Abnormality Detection And Classification Using Artificial Neural Network

R.Rajakumari, L.Kalaivani

Abstract : Among women community, the most dangerous disease is breast cancer. If it is detected in the early stage, the women will be rescued by giving proper treatment. The early detection is possible only by screening in regular interval. It will decrease the mortality rate. Mammography is a specialized medical imaging phenomenon that uses a low-dose x-ray system to see inside the breasts. It is called as mammogram, given support to the early detection and diagnosis of breast diseases in women. In this paper, an automated system is proposed to classify the breast tissues as normal or benign or malignant. Artifacts in the images are removed using Gaussian Mixture Model. Contrast-Limited Adaptive Histogram Equalization (CLAHE) algorithm is used to improve the appearance of the image. The features of the region of mammogram are extracted using hybrid feature extraction which includes Gray Level Co-occurrence Matrix (GLCM), texture and gradient. The features such as contrast, correlation, energy, homogeneity, global mean, uniformity, entropy and skewness are the best features that guarantee the improvement of classification with less feature dimension. K-Means clustering based segmentation is performed to identify the abnormality in the mammogram. The MIAS database images are considered for the evaluation. The feed forward Neural Network classifier is used for classification. Based on the classifier, the given input image is classified as normal or benign or malignant image. From the results, it shows that the proposed breast cancer identification method offers high accuracy and low complexity than the all other existing method.

Keywords : Mammogram, Gray Level Co-occurrence Matrix, Feed forward Neural Network, Receiver operating characteristic curve.

---

1 INTRODUCTION

Breast cancer transpires as one of the main source of deathly diseases among ladies around the world. It is a neoplastic disease, where normal body cells can be transformed into malignant ones. Benign tumor is starting stage and malignant tumor is severe stage. However, there is confirmation that early recognition and treatment can raise the survival rate of breast cancer patients. The guidelines for early detection of breast cancer include breast self-exams (BSE), clinical breast examination (CBE) and screening mammogram. Screening mammography is the most common imaging procedure for diagnosing breast cancer usually among women who have no complaints or symptoms of breast cancer. The main aim of screening mammogram is to detect cancer in starting stage which is to be felt by a woman or her physician. Many pattern recognition algorithms and decision making systems are implemented for breast cancer diagnosis. To interpret and recognize the pattern of the mammogram abnormality, Computer Aided Diagnosis (CAD) is used to help the radiologist. To discriminate normal, benign, and malignant stages of the cancer, the contrast enhancement, feature selection, histogram and gray level co-occurrence matrix (GLCM)[9] are used to extract the suspicious area. To select the best feature, correlation based feature selection (CFS) is used. Some researchers introduced, five individual feature-ranking methods [1] including fisher score, minimum redundancy-maximum relevance, relief-f, sequential forward feature selection, and genetic algorithm[1] for sorting the extracted features and fractal measures[8] for selecting the features with highest ranking to setup a classifier was implemented.

- *R.Rajakumari1 and L.Kalaivani2*
- *Assistant Professor (SG), Department of Computer Science and Engineering, National Engineering College, Kovilpatti. rajakumarinatarajan@gmail.com*
- *Professor, Department of Electrical and Electronics Engineering, National Engineering College, Kovilpatti*
- *anuprakad@gmail.com*

Also, the researchers used three different pattern classifiers Support Vector Machine [3], Linear Discriminate Analysis and Bayes Linear Classifier methodologies were used for calculation of performance evaluation measures. The other work contributed as, the obtained features are dimensionally reduced using Principal Component Analysis and classified through Support Vector Machine method. Mammogram plays a vital role in identifying the abnormal region of breast. Different pattern recognition algorithms and decision making systems were implemented for identifying the abnormality in mammograms. Wenfeng Han et.al,[6] has identified the masses in digital mammogram by iso-contour map method. Extraction of feature is based on textural and shape features. The features are selected by Correlation based Feature Selection (CFS) scheme and classified further to decide the masses or non-masses. N.M.Sangeetha et.al.[7] described a method for detection of breast abnormalities using image processing techniques are Shape extraction and Boundary of mass or lesion, So that signatures can be assigned for detection of breast abnormalities on masses and micro calcification of breast. Edge of a breast mass is one of the indicators of breast abnormality detection [2]. In a mammogram, round and circumscribed masses indicate benign changes and malignant masses usually has speculated (irregular) boundary. B. Verma et al[4] has formed a computer aided diagnosis system to determine a framework for computerize mammograms that give emphasis to neural-hereditary calculation characteristic determination technique and acquired precision was 85% on mammograms from Digital database for screening mammography (DDSM). Pavankumar M.Ahuja et.al.,[5] told that a benign neoplasm is smoothly marginated having regular shape, whereas a malignancy is characterized by an indistinct and irregular border that becomes more speculated with time. Because of the slight differences in X – ray attenuation between masses and benign glandular tissue, they appeared with low contrast and often very blurred. Micro-calcifications are small deposits of calcium that appear as small bright spots in the mammogram. They proposed a staging approach on mammography images using gradient magnitude sobel filter for cancer tumor mass segmentation. After Enhancement, the first step of the cancer signs detection should be a segmentation procedure could distinguish masses and micro-calcifications from background tissue followed by fuzzy classifier. Ankita Satyendra Singh, Prof. M. M. Pawar introduced probabilistic neural network [10] for classification.

2 PROPOSED METHODOLOGY

Various image processing operations on mammograms is performed. The analysis of mammograms is done using suitable techniques. The proposed approach is modeled into four parts image enhancement or improvement, segmentation, feature extraction and classification. The complete architecture is shown in Fig. 1 and their processes are described as follows.

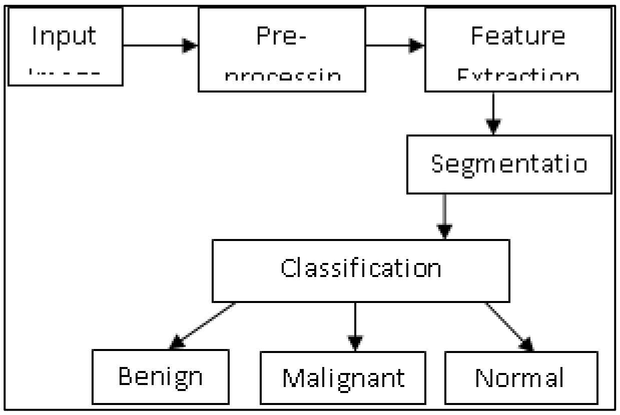
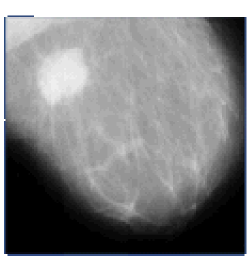


***Fig.1**. Block diagram*

2.1 Pre-processing

Preprocessing is needed to enhance the image for further processing. The original image is a RGB image which is to be converted into gray image for further process. Artifacts in the image may affect the quality of the image. These are removed using Gaussian Mixture Model and the Contrast-Limited Adaptive Histogram Equalization algorithm is used to improve the appearance of the image. Fig. 2 shows the input image and the preprocessed image is shown in Fig. 3.



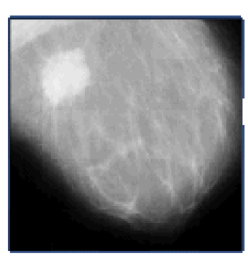


***Fig. 2**. Input image* ***Fig. 3**. Enhanced image*

2.2 Segmentation

The enhanced image is given to k-means segmentation. The algorithm needs the following steps. First it will compute the mean of each cluster. Then compute the distance of each point from each cluster by computing its distance from the corresponding cluster mean. Finally it will assign each point to the cluster it is nearest to. The above steps are continued till the sum of squared within group errors cannot be lowered any more. Finally, the segmentation output is obtained as shown in Fig. 4.

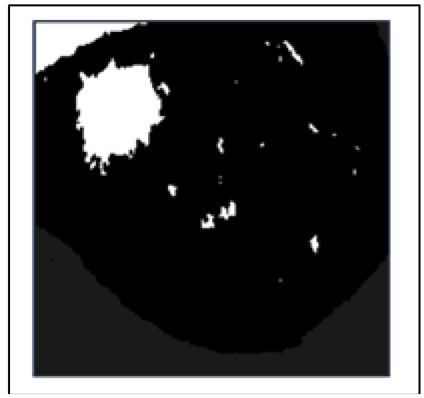


***Fig. 4**. Segmented image*

2.3 Feature Extraction

The MIAS dataset contains 322 images. Among those images, sample 100 images are taken and their GLCM, texture and gradient features are extracted. In texture feature, global mean is extracted and the gradient features like uniformity, entropy and skewness are measured. For simplicity, the features for only six images are listed in Table 1.

Features	*image1*	*image 2*	*image 3*	*image 4*	*image 5*	*image 6*
Contrast	0.0255	0.0398	0.0449	0.0398	0.0428	0.0321
Correlation	0.9961	0.9954	0.9959	0.9965	0.9947	0.9962
Energy	0.537	0.4421	0.48	0.4268	0.3341	0.3602
Homogeneity	0.9916	0.9861	0.9887	0.9892	0.9874	0.992
Global mean	0.0274	0.0203	0.0192	0.0168	0.0155	0.0143
Uniformity	0.004	0.0051	0.0061	0.0052	0.004	0.0039
Entropy	-57.1567	-74.3899	-63.3061	-61.8422	-61.2514	-61.3607
Skewness	13.0334	8.3763	9.141	9.5893	11.7728	12.6526

***Table1:** Statistical features of mammogram images*

2.4 Classification

Table1: Statistical features of mammogram images

2.4 Classification

To classify the images as normal, benign and malignant, a two-layer feed-forward network is used. The network is trained with scaled conjugate gradient propogation. Conjugate gradient methods (CGMs) are general purpose second order techniques which will help to minimize the goal functions of several variables, with sound theoretical foundations. The term Second order means that these methods make use of the second derivatives of the goal function, while first-order techniques like standard backpropagation only use the first derivatives. A second order technique generally finds a better way to a local minimum than a first order technique, but the computational cost is higher than that. CGMs are iteratively trying to get closer to the minimum. But the standard backpropagation always proceeds down the gradient of the error function; a conjugate gradient method will proceed in a direction which is conjugate to the directions of the previous steps. This is a kind of Levenberg-Marquardt method and is done by setting:

$$s_k = \frac{E'(\omega_k + \sigma_k.P_k) - E'(\omega_k)}{\sigma_k} + \lambda_k.\omega_k \qquad (1)$$

Scaled Conjugate Gradient (SCG) has been shown to be considerably faster than standard backpropagation and than other CGMs. As equation (1) and λ_k are computed from their respective values at step k-1, SCG has two parameters, namely the initial values σ_k and λ_k . Their values are not critical but should respect the conditions σ_1 and λ_1 . Empirically Moller has shown that bigger values of $0< \sigma_1 \leq 10^{-4}$ can lead to a slower convergence. The third parameter is the usual quantity $0< \lambda_1 \leq 10^{-4}$. To introduce nonlinearity in the model, Sigmoid functions are used in artificial neural networks. A linear combination of its input signals is computed by the neural network, and sigmoid function is applied to the result. Artificial neurons take in a set of weighted inputs and produce an output through an activation function whereas a hidden layer in an artificial neural network is a layer in between input layers and output layers. Hidden units are the nodes that are situated between the input nodes and the output nodes. Hidden units allow a network to learn non-linear functions and to represent combinations of the input features. Given too many hidden units, a neural net will simply memorize the input patterns and for too few hidden units, the network may not be able to represent all of the necessary generalizations. The goal is for achieving a balance between correct responses for the training patterns and to get correct responses for new patterns. This provides a balance between memorization and generalization. The network will be trained until it reaches an acceptable error rate. Sigmoid hidden neurons and linear output neurons can be fit to multi-dimensional mapping problems arbitrarily well, given consistent data and enough neurons in its hidden layer.

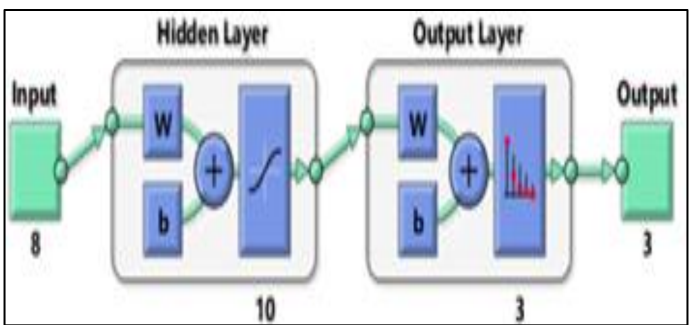


***Fig. 5**. Proposed system Architecture*

The dataset consists of 100 samples. To train the network 70 percentage of data is used, the 15 percentage of data is used for validation and the remaining 15 percentage of data is used for testing the result. And the output neurons set as 3, which is equivalent to the elements in the target vector. The target vector contains three vectors such as normal, benign and malignant. The network used 10 numbers of hidden layers. The architecture is shown in Fig. 5. The data presented to the network during training, it is adjusted according to its error. Validation samples are used to measure network generalization and to halt training when the improvement of generalization is stopped; the cross entropy error (CE) of the validation samples is increased. If there is no effect on training, the testing data provide an independent measure of network performance during and after training. A good classification results minimized cross entropy. If it is zero, it indicates there is no error. Also the percent error indicates the fraction of samples which are misclassified.

Process	Sample data	Cross Entropy	Percent Error
Training	70	$8.12155e^{-1}$	$31.42857e^{-0}$
Validation	15	$1.32278e^{-0}$	$20.00000e^{-0}$
Testing	15	$1.35822e^{-0}$	$33.33333e^{-0}$
Training	70	$8.12155e^{-1}$	$31.42857e^{-0}$

***Table 2:** Performance metrices: Cross entropy and Percent error*

Multiple time training the dataset will generate different results due to the different initial conditions and sampling. In this work, the data chosen is a random process. Hundred iterations are taken place to achieve the maximum performance, which is shown in Table 2. The confusion matrix describes the performance of a classifier on a set of test data for which the true values are known. The specificity is measured as the correct negative predictions divided by the total number of negatives. This is called as true negative rate. Sensitivity of a test is the probability of its yielding true positive results. The confusion matrix for the proposed work is shown in Fig. 6.

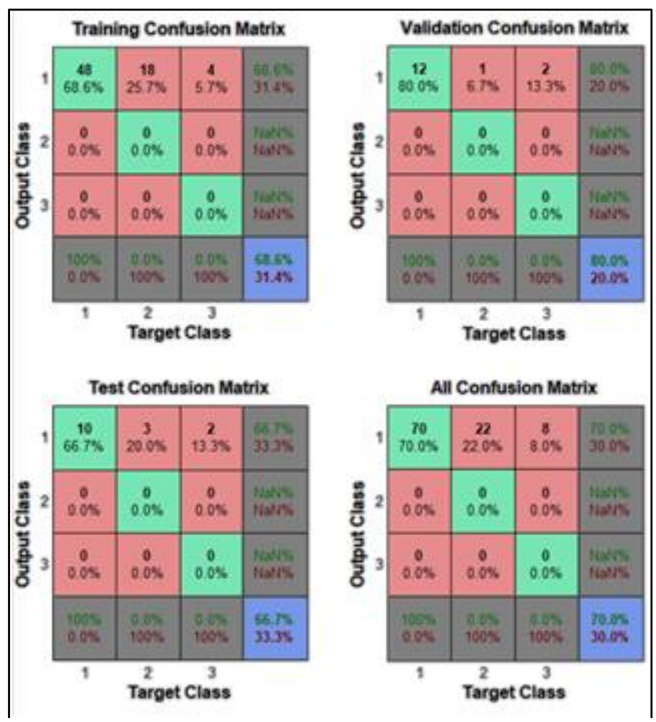


***Fig. 6**. Confusion matrix*

The receiver operating characteristic curve (ROC) represented the specificity and sensitivity of the different classes of data as shown in Fig. 6. The testing result shows that the classification rate is 96% and if the dataset increased, the accuracy will be increased.

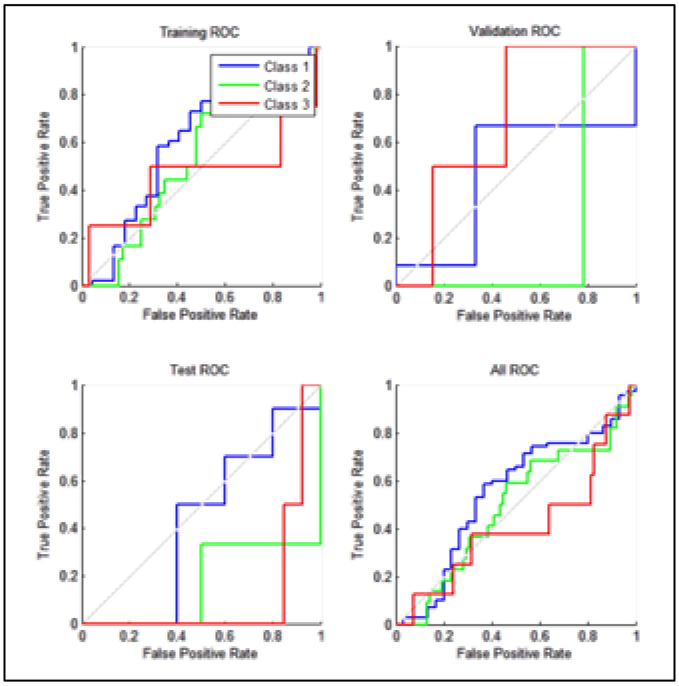


***Fig. 7**. ROC curve*

3 CONCLUSION

In this work, the detection of abnormalities and classification of mammogram image based on a two-layer feed-forward neural network is performed. This work utilizes MIAS dataset. The performance metrics indicates that the work yields better result. Contrast-Limited Adaptive Histogram Equalization (CLAHE) algorithm is used to enhance the image. The salient features such as contrast, correlation, energy, homogeneity, global mean, uniformity, entropy and skewness are the best features that guarantee the improvement of classification with less feature dimension. Based on the classifier, the given input image is classified as normal or benign or malignant image. The confusion matrix and the ROC curve describe the performance of a classifier. From the results, it shows that the proposed breast cancer identification method offers 96% classification accuracy and low complexity than the all other existing method. In future, this work can be extended to the whole dataset(MIAS) and that will provide best performance.

4 REFERENCES

[1] Anette Kniberg and David Nokto, “ A Benchmark of Prevalent Feature Selection Algorithms on a Diverse Set of Classification Problems”, Degree project in medical engineering, second cycle, stockholm sweden 2018

[2] Sarmistha Chakraborty, Mrinal Kanti Bhowmik, Anjan Kumar Ghosh and Tannistha Pal, Automated Edge Detection of Breast Masses on Mammograms, 978-1-5090-2597-8/16/$31.00 © 2016 IEEE.

[3] G.R. Jothilakshmi and Arun Raaza, “ Effective Detection of Mass Abnormalities and its Classification Using Multi-SVM Classifier with Digital Mammogram Images”, IEEE International Conference on Computer, Communication, and Signal Processing (ICCCSP-2017)

[4] A Computer-Aided Diagnosis System for Digital Mammograms Based on Fuzzy-Neural and Feature Extraction Techniques, Brijesh Verma and John Zakos, IEEE Transactions on Information Technology in Biomedicine, Vol. 5, No. 1, March 2001.

[5] Pavankumar M.Ahuja,R.D.Ghongade and Dr.D.G.Wakde, “Detection and Diagnosis of BreastTumors using Digital Mammogram Images: A fuzzy Approach”, International Journal of Science and Research (IJSR), Volume 5 Issue 12, December 2016.

[6] Wenfeng Han, Jianwei Dong, Yuting Guo, Ming Zhang and Jianzhong Wang, “ Identification of Masses in Digital Mammogram Using an Optimal Set of Features”, IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications 2011.

[7] N.M.Sangeetha and Dr.(Mrs).D.Pugazhenthi, “A novel shape and boundary extraction image processing technique for detecting the breast abnormalities using digital mammogram -Labview implementation”, Online International Conference on Green Engineering and Technologies (IC-GET), 2016.

[8] R. Dobrescu, L. Ichim, and D. Crisan, “Diagnosis of Breast Cancer from Mammograms by Using Fractal Measures,” International Journal of Medical Imaging, vol.1, no. 2, pp. 32-38, 2013.

[9] A.K. Mohanty, S. Beberta and S. K. Lenka, “Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammograms,” International Journal of Engineering Research and Applications (IJERA), vol.1, no. 3, pp. 687-693, 2011

[10] Ankita Satyendra Singh, Prof. M. M. Pawar, “Mass Classification of Mammogram Images using Selected Textural Features with PNN Classifier”, 2nd National Conference on Recent Advances in Engineering and Technology (NCRAET_17), Vol.5, Special Issue 4, June 2017.