[1] P.Narasimhaiah

[2] G. Sreenivasula Reddy

[3] C. Nagaraju

Advancements in Breast Cancer Detection: Harnessing Artificial Neural Networks for Improved Accuracy

Journal of Electrical Systems

Abstract: - Female breast malignancy is the exceedingly prevalent reason for the demise of women around the world. Women who are revealed to have breast cancer earlier in life get a lower death rate from the disease and increase the life expectancy of patients. Mammography screening is one of the effortless, efficient, and affordable ways to identify breast cancer in advance. The early investigators pioneered many methods based on statistical measurements and textural traits for the earliest identification of carcinoma of the breast. Due to artefacts, noise, pectoral muscles, and irregular illumination, the accuracy of cancer prediction in these works is relatively low. The accuracy of predictions made by employing textural characteristics for forecasting breast cancer in earlier work is 83.33%. The research proposal processes of mammograms to remove noise, artefacts, pectoralis, and inconsistent illumination in an endeavor to increase forecast accuracy. The proposed research uses an Artificial Neural Network (ANN) to classify breast masses as benign or malignant based on geometric pattern features. Its prediction accuracy is 86.67%, which is superior to research studies based on textural and statistical characteristics of breast mammograms.

Keywords: Benign, Cancer, Classification, Malignant, Mammograms, Prediction, Screening.

I. INTRODUCTION

Female breast malignancy is the most prevalent reason for demise for women around the world. The death rate in women due to tumors in the breast is 2.5 (that is 1 in 39 women will die). The mortality rate can be further decreased by the premature finding of tumors in the breast. One of the simple, efficient, and cost-effective techniques to detect breast cancer early is mammography screening. Previous works have shown that premature finding of tumors in the breast through periodic mammography screening of asymptomatic women can reduce mortality. Now mammography is not only the most sensitive and specific method for finding cancer in the breast but also the most pragmatic method for clocking [1]. The mammographic clocking programs' accuracy and efficiency should be improved for the premature finding of the sign of a tumor in the breast, several previous investigations have concentrated on techniques developed for computer–aided diagnosis to assist radiotherapists in identifying breast tumors, works including computational intelligence [2-4] and image analysis [5,6].

The several abnormalities visible in mammograms are an increase in calcification, tissue density, and masses. The masses are inspected for position and shape. Usually, malignant masses are high-density lesions, whereas benign masses are low-density, lucent-centered lesions. Smoothed, circumscribed and regular shape masses are benign, whereas rough, speculated, irregular shapes, and surrounded by radiating patterns of linear spicules masses are malignant.

Fractal analysis is used to characterize the shape and assess the texture of images [7-10]. The concept of fractal analysis is applied to mammograms to compare the shape of benign and malignant masses. Quantify the complexity or irregularity of an object boundary using fractal dimension [21]. Fractal objects have complex structures and irregular shapes and can't be represented by Euclidean dimensions. The shape of naturally occurring objects can be represented quantitatively with the help of fractal geometry. The breast tumor complex boundary is quantitatively characterized by applying the fractal concept [11].

Mammograms are interpreted and analyzed by developing Computer-Aided Detection (CAD) systems to accompany radiotherapists in identifying breast tumors. Present techniques use demographic characteristics or appearance characteristics. Demographic characteristics-based techniques utilize histograms or grey values of mammograms to categorize lumps [13]. Classification of masses based on the histogram technique is not used

[1]*Corresponding author: Research Scholar: Department of CSE, YSR Engineering College of YVU, Proddatur, India narasimhareddypolu@gmail.com

[2] Professor, Department of CSE, Chaitanya Bharathi Institute of Technilogy Proddatur, India, sreenu.gurrampati@gmail.com

[3]Professor: Department of CSE,YSR Engineering College of YVU Proddatur, India cnrcse@yahoo.com

fully because it differs in having noise or improvement [12]. Noise can be reduced to some level using some of the demographic quantities. The accuracy of methods based on demographic measures is up to 70% [14-16]. To categorize masses as either diseased or un-diseased, various approaches based on the shape characteristics of lesions were developed [20]. The methods based on texture features categorizes the breast masses as either diseased or un-diseased, consider only circumscribed, and contemplated masses as abnormal categories, but do not consider lobular, structural distortion and mammographic images with intricate masses. Researchers have felt that both shape and smoothness characteristics are equally significant and a combination of these features is used to categorize the breast masses [17]. Shape features of the mass have been used to classify breast masses, but it is not clear which type of lump is considered and the prediction precision is low [18].

The accuracy of the breast masses classification method based on statistical and textural measures is 83.33% [19]. However, these methods provide different results when mammographic images have noise or it is over-improved. The mass shape of the breast can be distinguished by seeing and this can be used to analyses and interpret the mammograms in CAD-based systems. The proposed research work classifies the breast tumor as Benign or Malignant based on geometric structural pattern characteristics of masses using ANN to produce an accuracy of 86.66%. The performance of these systems is affected by the existence of noise, pectoralis, artefacts, and uneven lighting of mammograms. In the proposed work to improve the prediction performance, mammograms are improved by eliminating the noise, artefacts, and pectoral muscles, and improving the contrast.

II. PROPOSED METHOD

A. Flow Diagram

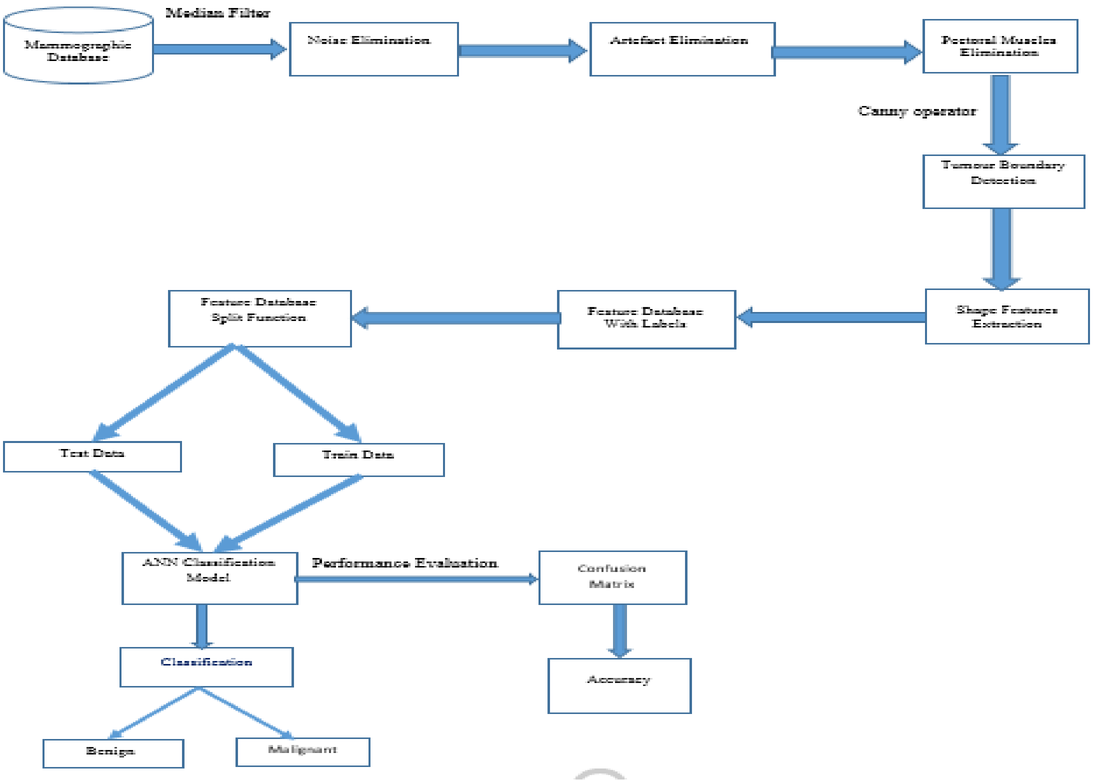

Figure 1. Proposed method flow diagram.

The proposed method steps are illustrated in Figure 1. To enhance the performance of the proposed technique, mammographic images of the dataset should be prepared to process. Pre-processing of the proposed method includes eliminating noise, artefacts, and pectoral muscles, and betterment of mammogram contrast.

B. Noise elimination

Noise is an erratic discrepancy of image intensity that seems as granules in the image. The noise due to dust particles during the image accusation is speckle noise and multiplicative noise. The speckle noise is modelled as

$$I_{speck} = Iorg + \quad N * Iorg \tag{1}$$

Where I_{org} is the original image, N is a noisy image with zero mean and V variance, and I_{speck} speckle noised image.

During breast tumour diagnosis, interpretation of mammograms is difficult for the radiotherapist and the radiologist may confused in cancer diagnosis when the mammogram images are noisy. During the denoising of a mammogram, noise should be eliminated preserving the important features of the image. The linear denoising method does not eliminate noise completely and also produces the blurring effect. Multiplicative noise like speckle noise can be eliminated using a nonlinear median filter. A square window of size 3×3 is taken in a median sieve, and the middle pixel of the square window is restored with the median value of all the pixels of the square window. Applying a median purifier to a window size of 3×3 is shown in Figure 2.

22	34	67
18	56	45
27	65	78

22	34	67
18	45	45
27	65	78

Figure 2. Original values of 3×3 window 3×3 window after applying a median purifier.

An ascending order of 3×3 window values is 18, 22, 27, 34, 45, 56, 65, 67, and 78. The median of these values is the middle value, which is 45. In the entire mammographic image, this 3×3 window is moved from the northeast corner to the southwest corner and the middle value of the window is replaced with the median value of the window values. Figure 3.a illustrates a noisy mammogram and Figure 3.b illustrates a mammogram after eliminating noise using a median filter.

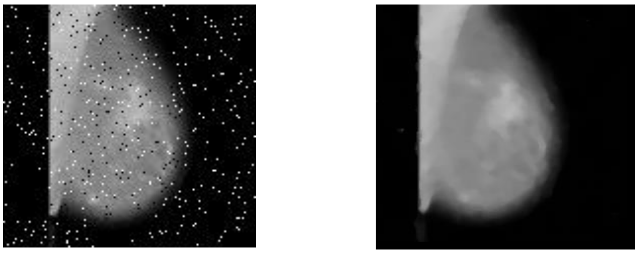

Figure 3. a) Noisy Mammogram b) Mammogram after eliminating noise

C. Artefacts elimination

Artefacts are present in mammograms in the form of identification labels, markers, and wedges in the hidden non-breast region. Usually, these artefacts are radiopaque in the sense that they are not transmitted by X-rays and other radiations. While interpreting mammograms to diagnose breast cancer these may affect prediction performance. To improve prediction performance the artefacts should be removed. The threshold binary connected component (TBCC) algorithm is used to eliminate artefacts. Figure 4 shows a mammogram image with an artefact and an artefact-eliminated mammogram with TBCC. This algorithm's steps are as follows.

1. Provide a grey mammographic image as input.
2. The standard deviation technique is used to find the threshold value of mammographic images.
3. Binarize mammographic images using the above threshold value.
4. Label the binary mammographic image and using equivalence classes find connected components.
5. Find connected components with maximum and second maximum area.
6. Compute the ratio of the maximum and second maximum connected components.

7. If the ratio computed in step 6 is higher, then the highest area component is kept and all other areas are removed, else highest area and second highest area component is kept and all other areas are removed.
8. Convex hull for each pixel in the image is computed and all pixels within the convex hull are set to one.
9. The image produced in step 8 is multiplied by the actual image to produce an image without artefacts.

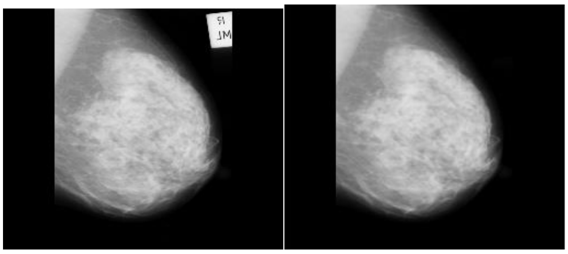

Figure 4. a) Mammogram with artefact b) Mammogram after eliminating artefact

D. Pectoral Muscles Elimination

Pectoral muscles are triangular-shaped regions in the upper portion of the mammographic image that have brighter pixels. Pectoral muscle density is approximately similar to the dense masses of the breast image. Pectoralis identification and removal play a major role in breast tumor identification and increase prediction precision. To simplify this process, instead of applying separate methods for the left and right MLO of mammograms, a method is designed to identify pectoralis in the right MLO of mammograms. If the mammogram is left MLO, then flip it. The output mammographic image of noise and artefact elimination is taken and segmented into the pectoral muscles region and background region. For this binary conversion global thresholding and grey-level thresholding are used. For better results, canny edge detection merged with connected components labelling is used. Figure 5 shows a mammogram with pectoral muscles and a mammogram after removing pectoral muscles. The segmented global and grey threshold connected component labelling (SGGTCCL) algorithm steps are as follows:

1. Input Image ← Noise and artefact eliminated mammographic image.
2. If the Input Image is not right MLO mammographic image
Then Input Image ← flip (Input Image)
3. Value ← 1- global-thresh(Input Image)
4. B_Img ← Img2bimg(Input Image, Value)
5. BB_Img ← Img2bimg(B_Img, gray thresh, std)
6. CEB_Img ← canny_edge(BB_Img)
7. X ← BB_Img – B_Img)
8. X ← bwlabel(X,8)
9. X← find(X==1)
10. X← find(X!=0)

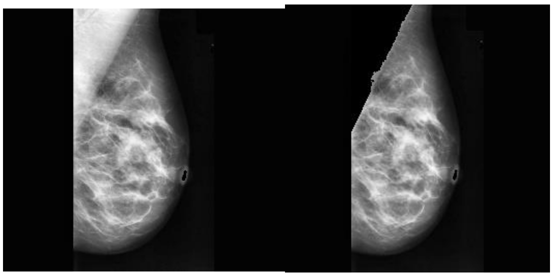

Figure 5. a) Mammogram with pectoralis b) Mammogram after eliminating pectoralis

E. Mass boundary detection

Mammographic images of the dataset contain both benign and malignant masses. These masses boundaries of mammographic images can be detected using a canny edge detector. Figure 6 shows mammograms with masses and mammograms with the boundary of masses using a canny edge operator. The canny edge operator is sensitive to noise. Prior to using the canny edge operator, the mammographic image convolved with a 2D Gaussian filter kernel to eliminate noise. The equation Gaussian kernel of size (2n+1)×(2n+1) is given by:

$$G_{xy} \;== \frac{1}{2\pi\sigma^2}\exp\left(-\frac{\left(x-(n+1)\right)^2+\left(y-(n+1)\right)^2}{2\sigma^2}\right);\, 1 \leq x, y \leq (2n+1) \tag{2}$$

The larger the Gaussian kernel size, edge detector low to noise sensitive. A good size of Gaussian kernel is 5×5 given below:

$$G_{5\times5} = \frac{1}{159}\begin{bmatrix} 2 & 4 & 5 & 4 & 2 \\ 4 & 9 & 12 & 9 & 4 \\ 5 & 12 & 15 & 12 & 5 \\ 4 & 9 & 12 & 9 & 4 \\ 2 & 4 & 5 & 4 & 2 \end{bmatrix}$$

The Canny edge detection operator returns the first derivative in the parallel direction and perpendicular direction as (G_x) and (G_y) respectively. From these values edge slope and angle can be computed as follows:

$$G = \sqrt{G_x^2 + G_y^2} \tag{3}$$

$$\theta = atan2\left(G_y, G_x\right) \tag{4}$$

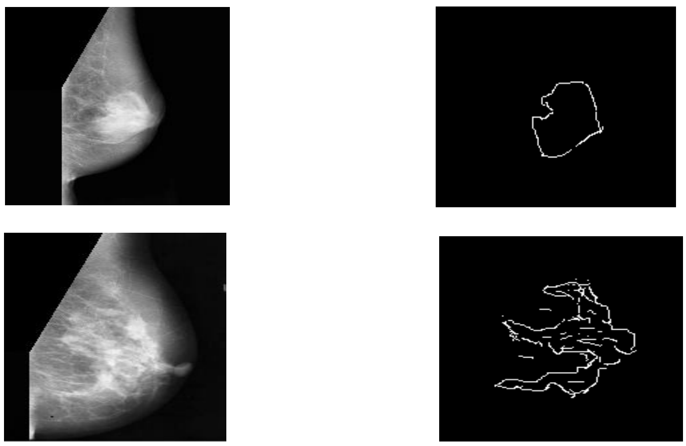

Figure 6. a) Mammograms with mass b) Mass boundary detection

F. Extracting geometric shape features

The masses of the mammographic images are either benign or malignant. Generally, the shape of benign mass is circular or elliptical and mass edges are glossy and delineated. Whereas, malignant masses are uneven outlines and have vague, micro-lobulated or contemplate borders. To discriminate Benign and Malignant masses we considered the count of various geometric shapes like lines, triangles, and squares. Taking a kernel size of 3×3 as shown in Fig. 7, compute the sum of all pixels intensities that lie in various geometric patterns and store in the corresponding array. The various geometric patterns considered in the proposed work are lines, triangles and squares.

1	2	4
8	16	32
64	128	256

Figure 7. 3×3 Kernel

From the mammographic dataset take each image, from the top left consider the kernel of size 3×3 shown in Figure 7 and multiply with the corresponding 3×3 mammographic image array pixels compute the sum of intensities of pixels and check this value will match the with any value of line array, if match increment line count by one, otherwise match the value with any value of triangle array, if match increment triangle count, otherwise match the value with any value of square array, if match increment square count. Move the 3×3 kernel in the entire mammographic image from the top left corner to the bottom right corner calculate the sum of pixels intensities and match with the values of the line array or triangle array or square and increment the corresponding count by one. This process repeats for every mammographic image in the dataset and finds line count, triangle count, and square count and stores them in the feature database. Various lines, triangles and squares in 3×3 kernel and their intensities are shown in Figure 8.

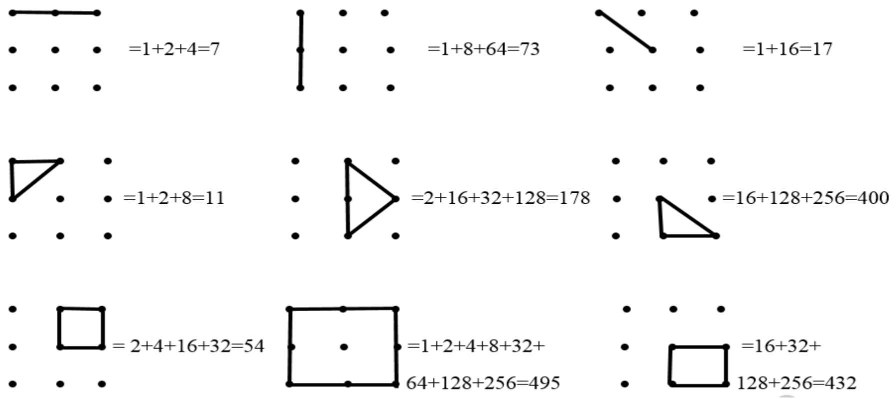

Figure 8. Geometric pattern in 3×3 window and their corresponding intensity values.

The mathematical representation of the line is $px + qy + r = 0$ (5)

where p, q, and r are the coefficients, $-\frac{p}{q}$ is the slope of the line $-\frac{r}{q}$ is the Y-axis intercept.

A triangle is a plane surface enclosed by three-line boundaries. It is a plane and its mathematical representation is

$$px + qy + rz + s = 0 \tag{6}$$

Where *p, q, r,* and *s* are the coefficients.

A square is a plane surface enclosed by four-line boundaries. It is a plane and its mathematical representation is

$$px + qy + rz + s = 0 \tag{7}$$

Where *p, q, r,* and *s* are the coefficients.

The feature database obtained from mammographic images is in below Table 1.

Table 1. Geometric pattern feature database of 100 mammograms

S. No	Lines	Triangles	Squares	Label
1	396	237	143	Benign
2	244	70	142	Malign
3	174	113	136	Benign
4	250	79	139	Malign
5	515	259	205	Benign
6	239	97	123	Malign
7	133	72	131	Benign
8	188	72	116	Malign
9	152	111	77	Benign
10	272	130	161	Malign
11	156	101	102	Benign
12	258	58	102	Malign

13	177	104	92	Benign
. . .	. . .	. . .	. . .	. . .
94	199	37	139	Malign
95	255	149	132	Benign
96	246	112	200	Malign
97	261	211	90	Benign
98	365	232	197	Malign
99	696	228	215	Benign
100	205	56	131	Malign

G. ANN model

Once the boundary of mass is determined, the main task of the proposed work is the categorization of the breast lymphoma into Benign mass or Malignant mass. To classify the breast masses various supervised machine learning algorithms are there. In our proposed work to classify breast masses artificial neural network learning model is used. When we use machine learning to predict whether masses are either benign or malignant, if the prediction is incorrect human intervention is needed to make some parameter adjustments. Whereas artificial neural network models are automatically adjustable to improve performance. Neural networks are arranged in a layered fashion, which makes to learn and make intelligent decisions on their own. Simple Neural Network has Input-Layer, Hidden-Layer, and Output-Layer. A neural network's problem-solving and computational capabilities can be enhanced by raising the number of hidden layers. NN can tolerate errors in a better way and has the ability of parallel processing in which multiple tasks are processed simultaneously. The architecture of ANN is illustrated in Figure 9.

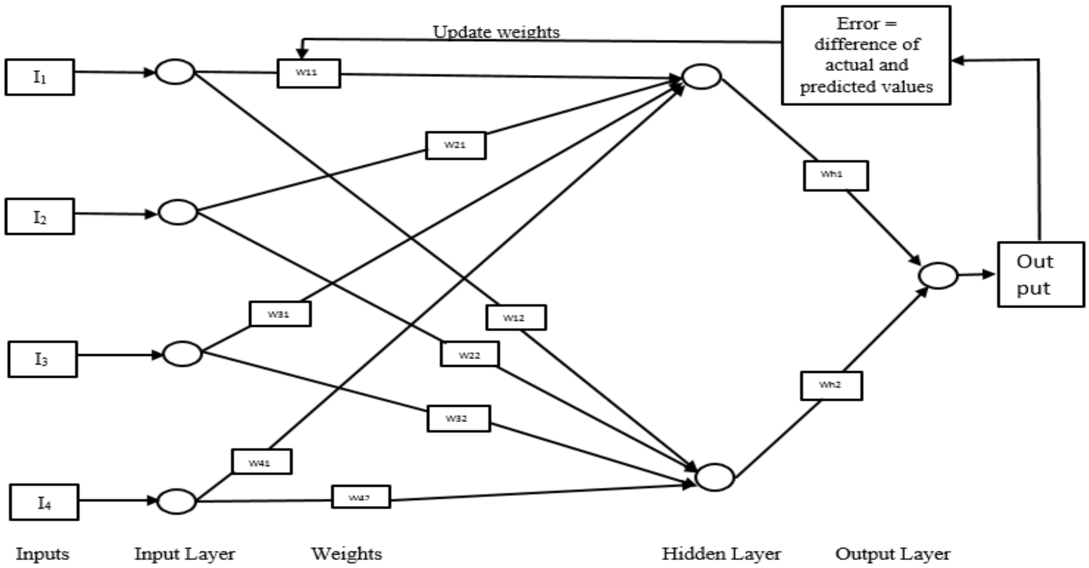

Figure 9. The architecture of simple ANN

Each neuron of the hidden layer takes input and computes output as shown in equation 8.

$$y = g(\sum_{i=1}^{3} w_i \times I_i + dis) \tag{8}$$

Where w_i is the weight, x_i is the input, *dis* is the distortion and *g()* is the rectified linear unit (ReLU) triggering function.

The mathematical model of the ReLU function is shown in equation 9.

$$g(x) = \max(0, x) \tag{9}$$

Each neuron of the output layer computes the output by taking the output of the previous layer as input as shown in Equation 10.

$$Output = h(\sum_{i=1}^{2} h_{iout} \times w_i + dis) \tag{10}$$

Where h_{iout} is the output of i^{th} node of the preceding hidden layer, w_i is the weight of the i^{th} node, *dis* is the distortion and *h()* is the sigmoid triggering function. Equation 11 describes the mathematical model for the sigmoid function.

$$h(x) = \frac{1}{1+e^{-x}} \tag{11}$$

III. Experimental Results and Discussion

The proposed method achievement can be analyzed by considering the publically available benchmark dataset All MIAS for experimentation. The mammographic images present in the All MIAS dataset are 322, of which 207 are normal mammographic, 62 benign masses, and 53 malignant masses. For experimentation, only benign and malignant masses mammograms are considered. The experimental results vary due to noise, artefacts, and pectoralis in mammographic images considered for experimentation. To increase the achievement of the proposed method noise, artefacts, and pectoralis present in the mammographic images should be eliminated. The median filter is used to eliminate noise in the proposed method. TBCC and SGGTCCL algorithms are used to remove artefacts and pectoral muscles and Canny Operator is used to detect tumor boundaries. Feature database is obtained by applying a geometric shape feature extractor to mammographic image data that is free from noise, artefacts, and pectoral muscles, and stored as a CSV file. Simple ANN is used to classify mammographic images of All MIAS datasets based on the geometric pattern feature database. A confusion matrix is used to analyze the achievement of the proposed technique. An error matrix is a table shown in Figure 10.

	Actual	
Predicted	TP	FP
	FN	TN

Figure 10. Error Matrix

True Optimistic (TP): Forecasted as Optimistic, but really Optimistic.

False Optimistic (FP): Forecasted as Optimistic, but really Pessimistic.

True Pessimistic (TN): Forecasted as Pessimistic, but really Pessimistic.

False Pessimistic (FN): Forecasted as Pessimistic, but really Optimistic.

The proposed method is analyzed using the error matrix in terms of rate of correctness, rate of error, exactness, recall, F1score, specificity, and sensitivity measures. These measures are defined as follows:

Accuracy: The ratio of correctly predicted observations of test data to all test data. This parameter is stated in Equation 12.

$$\text{Accuracy} = \frac{TP+TN}{TP+TN+FP+FN} \tag{12}$$

This measure determines how the model predicts correctly.

Error rate: The ratio of all incorrect predictions to all test data. This metric is stated in Equation 13.

Rate of Error = $\frac{FP+FN}{TP+TN+FP+FN}$ (13)

The model is best if this parameter value is 0.0 and worst when 1.0.

Precision: The ratio of correctly predicted positives over the total predicted positives. This parameter is computed as in equation 14.

Precision = $\frac{TP}{TP+FP}$ (14)

This metric evaluates the learning model's positive prediction accuracy.

Recall: The ratio of correct positive predictions over total actual positives in test data. This metric is stated in Equation 15.

Recall = $\frac{TP}{TP+FN}$ (15)

This measure evaluates the learning model's ability to identify the actual true results.

F1score: The harmonic mean between precision and recall is f1score. This parameter is expressed as in equation 16.

F1score = $\frac{2(p*r)}{p+r}$

(16)

This metric can be used to determine the effectiveness of a learning model when more than one learning model is there.

Sensitivity: The ratio of correct positive prediction over actual positives in test data. This measure is stated in Equation 17.

Sensitivity = $\frac{TP}{TP+FN}$ (17)

Specificity: The ratio of correct negative predictions over the actual total negatives in test data. This metric is stated in equation 18.

Specificity = $\frac{TN}{TN+FP}$ (18)

The database images are split up into pedagogy data and examined data in a 7:3 ratio. The ANN model is trained with training data. The trained ANN model is tested with testing data. The error matrix of the present mechanism based on statistical and textural features of mammograms and the newly proposed mechanism are delineated in Figures 11 and 12 respectively. The various statistical performance measures of the present method and the proposed methodology are delineated in Table 2. All the statistical parameters of the proposed mechanism are better than those of the present mechanism. Figure 13 shows the comparison of all these statistical parameters. From this graph in terms of all these parameters, the proposed methodology is much better than that of the present methodology. At the time of pregnancy breast images have more line patterns and this may miss lead to showing malignancy in the breast in the proposed system.

	Actual	
Predicted	13	4
	1	12

Figure 11. Existing method error matrix

	Actual	
Predicted	13	4
	0	13

Figure 12. Proposed method error matrix

Table 2: The performance measures of the present methodology and our proposed methodology are shown in the below table.

Method	Accuracy	Error rate	Precision	Recall	F1score	Sensitivity	Specificity
Existing	0.83333	0.1667	0.83333	0.83333	0.83333	0.84389	0.84389
Proposed	0.86667	0.1333	0.86667	0.86667	0.86667	0.88235	0.88235

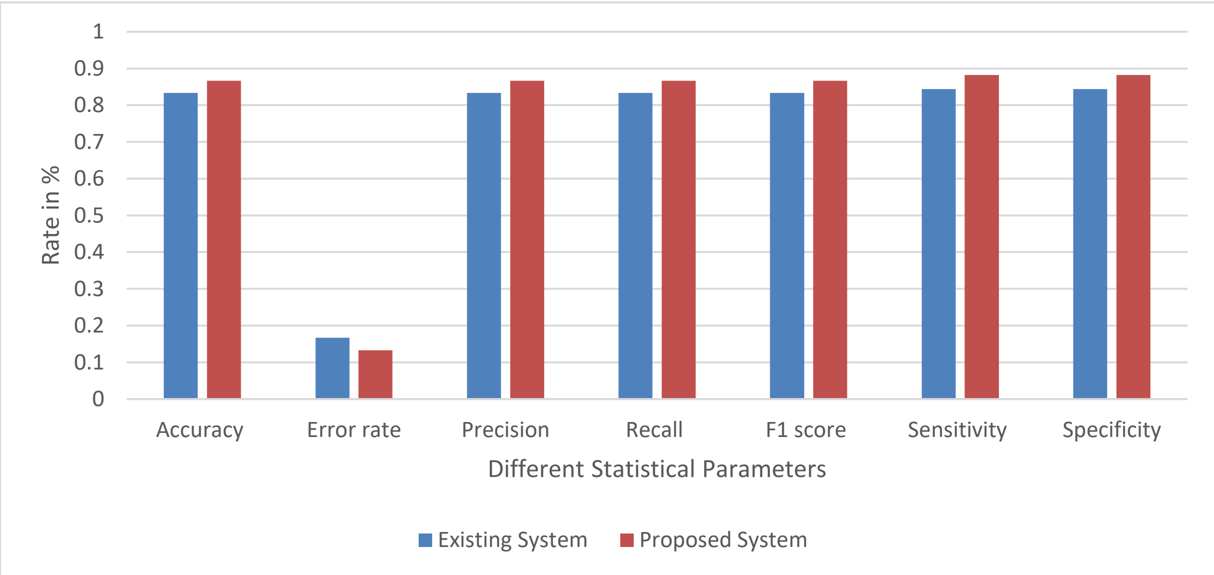

Figure 13. Comparison of different statistical parameters of existing and proposed methods.

The ROC curve shows the performance ANN classification model at different classification thresholds. This curve plots true positive rate and true negative rate. The true positive rate (TPR) is stated in Equation 19 and false positive rate (FPR) is stated in Equation 20.

$$TRP = \frac{TP}{TP+FN} \quad 19$$

$$FPR = \frac{FP}{FP+TN} \quad 20$$

Reducing the classification threshold results in a higher number of positive classifications, which raises the number of True Positives and False Positives. The ROC curve of existing and proposed methods are shown in Figure 14 and Figure 15 respectively.

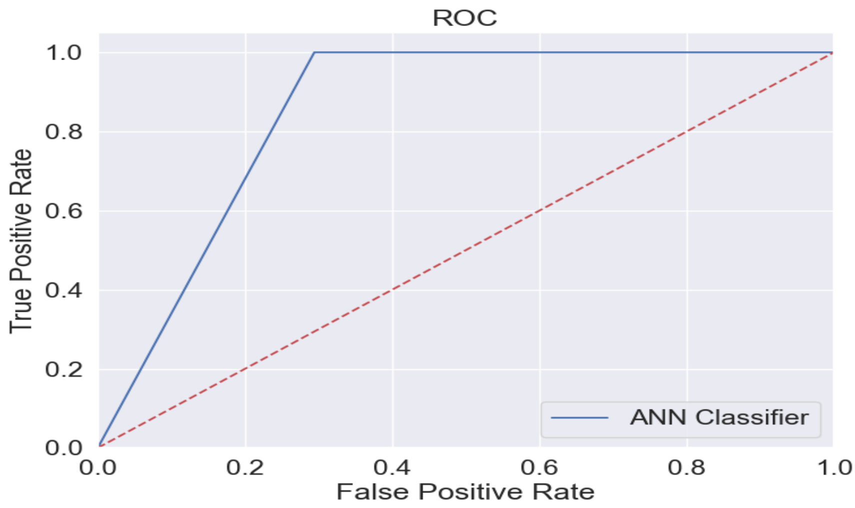

Figure 14. The ROC curve existing method ANN classifier.

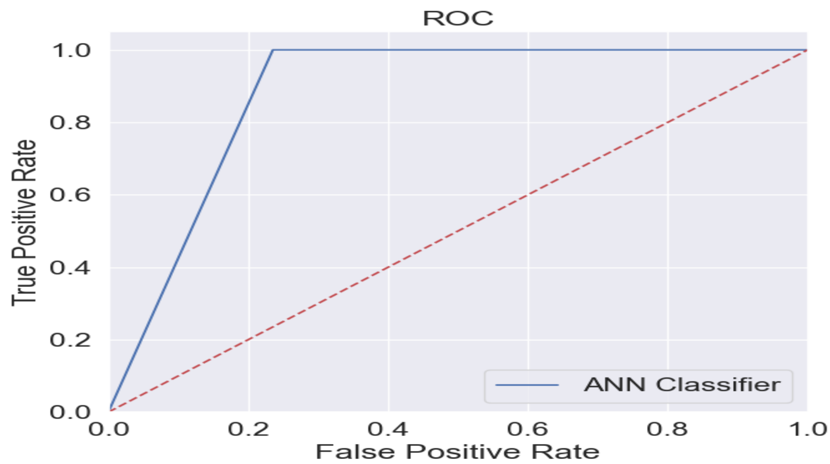

Figure 15. The ROC curve of proposed method ANN classifier.

The ROC AUC score of existing and proposed methods are 0.85294 and 0.88235 respectively. The ROC AUC score indicates the model's ability to generate relative ratings that aid in differentiating between positive and negative occurrences. Instead of producing well-calibrated probability estimates, ROC AUC is particularly helpful when the objective is to rank predictions in order of confidence level. The area under the curve of ROC remains constant and summarizes the results across several classification levels. For data that is unbalanced, it is an appropriate evaluation metric.

IV. CONCLUSION

Screening mammography is an efficient approach for premature disclosure of breast tumors. For radiotherapists, mammographic image analysis becomes difficult due to the presence of noise, artefacts and pectoralis. The proposed work eliminates noise, artefacts and pectoralis effectively and with this radiologists analyze the mammograms more perfectly to find cancer. In this proposed work features extracted are based on the geometric patterns present in masses of mammographic images. The breast tumor is categorized as Benign or Malignant based on these features and using ANN. The prediction accuracy of the proposed work is 86.67% and is better compared with the works based on prediction with the statistical and texture features.

REFERENCES

[1] American Journal of Roentgenology (AJR) (2010) ‘Computer-aided detection improves early breast cancer identification’, available at http://www.ajronline.org (accessed on 5 February 2010.

[2] Abdullahi Isa, Iliyas Ibrahim Iliyas and Muhammad Lefami Zarma “Computational Intelligence Approaches for Enhancing Biomedical Image Processing Applications Based on Breast Cancer”, Biomedical Signal and Image Processing - Advanced Imaging Technology and Application, December 2022.

[3] C. Kaushal, S. Bhat, D. Koundal, A. Singla “Recent Trends in Computer Assisted Diagnosis (CAD) System for Breast Cancer Diagnosis Using Histopathological Images’, **IRBM** Volume 40, Issue 4, August 2019, Pages 211-227.

[4] Bushra Mughal, Muhammad Sharif & Nazeer Muhammad “**Bi-model processing for early detection of breast tumour in CAD system",** The European Physical Journal Plus, Published: 15 June 2017.

[5] Osta, H., Qahwaji, R. and Ipson, S. (2008) ‘Comparisons of feature selection methods using discrete wavelet transforms and support vector machines for mammogram images’, in *5th International Multi-Conference on Systems, Signals and Devices*, pp.1–6.

[6] Surendiran, B. and Vadivel, A. (2009) ‘Classification and regression Tree classifier for classifying benign and malignant masses in digital mammogram using shape features’, in *Proc. Of International Conference on Computing Technologies (ICONCT'09)*, India, pp.47–52.

[7] Ivarani Routray; Nrusingha Prasad Rath “Textural Feature Based Classification of Mammogram Images Using ANN”, 2018 9th International Conference on Computing, Communication and Networking Technologies (ICCCNT).

[8] D. Casanova, J.B. Florindo, M. Falvo, O.M. Bruno “Texture analysis using fractal descriptors estimated by the mutual interference of colour channels", Information Sciences, Volumes 346–347, 10 June 2016, Pages 58-72.

[9] Suhad Abdul Rahman Yousif, Arkan Jassim Mohammed and Nadia Mohammed Ghanim Al-Saidi "TEXTURE IMAGES ANALYSIS USING FRACTAL EXTRACTED ATTRIBUTES", International Journal of Innovative Computing, Information and Control, Volume 16, Number 4, August 2020.

[10] Haipeng Li ,*, Ramakrishnan Mukundan and Shelley Boyd ” Novel Texture Feature Descriptors Based on Multi-Fractal Analysis and LBP for Classifying Breast Density in Mammograms”, . J. Imaging 2021, 7, 205.

[11] João Batista Florindo, Odemir Martinez Bruno “Fractal descriptors based on Fourier spectrum applied to texture analysis”, Physica A: Statistical Mechanics and its ApplicationsVolume 391, Issue 20, 15 October 2012, Pages 4909-4922.

[12] Cheng, H.D., Cai, X., Chen, X., Hu, L. and Lou, X. (2003) ‘Computer-aided detection and classification of microcalcifications in mammograms: a survey’, *Pattern Recognition*, Vol. 36, No. 12, pp.2967–2991.

[13] Vibha, L., Harshavardhan, G.M., Pranaw, K., Deepa Shenoy, P., Venugopal, K.R. and Patnaik, L.M. (2006) 'Classification of mammograms using decision trees, in *IDEAS 2006*, pp.263–266.

[14] Zhang, P., Verma, B. and Kumar, K. (2005) ‘Neural vs. statistical classifier in conjunction with genetic algorithm based feature selection’, *Pattern Recognition, Elsevier*, Vol. 26, No. 7, pp.909–991.

[15] S Mohan Kumar, Dr .G. Balakrishnan ME., PhD “Statistical Features Based Classification of Microcalcification in Digital Mammogram Using Stochastic Neighbor Embedding”, International Journal of Advanced Information Science and Technology (IJAIST) ISSN: 2319:2682 Vol.1, No.7, November 2012.

[16] Mohamed Meselhy Eltoukhy, Ibrahima Faye, Brahim Belhaouari Samir, "A statistical-based feature extraction method for breast cancer diagnosis in digital mammogram using multiresolution representation”, **Computers in Biology and Medicine**Volume 42, Issue 1, January 2012, Pages 123-128.

[17] Abdulkader Helwan, Member, IAENG, Rahib Abiyev "Shape and Texture Features for the Identification of Breast Cancer", Proceedings of the World Congress on Engineering and Computer Science 2016 Vol II WCECS 2016, October 19-21, 2016, San Francisco, USA.

[18] Surendiran B, Vadivel Ayyasamy "Mammogram mass classification using various geometric shape and margin features for early detection of breast cancer", International Journal of Medical Engineering and Informatics 4(1):36-54, February 2012.

[19] B. Surendiran and A. Vadivel "CLASSIFYING BENIGN AND MALIGNANT MASSES USING STATISTICAL and TEXTURAL MEASURES", ICTACT JOURNAL ON IMAGE AND VIDEO PROCESSING, NOVEMBER 2011, VOLUME: 02, ISSUE: 02

[20] Sardar Mehboob Hussain , Domenico Buongiorno , Nicola Altini, Francesco Berloco, Berardino Prencipe, Marco Moschetta, Vitoantonio Bevilacqua, and Antonio Brunetti "Shape-Based Breast Lesion Classification Using Digital Tomosynthesis Images: The Role of Explainable Artificial Intelligence", Appl. Sci. 2022, 12, 6230.

[21] Deepa Sankar, Tessamma Thomas "Fractal Features based on Differential Box Counting Method for the Categ]orization of Digital Mammograms", International Journal of Computer Information Systems and Industrial Management Applications (IJCISIM) ISSN: 2150-7988 Vol.2 (2010), pp.011-019.

[22] P. Narasimhaiah, C. Nagaraju, "*Breast Cancer Screening Tool Using Gabor Filter-Based Ensemble Machine Learning Algorithms", "*Published in International Journal of Intelligent Systems and Applications in Engineering", Volume – 11, Issue – 2, Pages 936 – 947.

[23] P. Narasimhaiah**,** C. Naga Raju " Machine Learning Technique for Prediction of Breast Cancer " published in International Journal on Recent and Innovation Trends in Computing and Communication', ISSN: 2321-8169 Volume: 11 Issue: 7s, Published: 05 June 2023, PAGE NO: 368-380.