Gradually the development of micro-calcification in the breast which leads to abnormal tissues leading to cancer cells in the breast. It is also called metastasis.

C. Segmentation

The image segmentation is a process of extracting a required object from the mammogram that is consistent and homogeneous in some characteristics. Image segmentation is indeed an important process in early diagnosis of cancer tumor and treatment planning. The novel segmentation methods are deployed to extract the anatomical structure and tumor from breast medical image. The image segmentation algorithm can be categorized into three generation the first, second and third. The first generation algorithms are based on thresholding (Otsu), Seed point selection (Region Based

Manuscript received February 20, 2021; revised May 3, 2021.

Shwetha S. V. is with Visvesvaraya Technological University, India.

Dharmanna L. is with the Dept. of ISE, SDMIT, Ujire, India (e-mail: dharmannasdmit@gmail.com).

Basavaraj S. Anami is with Institute of Technology Hubballi, India.

Mohamed Rafi is with the Department of CSE at UBDTCE, India.

DOI: 10.7763/IJCTE.2021.V13.1298

Segmentation), edge based tracing methods (Prewitt, Canny, Sobel, LoG, Roberts). The second generation algorithms incorporate uncertainty and optimization (Clustering algorithms, water shed transformation, Markov random filled). The third generation algorithm considers the prior information in segmentation(Artificial Neural Network, graph cut approach and Atlas algorithm) . The proposed work describes the pros and cons of the algorithms for computer based diagnosis system. The current work also focused on better understanding of various segmentation algorithms and its characteristics for breast medical image.

II. Literature Review

This section reveal about several authors work on segmentation methods for the natural and medical images for the extraction of Region of Interest (RoI). The segmentation algorithms are classified into three generations. In this section author broadly categorized into mainly two types, edge based segmentation [7] and region based segmentations. The section 2.1 describes edge based segmentation adopted for the various object’s edge segmentation work and section 2.2 describes the work of the region based segmentation.

A. Edge Based Segmentation

Desise Gulito *et al.* explore segmentation of malignant tumor from mammogram using fuzzy region growing method instead of crisp sets method, this method begins with a particular seed point and adopted a fuzzy member function that is statistical measurement of region being grown. The author had been adopted this method to segment several mammogram images and experimental results obtained are very closer to the tumor boundary drawn by the oncologist. The information available around the tumor is retained and computed the statistical measures [8]. The inhomogenity feature of the tumor gives the potential to classify the masses either benign or malignant. However, the author didn’t not attempt on accuracy of segmentation and classifications of the tumor.

Kamal kannan J *et al.* presented an identification of abnormalities in breast digital mammogram in order to detect cancer tumor in breast [9]. In this work the author performed filtering and segmentation of tumor in the mammograms. The enhanced mammogram by Guassian and Laplacian filters and applied with Otsu segmentation Technique for the extraction of tumor from the image. The author also presented double assessment of the malignant tumor for the better diagnosis system; hence oncologist takes the help of CAD (Computer Aided Diagnostic) system. The author did not develop a fully automated diagnosis system. The author also didn’t provide statistical feature based classification.

The authors Sham Levis and Aijuan Donge [9] reveal the study on water shed transformation for the segmentation of tumor in the mammogram from Mammographic Image Analysis Society (MIAS) [10] database. In this work the authors initially selected foreground and background markers, then adopted proposed algorithm to isolate the tumor region from its surrounding tissue. The approach was based on pixel density variations that were available in all mass tumors.

Dr. D Manimegalai *et al.* investigated a study on mammography image, in which the tumor region was separated by morphological Top Hat filter image transformation method [11] for the segmentation of the tumor. The author also intentionally performed pre-processing technique in order to highlight the contrast between tumor area and other portion of the mammogram. Then author had been extracted the first order features like Grey Level Cooccurence Matrix (GLCM) features, discrete wavelet transform, run length and high order gradient extracted from the tumor. The database was created for this feature set that is fed to the classifier algorithm such as Support Vector Machine (SVM) [12] for automatic classification. In this study the author had involved 322 images [13] with ground truth results, but did not presented on shape based features and clustering. Another disadvantage of this work is the high false positive rate.

B. Area Based Segmentation

Xiuo-ping–zhang had presented a novel approach to segment the suspicious abnormality area using wavelet packet transforms [14]. This method generated an image phase in which required object was separated from the background and also tumor region had been segmented. The author incorporated multi scale region based segmentation technique for the extraction of boundary information of detected area accurately [15]. In this work the author did not experiment on other abnormalities like micro-calcifications in mammograms, hence extensive evaluation and also refining of the algorithm is required.

Amar nedra *et al.* presented a work on classification of the breast abnormalities in digital mammograms via linear support vector machine method. The main objective of this work is to distinguish between two kinds of patients those who have cancerous and non cancerous tumor by processing digital mammograms through linear SVM technique. The experimental work inculcated segmentation of breast tissues using k-means followed by feature extraction using SURF [16] and finally classified through SVM classifier.

Arnau and Oliver *et al*. proposed a method called one shot segmentation of breast pectoral muscles and background in digitized mammograms. The author solved the problem by an approach of dividing the work it into two halves [17], first one is removal of background from the mammogram and second step was separation of pectoral muscles. For the experiment the images were used from MIAS database. But the author did not elaborated on identification and classification of cancerous tumor.

Nana Ramadijanti, *et al.* proposed technique using hierarchical k- means on mammograms. In this work author used valid tracing to get optimal number of clusters via hierarchical k-means clustering to obtain different clusters and each cluster is labeled as components. The author work demonstrated an error of 61.1% [18] by performing system testing on 36 data sets. In this work the author identified the tumor accurately based on the shapes of the tumor likely circle and oval with a well defined margin.

Damian Valdés-Santiago, *et al.* presented a work on the images with low contrast and diversity in breast anatomy that lead to unclear border of the suspicious anomalies for the visualization [19]. In order to solve the problem the mammographic mass segmentation was done using fuzzy c-means and classification by decision tree technique. Fuzzy c-means segmentation was enhanced by using the image histograms. The features were selected of the region of interest using GLCM. And the author had claimed 90% sensitivity and 70% specificity.

Leson *et al.* worked on image segmentation frame work for extracting tumors from breast MRI (Magnetic Resonance Imaging). The author had solved manual delineation of ROI and huge labeling of each component. In order to solve the problem the author has adopted a super voxel strategy for segmentation (a semi supervised). A supervised learning step was adopted for the location of tumor patches. Still author has to perform on fully automated system.

III. Segmentation Methods and Materials

Several segmentation methods adopted for the current research work are described through the Fig. 2. Data flow diagram of segmentation methods. The data flow diagram, has four phases. In the first phase, database is designed for breast and natural images by collecting images from local hospitals and well known databases MIAS, DDSM and Berkeley. In the phase two the images which are directly captured from medical instruments and collected from the database. These are noisy and PSNR value on an average 10 db. These images are enhanced using modified Gabor filters. In phase three segmentation algorithms adopted for the work has been described, details of each segmentation algorithms has been described in the following sections. In the fourth phase the result of all segmentation algorithms has been stored on the database and metrics of each algorithm also computed.

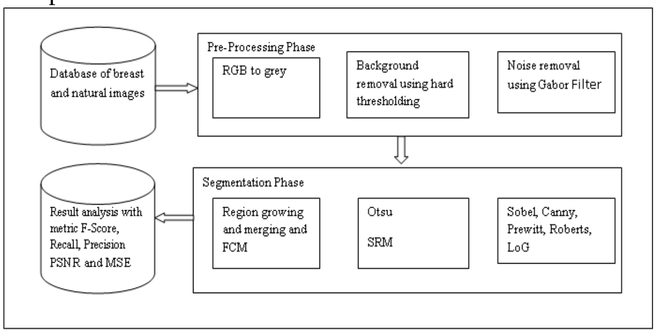

Fig. 2. Dataflow diagram of segmentation methods.

A. Region Based Segmentation

The region based segmentation like SRM, FCM, Otsu have been discussed in following section.

1) Fuzzy c-means (FCM)

The fuzzy set is defined as every pixel of the element of the mammogram images.FCM methods well defined. Various authors, fuzzy membership function and fuzzy region. The algorithm is implemented for the dimension of data, clusters, membership function and termination criteria. The algorihm for membership function as follows,

Algo. MembershipFunc()
If $|P-\boldsymbol{\mu}|<=\Delta_{max}$ then $\sigma=1$
Elseif $|P-\mu\,| >\Delta_{max}$ then $\sigma=0$
Else $\sigma = 1/(1+\Gamma(|P-\boldsymbol{\mu}|))$
end

2) Region growing algorithm

The algorithm is carried out through closer intensity of pixel in the given image that is to be extracted. The region that are being extracted by applying the following properties of the algorithm,

1. The summation of all sub regions area should be equal to whole image
2. Every region must be continuous and connected
3. The pixel should belong to only one region of the image
4. Each region should satisfy the homogeneity property
5. Two neighboring regions should not have common pixel

Region Growing and Merging based algorithm have four methods such as region growing, splitting, merging and Region splitting and merging.

Each of the above procedure are implemented with the following steps,

a) Algo. region growing ()

Step 1: Similar intensity pixels are grouped together for the extraction of region on some predefined conditions
Step 2: The seed point (x1, y1) is selected initially and region will start growing from this point
Step 3: Consider the next pixel randomly from the image that is to be segmented and nearest neighbors are observed depending on the type of connectivity (4-connectivity, 8-connectivity or m-connectivity)
Step 4: The neighboring pixel was selected from the region as (x1, y1) satisfying the homogenic property of a region
Step 5: A new pixels called (x2, y2) is considered as a member of the region. Then the surrounding pixels of the selected pixel (x2, y2) are examined and enhanced the area of the region
Step 6: This method continues in the same way until no more new pixel is accepted for the current region; then labeled all the selected pixels that belongs to current region

b) Algo. region splitting ()

Step 1: In region splitting consider homogeneity property pixels that are grouped together
Step 2: If the homogeneity property is not satisfied then split the region into four equal sub regions
Step 3: Repeat step 1 and step 2 until all the regions of the image satisfies the given property
Step 4: The dividing strategy is as shown below
Step 5: The whole object will be represented by R, which is called parent node and then it is divide d into four leaf nodes R_1, R_2, R_3 and R_4. In these leaf nodes, node R_4 doesn't contain pixels which satisfy homogeneity condition, then R_4 is subdivided into four areas such as R_{41}, R_{42}, R_{43}, R_{44}
Step 6: All pixels in a particular region satisfies homogeneity property in the indivisible region

c) Algo. region merging ()

The algorithm Region merging is the other hand of the region splitting technique and is described as following steps,
Step 1: in this algorithm start from the pixel level and accept every region as homogenous region at any level of merging to examine its core adjacent homogeneous

area arranged in 2X2 manner, together must satisfy the homogeneous property. Step 3, if the property is satisfied all the pixels are merged to create a bigger region otherwise the region has left as they are.

d) Algo. split &merge ()

Step 1: whole image is split into four quadrants and continue splitting every quadrant further until all the subsection or region satisfies the homogeneity property.
Step 2: In a region merging each pixel of a smaller region merge regions into larger regions, if smaller regions satisfy the homogeneous property.
Step 3: if the homogenous region are small, region merging algorithm is better otherwise region splitting is better.
Step 4: In our application a combination of region merging and splitting is adopted.

The limitation of this algorithm is initial seed point need to be selected manually and also consume more time, so in order to overcome the drawback of this the Otsu algorithm have been discussed in the next section.

3) Otsu algorithm

Otsu method was invented by Nobuyki Otsu. This method accomplishes automatic image threshold. The method estimates a single threshold (pixel intensity) that divides the image into different classes such as foreground and background. This method is suitable for noisy image segmentation.

This method is to be applied for grey level images. The given pixel grey level to be divided into L discrete values and mean grey image is also divided into L values. The pair is formed (1) the pixel grey level and (2) mean of the surrounding pixel (x, y). Each pair is associated to one of L x L possible in 2-D bins. The f_{xy}: total number frequency of a pair (x, y) separated by all the pixels in the image N. That is defined by joint probability mass function in 2-D histogram.

$$p(x,y) = \frac{f(x,y)}{N} \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} p(x,y) \tag{1}$$

2D Otsu technique

The probability of two classes that can be indicated as,

$$W_0 = \sum_{x=0}^{s-1} \sum_{y=0}^{t-1} p(x,y) \tag{2}$$

$$W_1 = \sum_{x=s}^{L-1} \sum_{y=t}^{L-1} p(x,y) \tag{3}$$

The pixel intensity average value vectors of two classes and total vector obtained as,

$$\mu_0 = [\mu_{0x}, \mu_{0y}]^T = \left[\sum_{x=0}^{s-1} \sum_{y=0}^{t-1} i \sum_{W_0}^{p(x,y)} \sum_{x=0}^{s-1} \sum_{y=0}^{t-1} p(x,y)/W_0 \right]^T$$

$$\mu_1 = [\mu_{1x}, \mu_{1y}]^T = \left[\sum_{x=s}^{L-1} \sum_{y=t}^{L-1} xp(x,y)/W_1 \sum_{x=t}^{L-1} \sum_{y=t}^{L-1} yp(x,y)/W_1 \right]^T$$

$$\mu_T = [\mu_{Tx}, \mu_{Ty}]^T = \left[\sum_{x=0}^{L-1} \sum_{y=0}^{L-1} xp(x,y)/W_1 \sum_{x=0}^{L-1} \sum_{y=0}^{L-1} yp(x,y) \right]^T$$

In all the cases diagonal probability is must zero, hence it is easy to check

$$W_0 + W_1 \cong 1$$

$$W_0\mu_0 + W_1\mu_1 = \mu_T$$

The interclass matrix (discrete) is defined as

$$S_b = \sum_{k=0}^{1} w_k \left[w_k (\mu_k - \mu_T)(\mu_k - \mu_T)^T \right]$$

Algo.Otsu()
// input: grey scale image
//output: tumor region as foreground and other part is background
Max, s, t←0
for t← 0 to L-1 do
 for s← 0 to L-1 do
 Estimate tr(S_b)
 if tr(S_b)> Max
 Max←tr(S_b)
 S←s
 T←t
 return S and T
 end if
 end for
end for
close

4) Statistical region merging (SRM)

This algorithm works for the colored image samples and segment all the objects that are having different colored tissue. All the regions are grouped depending upon the merging rule, resulting in a list of several smaller objects, grouping a number of surrounding pixel depend upon their shades that belong within a specific threshold.

For example, with 10 values of pixel intensity of array x={1.7,1.8,1.9,3.2,4.9,5.1,5.2,5.6,9,10} with range of x<10. For the above pixel values, if the grouping criteria are only a threshold that defines the distance of the selected pixel intensity value should be within 0.3 ranges and mean should be applied. The detailed calculation as follows,

(1.7+1.8+1.9)/3=5.4/3=1.8
3.2=3.2/1=3.2
4.9=4.9/1=4.9
(5.1+5.2+5.3)/3=15.6/3=5.2
5.3=5.3/1=5.3
9=9/1=9
10=10/1=10

The obtained results will be 1.8, 3.2, 4.9, 5.2, and 5.3,9,10.

So in the above example the ten region of the sample merges into seven regions .

Algo. SRM()
Input: colored breast tissue sample
Output: segmented tissue sample
for k← 0 to q
 for i← 0 to m
 for j← n
 if(img(i, j)) <= threshold
 sum←sum + img (i, j)
 end if
 end for j
 end for i
 newsample←sum
end for k

B. Edge Based Segmentation

The section describes various edge based segmentation algorithms with the description of each in the following section.

1) Sobel & canny

The Sobel & Canny edge detection are based on gradient image. The first step is the convolution of image and a gradient kernel on X and Y direction after the threshold of the image. It is result of Sobel and Canny detector is non maximum suppression and hysteresis threshold and also used for tracing along the edges of the object.

2) Prewitt

The Prewitt method is used for image segmentation particularly for identifying edges of the object. In this method an approximation of the gradient for an image intensity function is computed. At every point of the image the Prewitt result is either the corresponding gradient vector or the norm of this vector. The method mainly depends on convolving the image with a tiny, separable an integer value filter in vertical and horizontal direction. This method produces high frequency variations in the object.

3) Roberts

The Robert operator is an edge based segmentation for detecting edges in the object. The idea behind this segmentation is that approximate the gradient of an image to discrete differentiation which is achieved by estimating the sum of the square of the difference between diagonally adjacent pixels.

4) LoG

The LoG of an image segmentation method focused on area of rapid intensity changes, hence the method can identify all the edges of the object. In this method Guassian, a special filter is applied to minimize its sensitivity to noise then LoG is deployed for the segmentation of edges. The grey level image is the input and resultant image is another edge based object. The kernel used for this method is as shown in Fig. 3,

0	-1	0
-1	4	-1
0	-1	0

-1	-1	-1
-1	8	-1
-1	-1	-1

Fig. 3. The two commonly used discrete approximations to the Laplacian Filter.

IV. Metrics to Measure Quality and Accuracy of Segmentation

The experimental work is carried out for various kinds of images. We have used the Berkeley Segmentation Dataset (BSDS500), DDSM (Digital Database for Screening Mammography) and Ultrasound breast image database. Segmented image accuracy was estimated in terms of Precision, Recall, F-score, MSE and PSNR with existing benchmark results. The gold standard segmentation is given by the experts and with respect to that the segmentation is compared. The edge detectors will provide a similar boundary map where every pixel value by considering the possibility of benign edges.

To calculate the metrics like Precision, Recall and F-score, the algorithm is given below,

Algo. Accuracy (Recall ,Precision,Fscore)
tp,fp,fn,tn←0
set m and n to the upper limit
Repeat i← 0 to n do
Repeat j← 0 to m dosed image is 1
Compare image element of the gold standard is 1 and proposed image is 1
Increment value of tp
Otherwise Compare image element of the gold standard is 1 and proposed image is 0
Increment the value of fp
Else otherwise Compare image element of the gold standard is 0 and proposed image is 0
Increment the value of tn
Else increment the value of fn
End of the loops
Estimate the Precision, Recall and FScore with well defined equations

The above algorithm estimated precision and recall values of the test images. The precision gives false positive for a test image with respect to benchmark whereas recall gives true positive rate. Apart from precision and recall, the F-score also computed which provides the harmonic mean of estimated values. The Table I describes the accuracy result computed for the various segmentation methods and in that RGM, FCM, Otsu and SRM shows better result. The value obtained from the result are shown in Fig. 4 to Fig. 12

Apart from that the MSE and PSNR of the segmentation is calculated for the mammograms to measure the quality and the algorithm is given below,

Algo. MSE & PSNR()
Input: read original image
Output: MSE and PSNR of the image
Enhance the original image with modified Gabor filter and label as input 2
Adjust dimension of input image1 and labeled as input image 2
sq←(double(input1)-double(input2))^2
MSE←∑∑sq/dimension of matrix(mXn)
PSNR←10*log $_{10}$ (255^2/MSE)
Display MSE and PSNR as output

V. Result

We have chosen several test image sets for the experiment from BSDS-500 datasets, DDSM and Ultrasound. The enhancement of the images is done by modified Gabor Filter and later segmented. The Precision, Recall, F-Score, MSE and PSNR were computed and tabulated for the segmentation techniques like Canny, Sobel, Roberts, LoG, Prewitt, Otsu, FCM, SRM and Region Growing &Merging against the gold standard database. Region growing and merging and human boundary marked for the benign and malignant tumor. In this experiment we have involved natural and medical images for the segmentation. We used one set of fixed parameter for all natural and medical images. The result shows the F value and speed of all nine algorithms. The algorithms like RGM, FCM, Otsu and SRM algorithms were modified for tumor segmentation accuracy. The FCM, SRM and Region growing and merging showed the best performance in the measure of F-Score. The standard algorithms like Canny, Sobel and Prewitt etc, are fast but poor in accuracy. Among learning based algorithms, FCM and SRM are the fastest one. The experiment was conducted in the laptop with Intel Core-2 dual Compaq 510 and the software MATLAB version 2018. The RGM segmentation is applied on benign and malignant mammogram and ultrasound images.

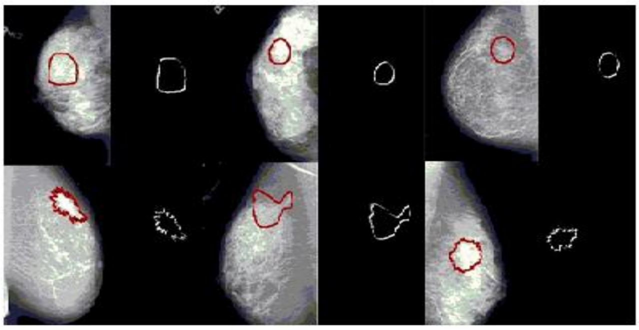

Fig. 4. The first row: benign tumor segmentation and second row: malignant tumor segmentation by region growing and merging.

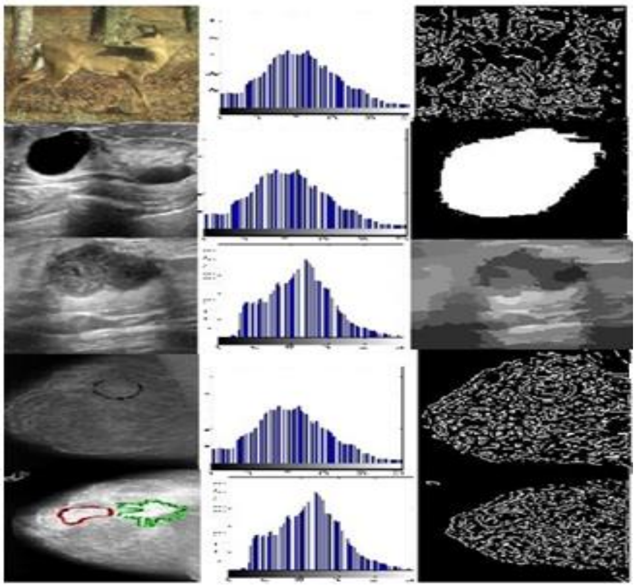

Fig. 5. (a)Input Image (b) Histogram (c)Segmented Image for SRM segmentation.

The first row of the Fig. 4 shows the benign samples and second row the malignant samples of three images out of ten segmented images. The Fscore, Recall, Precision of the segmented image is calculated and tabulated in Table II. The Fig. 5 to Fig. 12 showcases the segmentation of several types of images by FCM, Otsu, Canny, Sobel, Prewitt, LoG and Roberts respectively. The first column of the figures are the input images (natural image, ultrasound breast images with benign and malignant tumor, mammogram images with benign and malignant tumor) and the second column shows the histogram of the enhanced image and the third column shows the RoI (Region of Interest) extracted.

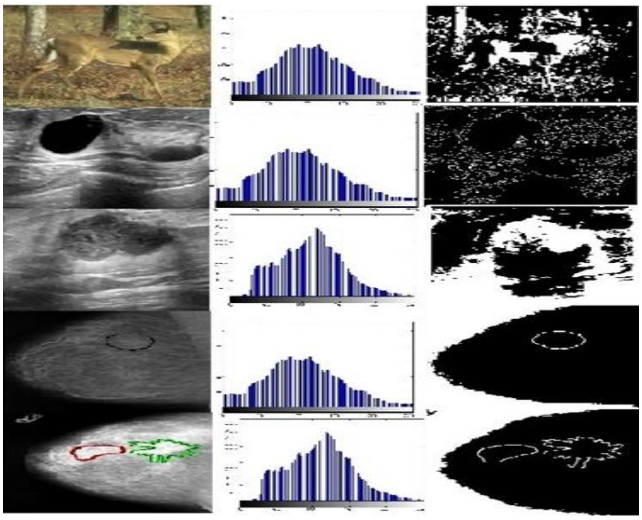

Fig. 6. (a)Input Image (b) Histogram (c)Segmented Image for FCM Segmentation.

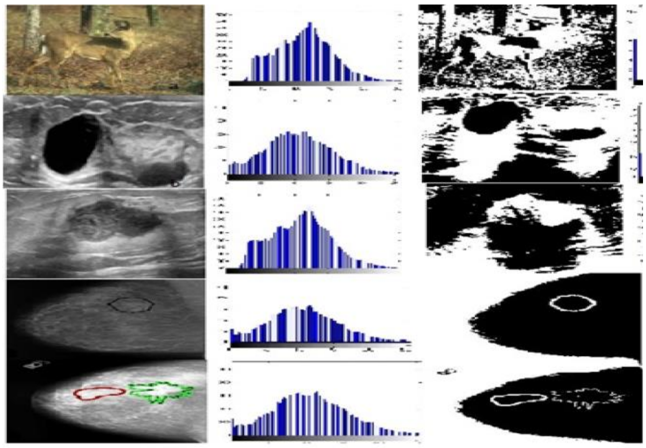

Fig. 7. (a) Input Image (b) Histogram (c) Segmented Image for Otsu Segmentation.

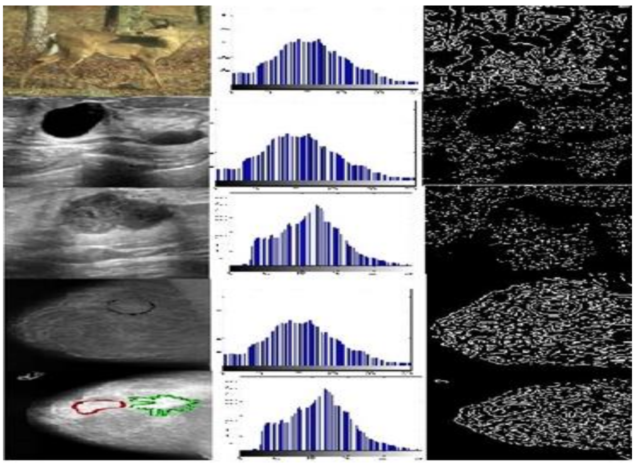

Fig. 8. (a) Input Image (b) Histogram (c) Segmented Image for Canny Segmentation.

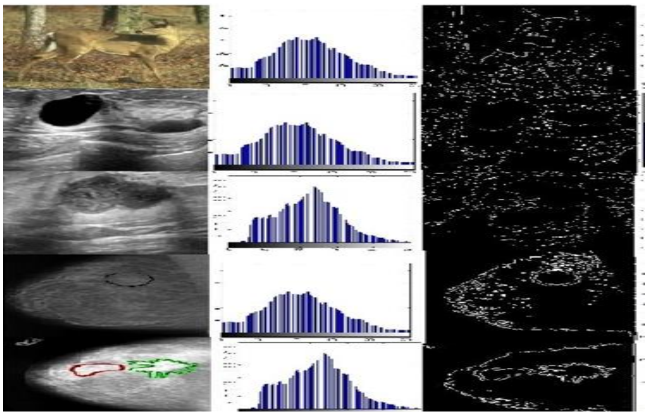

Fig. 9. (a) Input Image (b) Histogram (c) Segmented Image for Sobel Segmentation.

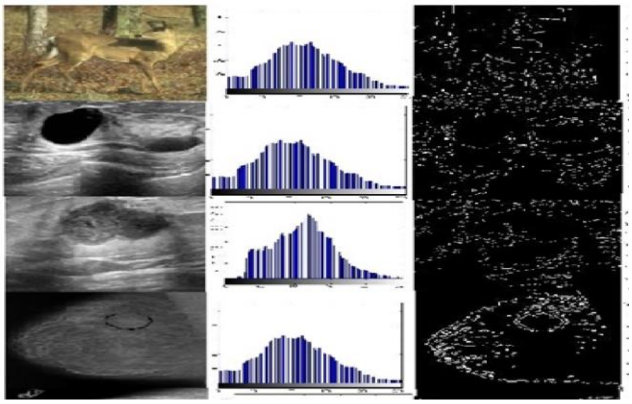

Fig. 10. (a) Input Image (b) Histogram (c) Segmented Image for Prewitt Segmentation.

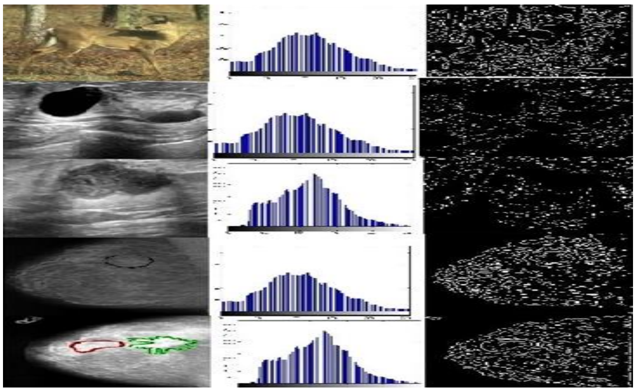

Fig. 11. (a) Input Image (b) Histogram (c)Segmented Image for LoG Segmentation.

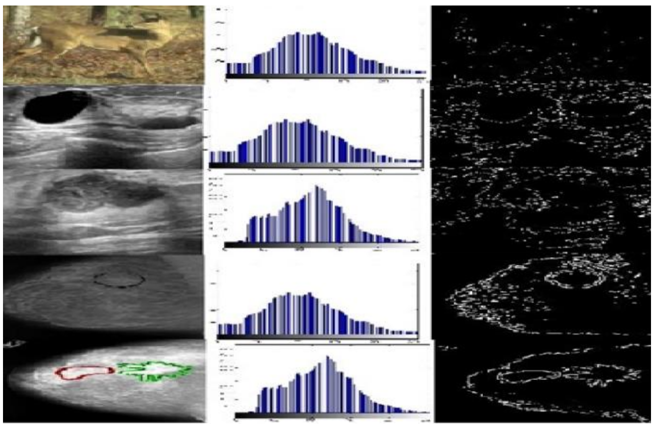

Fig. 12. (a) Input Image (b) Histogram (c) Segmented Image for Roberts Segmentation.

The Fig. 4 to Fig. 12 describes the segmentation results of different methods adapted in this current research work result of computed metrics such as F-Score, Recall, Precision, PSNR and MSE are shown in the Table I. Region growing &merging and FCM are to be considered for the segmentation of tumor in the mammogram images. Statistical Region Merging algorithm performs better for the segmentation of breast tissue samples to identify the stages of the cancer. For RGM, the individual values of segmentation are observed to be very close to the average value obtained in Table II for RGM. Similarly the other segmentation values are observed. The edge based segmentation like Canny, Sobel, Prewitt, LoG and Roberts are not applicable for the tumor segmentation in mammograms and ultrasound because edge based segmentation shows all the boundary lines of blood vessels, tumor edges etc which leads to lot of ambiguity in detection of tumor region exactly. Fig. 8 shows the segmented image having FScore value = 0.006 having accuracy of 0.06%. Similarly other edge based segmentations also have low accuracy compared to RGM, FCM and Otsu segmentation techniques.

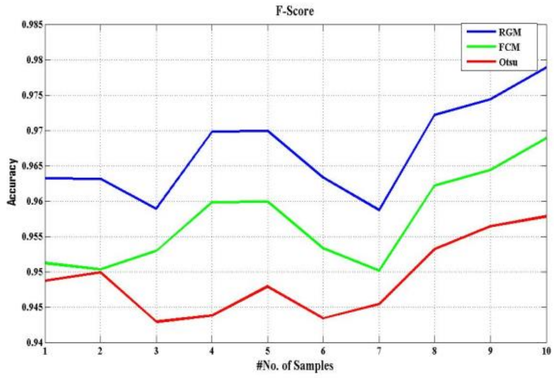

Fig. 13. F-score of region growing & merging, FC means and Otsu.

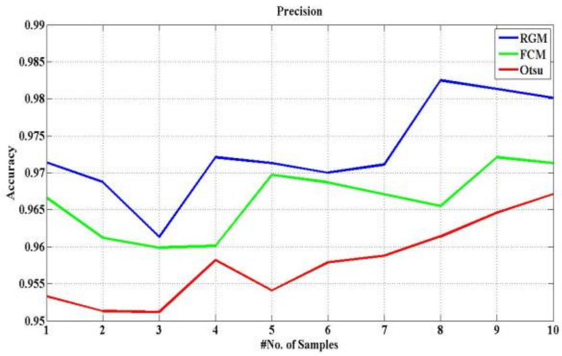

Fig. 14. Graph of precision for region growing & merging, FC means and Otsu Segmentation.

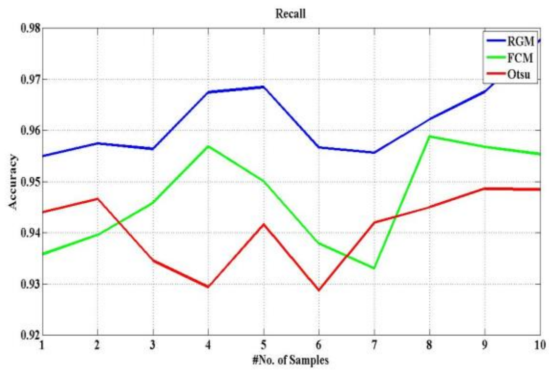

Fig. 15. Graph of Recall for Region Growing & Merging, FC Means and Otsu Segmentation.

The Fig. 13 to Fig. 15 depicts the accuracy metrics like Fscore, Recall and Precision of the three major segmentation algorithms (RGM, FCM, Otsu) against the number of samples (10 samples) as provided in Table II. The Otsu segmentation is based on the light intensity, and this method is not suitable for the image which contains foreground and back ground objects. But this method is having high speed for extracting the RoI from the mammogram. Region growing and merging segmentation algorithm is optimal for the homogeneity regions. But still in-homogeneity criteria suffer to handle appropriately and also this algorithm takes longer time over the Otsu, FCM and RGM. Hence RGM has an average of 96.72, FCM algorithm segments the region of interest different threshold and is difficult to judge which clustered part. The total run time of the experiment by Otsu is 3 seconds, FCM is 7 seconds and RGM is 14 seconds.

TABLE I: ESTIMATION OF VARIOUS SEGMENTATION METHODS WITH RESPECT TO F-SCORE, PRECISION, RECALL, MSE, PSNR

Sl.No	Segmentation Technique	Metrics	Benign Ultra sound image	Malignant Ultra sound image	Benign mammogram	Malignant mammogram	Deer image
1.	Region Growing & Merging	FScore	0.97	0.96	0.996	0.9925	0.9734
		Precision	0.98	0.96	0.996	0.9925	0.9734
		Recall	0.86	0.96	0.996	0.9925	0.9734
		MSE	78	120.85	38.89	127.9	121.6
		PSNR	40	39	35	38	36
2.	SRM	FScore	0.9863	0.0175	0.7865	0.7726	0.9335
		Precision	0.9863	0.0175	0.7865	0.7726	0.9335
		Recall	0.9863	0.0175	0.7865	0.7726	0.9335
		MSE	382.98	513.12	1822.95	176.29	3639.24
		PSNR	22.33	21.06	15.56	25.70	12.55
3.	FCM	FScore	0.93005	0.9795	0.95802	0.8891	0.9299
		Precision	0.93005	0.9795	0.95802	0.8891	0.9299
		Recall	0.93005	0.1795	0.95802	0.8891	0.9299
		MSE	936.71	1233.46	1096.66	205.38	9639.90
		PSNR	18.45	17.25	17.76	25.04	8.32
4.	Otsu	FScore	0.9329	0.9215	0.9412	0.9511	0.9288
		Precision	0.9329	0.9215	0.9412	0.9511	0.9288
		Recall	0.9329	0.9215	0.9412	0.9511	0.9288
		MSE	0.9373	1226.49	1121.47	206.49	13227.61
		PSNR	17.70	17.28	17.67	25.02	6.95
5.	Canny	FScore	0.0061	0.0264	0.9731	0.9976	0.9205
		Precision	0.0061	0.0264	0.9731	0.9976	0.9205
		Recall	0.0061	0.0264	0.9731	0.9976	0.9205
		MSE	2253.54	5107.12	750.12	1006.25	19715.47
		PSNR	14.64	11.08	19.41	18.14	5.22
6.	Sobel	FScore	0.0033	0.0249	0.97742	0.9969	0.9183
		Precision	0.0033	0.0249	0.97742	0.9969	0.9183
		Recall	0.0033	0.0249	0.97742	0.9969	0.9183
		MSE	845.04	1206.62	37.23	245.15	1206.62
		PSNR	18.90	17.35	32.46	24.27	17.35
7.	Prewitt	FScore	0.0027	0.0251	0.9745	0.9968	0.9184
		Precision	0.0027	0.0251	0.9745	0.9968	0.9184
		Recall	0.0027	0.0251	0.9745	0.9968	0.9184
		MSE	844.79	1206.99	37.91	245.13	24299.04
		PSNR	18.90	17.35	32.38	24.27	4.31
8.	LoG	FScore	0.0033	0.0253	0.9753	0.9985	0.9197
		Precision	0.0033	0.0253	0.9753	0.9985	0.9197
		Recall	0.0033	0.0253	0.9753	0.9985	0.9197
		MSE	845.70	1201.87	58.31	238.79	20871.98
		PSNR	18.89	17.37	30.51	24.38	4.97
9.	Roberts	FScore	0.0103	0.0293	0.9719	0.9950	0.9176
		Precision	0.0103	0.0293	0.9719	0.9950	0.9176
		Recall	0.0103	0.0293	0.9719	0.9950	0.9176
		MSE	848.78	1203.85	42.74	246.19	25663.24
		PSNR	18.88	17.36	31.86	24.25	4.07

TABLE II: F-SCORE, RECALL, PRECISION OF TEN SAMPLES

Image Sample	FScore			Recall			Precision		
	RGM	FCM	Otsu	RGM	FCM	Otsu	RGM	FCM	Otsu
1.	0.9612	0.9512	0.9487	0.9550	0.9358	0.9441	0.9714	0.9666	0.9533
2.	0.9603	0.9503	0.9499	0.9575	0.9396	0.9467	0.9687	0.9612	0.9513
3.	0.9589	0.9529	0.9429	0.9564	0.9459	0.9346	0.9613	0.9599	0.9512
4.	0.9698	0.9598	0.9438	0.9675	0.9569	0.9294	0.9721	0.9601	0.9582
5.	0.9699	0.9599	0.9479	0.9685	0.9501	0.9417	0.9713	0.9697	0.9541
6.	0.9633	0.9533	0.9434	0.9567	0.9379	0.9288	0.9700	0.9687	0.9579
7.	0.9587	0.9501	0.9454	0.9557	0.9331	0.942	0.9711	0.9671	0.9588
8.	0.9722	0.9622	0.9532	0.9622	0.9589	0.945	0.9285	0.9655	0.9614
9.	0.9744	0.9644	0.9564	0.9676	0.9568	0.9486	0.9813	0.9721	0.9646
10.	0.9789	0.9689	0.9578	0.9777	0.9554	0.9485	0.9801	0.9713	0.9671
Avg	**0.9672**	**0.9573**	**0.9489**	**0.9625**	**0.9470**	**0.9409**	**0.9730**	**0.9662**	**0.9578**

VI. Conclusion

The segmentation algorithm plays an important role for the diagnosis of breast cancer through ultra sound and mammogram images. Still there exist plenty of issues like accurate extraction of tumor region and its localization. The work includes the implementation of various number of segmentation. The diagnosis depends on the better segmentation that leads to the extraction of better features for the classification and diagnosis of breast cancer disease. Through this experiment the algorithms RGM, FCM and Otsu are the heart of the segmentation. In this work the strength and weakness of algorithms is analyzed through the parameters F-score, Recall, Precision. The image quality is measured by the metrics like MSE and PSNR. In our further work these algorithms are to be deployed and features need to be extracted for the detection of disease and for the appropriate classification using machine learning algorithms.

Conflict of Interest

The authors declare no conflict of interest.

Author Contributions

Ms. Shwetha S V had carried out the work of segmenting the tumor in the mammograms and ultrasound images of benign and malignant samples for the extraction of the tumor. Dr. Basavaraj S. Anami and Dr. Mohamed Rafi advised for the segmentation of the tumor by various segmentation algorithms and to compare with the other segmentation algorithms for the determination of best segmentation algorithms. Dr. Dharmanna L, the research supervisor had suggested to work on the Berkeley databases and to apply the suitable algorithms like edge based and area based algorithms for the extraction of the boundary of the tumor from mammogram. He suggested comparing the segmented tumor and analyzing with respect to PSNR, MSE, Recall, Precision and F-Score of the segmented image.

Acknowledgment

We thank hospitals at Davanagere and Bhadravathi for the samples. We also we thank to Visveswaraya Technological University for the support.

References

[1] G. M. Jayesh and S. P. Sankar, "Efficient preprocessing filters and mass segmentation techniques for mammogram images," in *Proc. 2017 IEEE International Conference on Circuits and Systems (ICCS)*, 2017.

[2] A. Q. Syed and K. Narayanan, "Detection of tumor in MRI images using artificial neural networks," *International Journal of Advanced Research in Electrical, Electronics and Instrumentation Engineering*, vol. 3, no. 9, pp. 11749-11754, 2014.

[3] L. Zhen and A. K. Chan, "An artificial intelligent algorithm for tumor detection in screening mammogram," *IEEE Transactions on Medical Imaging*, vol. 20, no. 7, pp. 559-567, 2001.

[4] R. Valliappan, P. Then, and P. Sumari, "Digital mammogram tumor preprocessing segmentation feature extraction and classification," in *Proc. 2009 WRI World Congress on Computer Science and Information Engineering*, 2009.

[5] H. T. Nguyen *et al*., "Detection of masses in digitised mammograms using dendronic image analysis," in *Proc. the Second Joint 24th Annual Conference and the Annual Fall Meeting of the Biomedical Engineering Society Engineering in Medicine and Biology*, 2002.

[6] A. Nuh *et al*., "Classification of breast masses in mammogram images using KNN," in *Proc. 2015 23nd Signal Processing and Communications Applications Conference (SIU)*, 2015.

[7] S. Neeraj and L. M. Aggarwal, "Automated medical image segmentation techniques," *Journal of Medical Physics/Association of Medical Physicists of India*, vol. 35, no. 1, p. 3, 2010.

[8] R. D. Ghongade and D. G. Wakde, "Computer-aided diagnosis system for breast cancer using RF classifier," in *Proc. 2017 International Conference on Wireless Communications, Signal Processing and Networking (WiSPNET)*, 2017.

[9] S. H. Lewis and A. Dong, "Detection of breast tumor candidates using marker-controlled watershed segmentation and morphological analysis," in *Proc. 2012 IEEE Southwest Symposium on Image Analysis and Interpretation*, 2012

[10] G. Qiu *et al*., "Automatic segmentation of micro-calcification based on sift in mammograms," in *Proc. 2008 International Conference on BioMedical Engineering and Informatics*, 2008.

[11] S. Nadia and A. Hlima, "Designing a new approach for the segmentation of the cancerous breast mass," in *Proc. 2016 13th International Multi-conference on Systems, Signals & Devices (SSD)*, 2016.

[12] H. Abdulkader and R. H. Abiyev, "ISIBC: An intelligent system for identification of breast cancer," in *Proc. 2015 International Conference on Advances in Biomedical Engineering (ICABME)*, 2015.

[13] F. M. Mohamed, D. Manimegalai, and S. Thaiyalnayaki, "Automatic detection of tumor subtype in mammograms based on GLCM and DWT features using SVM," in *Proc. 2013 International Conference on Information Communication and Embedded Systems (ICICES)*, 2013.

[14] X.-P. Zhang, "Multiscale tumor detection and segmentation in mammograms," in *Proc. IEEE International Symposium on Biomedical Imaging*, 2002.

[15] R. Kaur and A. Doegar, "Localization and classification of brain tumor using machine learning & deep learning techniques," *International Journal of Innovative Technology and Exploring Engineering*, vol. 8, issue 9, July 2019.

[16] N. Amara, M. Shoaib, and S. Gattoufi, "Detection and classification of the breast abnormalities in Digital Mammograms via linear support vector machine," in *Proc. 2018 IEEE 4th Middle East Conference on Biomedical Engineering (MECBME)*, 2018.

[17] O. Arnau *et al*., "One-shot segmentation of breast, pectoral muscle, and background in digitised mammograms," in *Proc. 2014 IEEE International Conference on Image Processing (ICIP)*, 2014.

[18] R. Nana, A. Barakbah, and F. A. Husna, "Automatic breast tumor segmentation using hierarchical k-means on mammogram," in *Proc. 2018 International Electronics Symposium on Knowledge Creation and Intelligent Computing (IES-KCIC)*, 2018.

[19] D. Valdés-Santiago, D. Mesejo-León, and Á. León-Mecías, "Multiple-level logarithmic wavelets for mammographic contrast enhancement: A statistical analysis for wavelet selection," *Computation Systems*, vol. 22, issue 2, pp. 621-637, 2018.

Shwetha S. V. is a research scholar, who is pursuing the Ph.D. in the stream of image processing under Visvesvaraya Technological University, Belagavi. She had completed her M.Tech. and B.E. in the year 2013 and 2008 respectively. She had 12 years of teaching experience at SDMIT Ujire. She had published her three research in Scopus indexed international journal and one in international conference. Her area of interest is medical image processing.

Dharmanna L. is an associate professor and the head of Department of ISE, SDMIT, Ujire. He had received his Ph.D. degree in the stream of medical image processing from Visvesvaraya Technological University, Belagavi in 2016. He had completed his M.Tech. and B.E. from VTU in the year 2006 and 2004 respectively. He has 15 years of teaching experience. He had published 30 research articles in several Scopus indexed international journals.

Basavaraj Anami is a well-known academician and researcher from India in the field of computer science and engineering. He is presently, the principal of KLE Institute of Technology Hubballi, an esteemed educational institution of Karnataka, India. He possesses a vast teaching experience of 32 years in computer science discipline. His research interests include image processing, speech processing and artificial intelligence. He has to his credit about 65 research papers published in various journals and conferences in India and abroad.

Mohamed Rafi is a professor in the Department of CSE at UBDTCE, Davanagere. His area of interest is image processing, data analytics, computer vision, database management, cyber security. He has record of 103 research publications in international and national journals and conferences. He has served as the vice principal at HSMIT, Tumkur; the chairman of department and several other committees.