ISSN: 0975-8585

# Research Journal of Pharmaceutical, Biological and Chemical Sciences

## Hybrid Clustering Scheme for the Classification of Lesions in Mammogram Images.

**V Vedanarayanan*.**

Department of ETCE, Sathyabama University, Chennai -119, Tamil Nadu, India.

## ABSTRACT

Breast Cancer is the major cause for around 30% death of women in this world. Detection of tumors by pathological tests takes more time in which the lesion increases its spreaders. To avoid this problem we propose a hybrid method for the classification of the lesions into benignant and malignant tissues. Hybrid Median Filter (HMF) is used for the noise removal. Then image enhancement is done by mean adjustment technique. K means algorithm and Adaptive Mean Shift Expectation Maximization (AMSEM) algorithm are proposed to segment the breast cancer images. In this work, an efficient approach to search for global threshold of image using Gaussian mixture model is proposed. Firstly, a gray-level histogram of an image is represented as a function of the frequencies of gray-level. Using Gray Level Co-occurrence Matrix (GLCM) feature extraction is done. A fuzzy-neural classifier is used for the classification of the mammogram images in to benignant and malignant. The proposed method evaluated the Mini MIAS data base images.

**Keywords:** hybrid clustering, lesions, mammogram, breast cancer

***Corresponding author***

## INTRODUCTION

Breast cancer is the major cause of death of women in the past few decades. The death rate of women due to breast cancer increases year by year. According to GLOBOCAN (WHO), for the year 2012, an estimated 70218 women died in India due to breast cancer, more than any other country in the world. Currently there are three methods for the detection of lesions in mammogram images which are mammography, MRI, ultrasonography. MRI is very expensive than other two methods. Ultrasonography is useful but its performance is limited to advanced breast cancer. So mammography is the best and economic method among these.

Digital mammogram images are chosen by most of the researchers for their studies [13][14]. By using such images for dense breast tissue also good contrast is resulted. The patient should not be exposed to radiation for long time, and so the image acquisition is very fast. Studies show that still there is a misclassification of about 10% to 25% is present. In this method a new fuzzy neural classifier is used. Accuracy, sensitivity values are obtained for the classifier output.

The rest of the paper is organized as follows. Section II describes about data collection. Section III explains the proposed method. Section IV describes the results obtained from the proposed system. Section V concludes the work.

### Data Set Collection

The images required for this study is taken from (MIAS) Mammographic Image Analysis Society [12]. MIAS is an organization of UK research groups interested in the understanding of mammograms and has a data base of digital mammograms. It includes radiologist's "truth" markings on the portions where abnormalities are present. All the images in MIAS data base are 1024×1024.

## METHODOLOGY

The proposed method is described in this section. The schematic diagram for the proposed method is shown in fig.1. This method consists of the following stages.

- Acquiring the mammogram image.
- Image filtering using Hybrid Median Filter (HMF).
- Contrast enhancement by using mean adjustment technique.
- Segmentation of the mammogram image.
- Feature extraction from the segmented image.
- Classification of the lesions into cancerous or non-cancerous masses.

### Image filtering

Image filtering is done by using a hybrid median filter (HMF) [11].it can be used for both color and gray scale images. No need of approximation from rgb to gray scale. HMF is a filter of non-linear classes which effectively removes impulsive noise and preserves edges. HMF takes two median values, one for horizontal and vertical elements and other for diagonal and off diagonal . HMF has two advantages that it preserves corners and repeated application will not excessively smoothen the image.

### Image Enhancement

In the proposed method mean adjustment technique is used for image enhancement. Mean adjustment is used to improve the contrast brightness and illumination.
Mean adjustment value (MA) is calculated using the equation,
MA= median threshold- mean(image pixel)

Image pixel value is adjusted using the equation, Img= Img + MA*(1-Img)

After finding the maximum and minimum intensity value from image mean adjustment is done by the equation,

Img=(Img-minimum)/(maximum-minimum)

After performing this step intensity enhanced image is obtained.

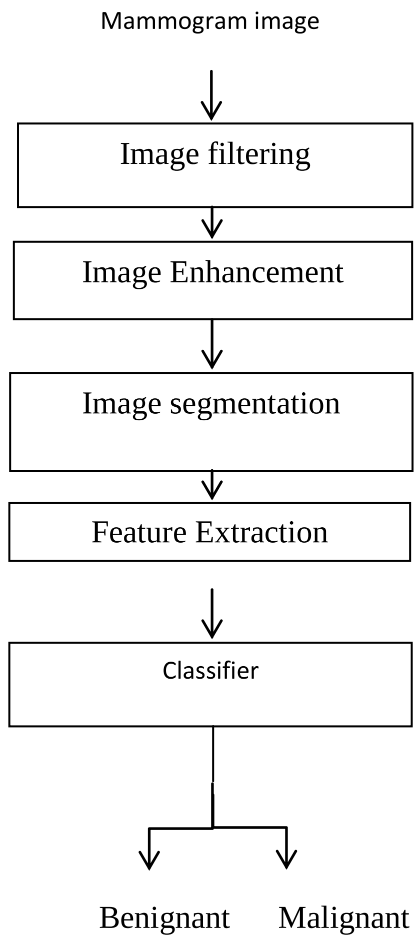



**Figure1: Schematic representation of the classifier**

**Image Segmentation**

A combination of k-means algorithm and Adaptive Mean Shift Expectation Maximization (AMSEM) is used for segmentation.

**K-means algorithm**

The k-means algorithm assigns feature vectors to clusters by the minimum distance assignment principle [10], which assigns a new feature vector x(q) to the cluster c(k) such that the distance from the x(q) to the center of c(k) is the minimum over all k clusters. The basic k-means algorithm is as follows:

X1, X2,..,Xp is the set of observations and M1,M2,…,Mp is the initial set of k-means.

- Put 1st k feature vectors as initial centers.
- Assign each sample vector to the cluster with minimum distance assignment principle.
- Compute new average as new center for each cluster.
- If any center has changed, then go to step 2, else terminate.

**AMSEM Algorithm**

Basic structure of the proposed algorithm is given as follows:

Start with random parameter values for the two models.
Iterate until parameter values converge:
Expectation step: assign points to the model that fits it best.

Maximization (M) step: update the parameters of the models using only points assigned to it.

In fact both steps are slightly more complicated, due to the assignment being continuous rather than binary valued. The following sections go into more detail regarding the E step and the M step.In the E step we compute for each data point two weights w1(i);w2(i) (the soft assignment of the point to models 1 and 2 respectively1). Again, we assume that the parameters of the processes are known.

After executing this AMSEM algorithm the image pixels are grouped together into 3 groups. In which the high intensity region shows the region where lesions are present. And features are extracted from that particular region.

**Feature extraction**

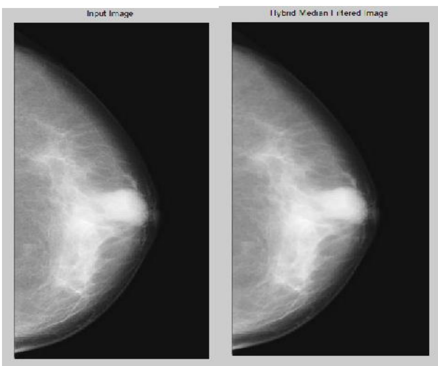
For the study 23 features are selected. One 1st order feature and 22 2nd order textural features are selected in this idea. Gray Level Co-occurrence Matrix (GLCM) is used for the feature extraction [8][5][4]. Area is the 1st order feature and entropy, energy, dissimilarity, contrast, correlation, homogeneity and 16 other features are measured [9][7][6]. The extracted features are given as the input for the classifier.

**Neural classifier**

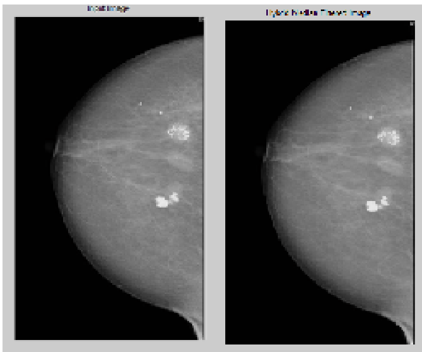
After extracting the 23 features, for finding whether there is any cancerous mass present in the mammogram image, these extracted features are the membership values for fuzzy neural classifier [2]. And the membership values will be the input for classifier. According to the membership values the images are classified into malignant or benignant.For this classification a 3 layer Back Propagation Network is used. It consist of an input layer 3 hidden layers and an output layer.

## RESULTS AND DISCUSSION

Mammogram images of 12 patients are used to estimate the performance of the proposed idea. Different training pairs with different resolutions are collected for training the neural network.

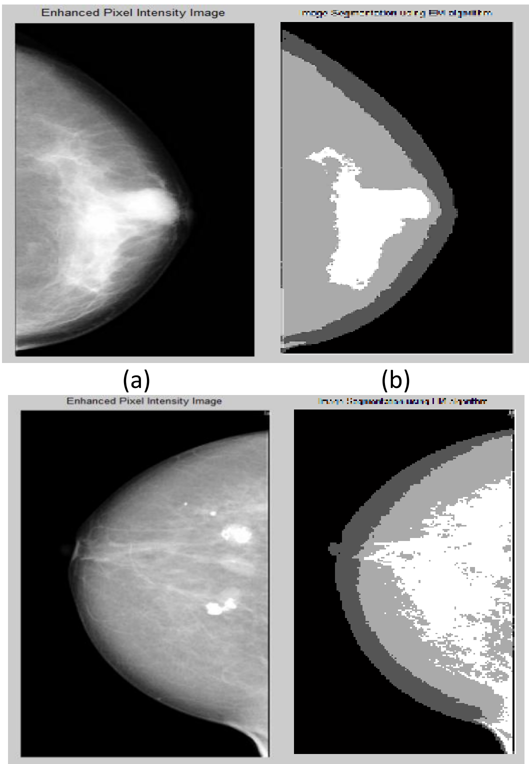



(a) (b)



(c) (d)

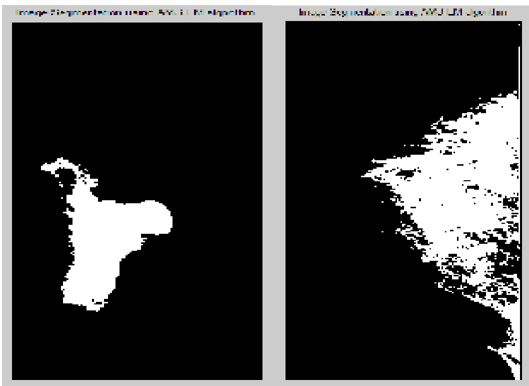
**Figure 2-(a),(c)input images; (b), (d) HMF output**




(a) (b)

(c) (d)

**Figure 3-(a),(c) shows the enhanced pixel intensity images (b), (d) shows the k-means segmentation outputs.**



(a) (b)

**Figure 4-(a),(b) shows the AMSEM method of segmentation tumor regions.**

The following picture is that of the Artificial Neural Network architecture (ANN). A three layer BPN with LM training algorithm is used. Performance of the network is evaluated based on the MSE value. Membership values for training are obtained from Fuzzification process.

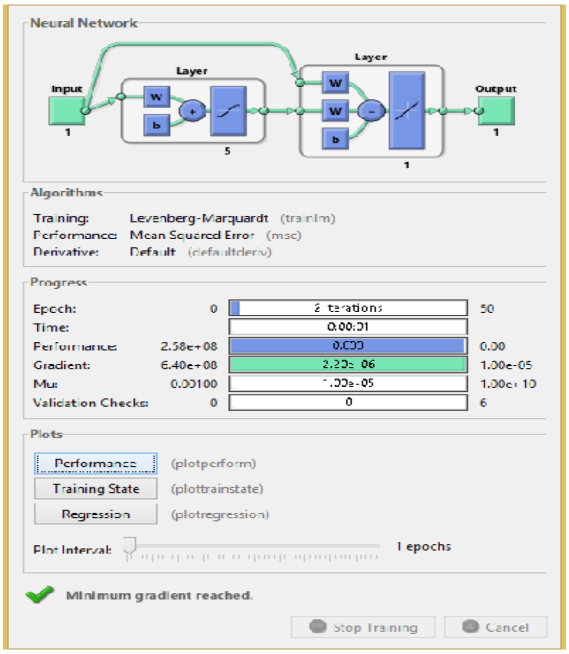



**Figure 5-Neural network tool box**

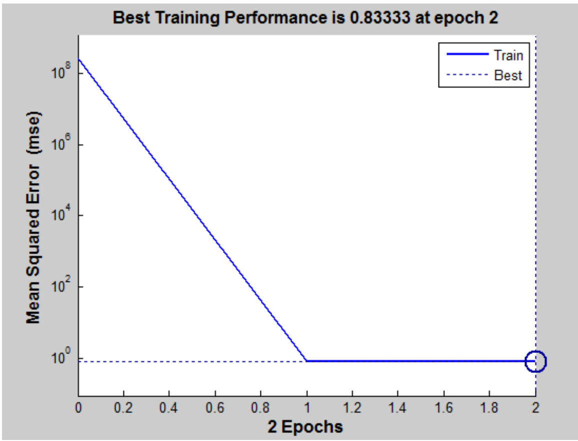



**Figure 6- best training performance graph**

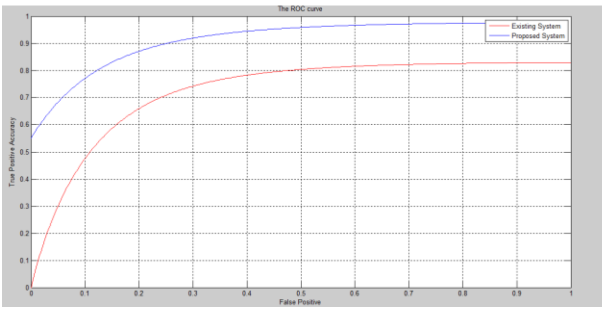


**Figure 7-comparison graph**

The performance graph of the ANN. Best training for the selected input image is reached at $2^{nd}$ Epoch itself.

Fig.7 given below is the comparison graph between false positive and true positive accuracy for the existing method and the proposed method [1]. From Fig 7 effective classification of the tumor can be noticed.

| Image | Accuracy | Sensitivity (out of 25) |
|---|---|---|
| Image 1 | 82.7192% | 25 |
| Image 2 | 88.7379% | 25 |
| Image 3 | 64.8748% | 25 |
| Image 4 | 78.3937% | 25 |
| Image 5 | 82.0163% | 25 |

**Table 1- accuracy and sensitivity**

## CONCLUSIONS

This research work provides a non-invasive method for detection of lesion and classification of lesions present in the digital mammograms. First the mammogram image is preprocessed using HMF and mean adjustment is done for contrast enhancement. Segmentation is performed by a combination of k-means algorithm and AMSEM algorithm. Features are extracted from the segmented image. Classification is done using a fuzzy based neural classifier. Based on the training, lesions are classified as benignant and malignant. Accuracy and sensitivity of the proposed technique is much better than the existing methods which we can infer from Fig 6. Further this research can be extended to training of more set of images and can be compared with other evolutionary computation methods like Genetic Algorithms etc.

## REFRENCES


[1] Samulski M, Karssemeier N. “Optimizing case based detection performance in a multiview CAD system.”DOI:10.1109/TMI.2011.2105886. Page(s):1001-1009

[2] SerifovicTrbalica.A, Demirovic.D, Prlaca.N, Cattin P.C. “Classification of benignant and malignant masses in breast mammograms”. Information and Communication Technology, Electronics and Microelectronics(MIPRO).2014 DOI:10.1109/MIPRO.2014Pages(s):228-233

[3] Radovic.M, Dokovic.M, Peuilic.A, Filipovic.N. “Application of data mining algorithm for mammogram classification.” Bioinformatics and Bioengineering(BIBE).2013.6701551 Page(s):1-4

[4] Nayak.A, Ghosh D.K, Ari S. “Suspicious lesion detection in mammograms using undecimated wavelet transform and adaptive thresholding.” Advanced Computing Technologies (ICACT), 2013. DOI:10.1109/ICACT,2013.6710546

[5] Mohamed Fathima M, Manimegalai D, Thaiyalnayaki S. “Automatic detection of tumor subtype in mammogram based on GLCM and DWT features using SVM.” Information Communication and Embedded Systems (ICICES) DOI: 10.1109/ICICES.2013.6508213. Page(s):809-813

[6] Korkmaz S.A, Eren.H “Cancer detection in mammograms estimating feature weights via Kullback-Leibler measure.” Image and Signal processing (CISP)2013DOI:10.1109/CISP.2013.6745208 Page(s):1035-1040.

[7] ParamkushamS.Rao, PrabhakarRao, “Early stage detection of breast cancer using novel image processing techniques.” (ICACT). DOI:10.1109/ICACT.2013.6710511 Page(s):1-5

[8] Nguyen ,V.D, Nguyen D.T, Nguyen H.L, Bui D.H, Nguyen T.D “Automatic identification of massive lesions in digitalized mammograms.” Communications and Electronics (ICCE) DOI:10.1109/CCE.2012.6315919. Pages:313-317

[9] Shen-Chuan Tai, Zih-Siou Chen, Wei-Ting Tsai, “An automatic mass detection system in mammogram based on complex texture features.” Biomedical and Health informatics, DOI: 10.1109/JBHI.2013.2279097 Page(s):618-627

[10] Hassan S.A, Sayed M.S, Farag F, “Segmentation of breast cancer lesion in digitized mammogram image.” Biomedical Engineering Conference. (CIBEC)DOI: 10:1109/CIBEC.2014. 7020928 Page(s):103-106.

[11] Rakesh M.R, Ajeya B, Mohan A.R “Hybrid mesian filter for impulse noise removal of an image in image restoration.” International journal of advanced research in electrical electronics and instrumentation engineering.Vol2,issue 10, October2013

[12] De oliveira Silva, Kardec Barros, Carvalho Santana. “A telediagnosis system for automatic detection of lesions in digital mammograms”. Bio signals and bio robotics Conference. DOI:10.1109/BRC.2014.6880970 Page(s):1-6

[13] Ganesan K, Acharya U.R, Chua C.K, Lim C.M, Abraham K.T. “One class classification of mammogram using trace transform functional.” Instrumentation and Measurement, IEEE TransactionDOI:10.1109/TIM.2013.2278562 Page(s):304-311

[14] Dae Hoe Kim, Jae Young Choi, Yong Man Ro. “Region based stellate features for classification mammographic speculated lesions in computer aided detection.” Image Processing (ICIP) DOI:10.1109/ICIP.2012.6467486. Page(s): 2821-2824.