ELM BASED CAD SYSTEM TO CLASSIFY MAMMOGRAMS BY THE COMBINATION OF CLBP AND CONTOURLET

S. Venkatalakshmi[1] and J. Janet[2]

[1]*Department of Computer Science Engineering, Panimalar Institute of Technology, India*
[2]*Sri Krishna College of Engineering and Technology, India*

Abstract

Breast cancer is a serious life threat to the womanhood, worldwide. Mammography is the promising screening tool, which can show the abnormality being detected. However, the physicians find it difficult to detect the affected regions, as the size of microcalcifications is very small. Hence it would be better, if a CAD system can accompany the physician in detecting the malicious regions. Taking this as a challenge, this paper presents a CAD system for mammogram classification which is proven to be accurate and reliable. The entire work is decomposed into four different stages and the outcome of a phase is passed as the input of the following phase. Initially, the mammogram is pre-processed by adaptive median filter and the segmentation is done by GHFCM. The features are extracted by combining the texture feature descriptors Completed Local Binary Pattern (CLBP) and contourlet to frame the feature sets. In the training phase, Extreme Learning Machine (ELM) is trained with the feature sets. During the testing phase, the ELM can classify between normal, malignant and benign type of cancer. The performance of the proposed approach is analysed by varying the classifier, feature extractors and parameters of the feature extractor. From the experimental analysis, it is evident that the proposed work outperforms the analogous techniques in terms of accuracy, sensitivity and specificity.

Keywords:
Breast Cancer, Microcalcification, Mammogram, Classification

1. INTRODUCTION

The most common kind of cancer among women is the breast cancer. Indian Council of Medical Research has reported that over 14.5 lakh women are diagnosed with breast cancer in 2016. In spite of the skyrocketing rise of the breast cancer affected patients, the awareness about breast cancer is inadequate. Hence, there is no surprise that the ICMR has accentuated that only 12.5% of the patients detect breast cancer at initial stage [1]. The priceless benefit being reaped by the patients, who detected the breast cancer during their initial stage is the lengthened lifespan. Today's world has assigned computers as the technological companion to mankind.

Computers are exploited in almost all the activities of mankind. The computer technology has rendered so many useful tools, through which the advanced medical technology can thrive effectively. Digital mammography is an advanced medical imaging technology, which acts as a screening tool for breast cancer diagnosis. Computer Aided Diagnostic (CAD) systems are meant to detect and highlight the suspicious regions in a mammogram. The physician monitors the suspicious regions being highlighted by the CAD system and makes final decision [2].

The microcalcifications are the minute deposits of calcium. In most cases, microcalcifications are benign and are harmless. However in certain cases, microcalcifications are malignant.

The category of microcalcification is decided by taking its location, shape and appearance into account. Hence, it is beneficial to present a CAD system, which can categorize the microcalcifications found in the mammograms. Generally, a standard CAD system has got three significant phases and they are medical image pre-processing, feature extraction and classification. Among these three phases, feature extraction is the deciding phase, as the classifier differentiation is done based on the feature sets. The speed and accuracy of a CAD system depends on the sharpness and compactness of the feature set [3].

This paper presents a CAD system for detecting and classifying the microcalcifications being present in the mammograms. The objective of this work is attained by incorporating significant phases such as mammogram pre-processing, segmentation, feature extraction and classification. The mammogram pre-processing phase prepares the image to make it suitable for the forthcoming image processing phases. The segmentation process aims at extracting the regions of interest from the mammograms. The next phase extracts the features from the regions of interest, which are so detected and the feature vectors are formed. Finally, the classifier makes decision about the class of microcalcification in the mammogram.

Initially, the proposed work pre-processes the mammogram images by employing adaptive median filter [4] to de-noise and smoothen the image. The mammogram image is segmented by means of Generalized Hierarchical Fuzzy C Means algorithm (GHFCM) [5], which is an enhanced version of the traditional FCM algorithm. The segmentation process is followed by the feature extraction, which is achieved by Completed Local Binary Pattern (CLBP) and Contourlet. The multiclass Extreme Learning Machine (ELM) is employed as the classifier and is responsible for deciding the kind of microcalcification. The merits of this work are as follows:

- The adaptive median filter de-noises, smoothens and preserves the edge information of the mammogram image.
- The GHFCM segments the pre-processed image, as it is proved as robust and efficient [4].
- The texture features are extracted from the segmented images by employing CLBP and contourlet.
- As ELM is utilized for classification, the classification task is proved to be fast, efficient and scalable.

The remaining content of the paper is organized as follows. The related literature with respect to microcalcification detection is reviewed in section 2. The proposed approach for detection and classification of microcalcification is presented in section 3. In section 4, the performance of the proposed work is analysed and compared with the existing techniques. The

concluding remarks of the proposed work are presented in section 5.

2. REVIEW OF LITERATURE

This section intends to review the related literature with respect to microcalcification detection. Texture is the most important quality of an image and the thus, several applications such as medical image processing, image based security applications, remote sensing make use of texture descriptors for feature extraction process [6]. Numerous works are there in the existing literature, which exploit texture feature for mammogram classification.

In [7] a mammogram classification model, which used wavelet and curvelet for feature extraction process is proposed. The features being extracted are sieved through by statistical t-test and the Support Vector Machine (SVM) is employed as the classifier. The work in [8] presents a mammogram classification scheme which extracts features from the regions of interest. The remarkable features are then selected by independent component analysis. Finally, the abnormal regions are detected by the neural network. A system to detect microcalcification in mammograms is presented in [9] and is based on wavelets. The texture features of first order statistical, co-occurrence matrices, run length matrices and energy measures are calculated. Finally, k-NN classifier classifies between the mammograms.

In [10], a texture feature based mammogram classification model is presented, which used Grey Level Co-occurrence Matrix (GLCM) features being extracted from curvelet bands. Simple logistic classifier is employed for the purpose of classification. The work presented in [11] utilized Local Binary Pattern (LBP) and Co-occurrence matrix to extract texture features from the regions of interest. Additionally, this work incorporates Leave One Out (LOO) strategy with the SVM, in order to attain good classification rate. This work is extended in [12], so as to reduce the false positive rates by utilizing LBP texture features. In [13], the regions of interest are segmented and then the features are extracted by texture and morphological feature extractors. This work has shown minimal false positive rate.

In [14], the Completed Local Binary Pattern (CLBP) and wavelets are utilized for feature extraction and then the classification process is followed. The work presented in [15] intended to classify normal and abnormal mammograms by means of extracting texture features through curvelet. The extracted texture features are then filtered by taking the Euclidean distance into account and the classification is achieved by k-NN. In [16], the normal and abnormal regions are detected by utilising multiresolution LBP features.

This paper strives to introduce an accurate microcalcification classification scheme by taking the related works into account. In order to achieve maximum accuracy, this work utilizes CLBP and contourlet for extracting rich set of texture features. As an added feather to the crown, the classification process is achieved faster when compared to the existing techniques, as the proposed work employs ELM as the classifier.

3. MAMMOGRAM CLASSIFICATION SCHEME

The proposed mammogram classification scheme is explored in this section. This section aims to present the overall idea and all the sub-phases of the proposed work in detail.

3.1 PROPOSED APPROACH

This paper presents a mammogram classification scheme, which differentiates between the normal, benign and malignant kind of cancer. The entire work is subdivided into sub-phases such as mammogram pre-processing, segmentation, feature extraction and classification. The image pre-processing step aims to remove noise and to improve the quality of an image. This work employs adaptive median filter which removes noise and controls the distortion effectively by smoothening the image. The pre-processed image is segmented by GHFCM, which is proved to be resilient against noise. The segmented images are then fed as input to the feature extraction stage. The texture features are extracted from the mammogram images by CLBP and contourlet and the feature vectors are formed. In the next step, the ELM is trained with the feature vector, such that the classifier can learn about the normal and abnormal regions of mammograms. The overall scheme of the proposed approach is presented in Fig.1.
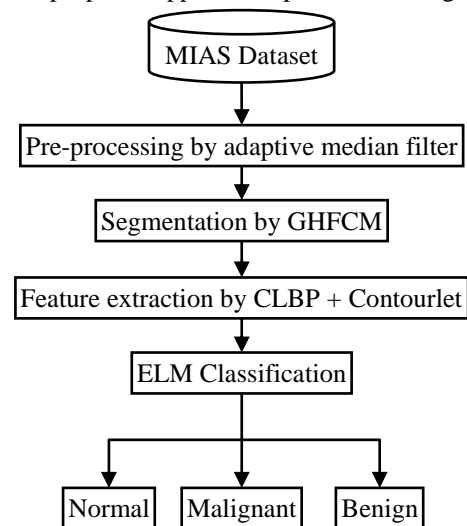

Fig.1. Overall scheme of the proposed work

In the training phase, the system is trained with the extracted feature vectors and the knowledge is imparted to the ELM. In the testing phase, an input mammogram image is passed. The usual processes such as pre-processing, segmentation and feature extraction are carried out and the feature vector is returned. The ELM takes this feature vector as input and compares it with the trained set of feature vectors. The outcome of ELM can be either one of normal, malignant or benign. By this way, the proposed work can assist physicians in taking final decisions. The proposed work is proved to be accurate, fast and reliable.

3.2 MAMMOGRAM PRE-PROCESSING

This work attempts to pre-process a mammogram image by means of adaptive median filter. This function of this filter is to perform spatial processing, so as to detect the noisy pixels of an image. The noisy pixel is detected by comparing the value of the

corresponding pixel to its neighbourhood pixels. Whenever a pixel is detected with unrelated value when compared to the neighbourhood pixels, then that particular pixel is substituted with the median value of the neighbourhood window. The advantage of adaptive median filter is that the size of neighbourhood window is not fixed, as in median filter. The Fig.2 depicts the pre-processing step.
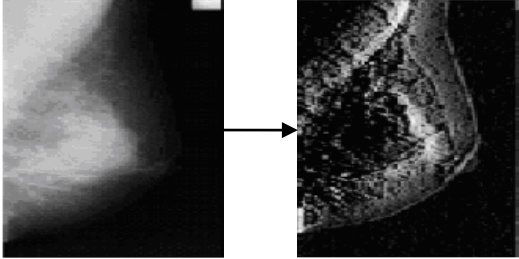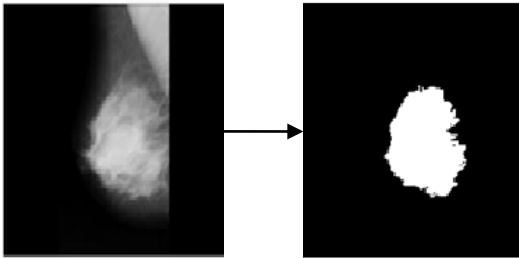

Fig.2. Mammogram image pre-processing

Fig.3. Mammogram segmentation

The centre pixel of the window is checked for the presence of impulse noise. Suppose, if the centre pixel is impulse then it is changed with the median value of the window. In case, if the centre pixel is not impulse, then the same value is maintained. The size of the window is extended, if the median value is found to be impulse. This means that the values are processed for a larger window and the median value is quite optimal. This process continues for all the pixels in the image which results in a good quality mammogram image. Besides this, the adaptive median filter conserves the image details. After the completion of image pre-processing, the mammograms are then segmented to make the whole process easier.

3.3 MAMMOGRAM SEGMENTATION

Segmentation itself is a separate area of research and fuzzy based segmentation techniques are quite popular. The main merit of fuzzy based segmentation techniques is the capability to preserve the details of an image [17,18]. The standard FCM algorithm is proved to be efficient only with the noise images because of the incorporation of Euclidean distance. The Euclidean distance doesn't handle spatial operations and is inefficient when it comes to noisy image. In order to overcome these issues, the proposed work utilizes sub-FCM instead of Euclidean distance. Additionally, the algorithm presented in [5] has enforced the utilization of generalized mean over membership, in order to club local spatial and component details. On the other hand, the distance function clubs the local spatial details with the observation details. The image segmentation process is presented in Fig.3.

This paper utilizes the algorithm GHFCM proposed in [5], which combines the Hierarchical FCM (HFCM) [19] and Generalized FCM (GFCM) [20] together. Thus, the benefits of both the algorithms are reaped by GHFCM. By taking both HFCM and GFCM together into account, the objective function of GHFCM is provided as follows.

$$Ob_f = \sum_{i=1}^{N}\sum_{x=1}^{X}\sum_{y=1}^{Y} a_{ix}^g b_{ixy}^h \sum_{d \in NH_i} l_d m_{dxy} \qquad (1)$$

The Eq.(1) can be written as,

$$Ob_f = \sum_{i=1}^{N}\sum_{x=1}^{X}\sum_{y=1}^{Y} a_{ix}^g b_{ixy}^h l_d \left\| k_c - \mu_{xy} \right\|^2 \qquad (2)$$

where $i = \{1,2,…,N\}$ is the dataset, which contains N data items. X is the count of clusters, Y is the count of subclasses. The degree of membership of k_i in x^{th} cluster is represented by a_{ix} and g is the weight exponent of the fuzzy membership function a_{ix}. b_{ixy} is the sub-membership that satisfies the conditions $\sum_{x=1}^{X} a_{ix} = 1$ and $\sum_{y=1}^{Y} b_{ixy} = 1$. l_d is the weighing factor which controls the impact over the distance between the corresponding and the centre pixel. NH_i is the neighbourhood of the ith pixel. k_c is the image intensity and μ_{xy} is the cluster centre. m_{dxy} is the sub-distance function, which is Euclidean distance. The following equations present the computation of a_{ix}, b_{ixy} and μ_{xy}.

$$a_{ix} = \frac{\left(\sum_{y=1}^{Y}\sum_{d \in NH_i} l_d b_{ixy}^h m_{dxy}\right)^{\frac{1}{(1-g)}}}{\sum_{p=1}^{X}\left(\sum_{y=1}^{Y}\sum_{d \in NH_i} l_d b_{ipy}^h m_{dxy}\right)^{\frac{1}{(1-g)}}} \qquad (3)$$

$$b_{ixy} = \frac{\left(\sum_{d \in NH_i} l_d b_{ixy}^h m_{dxy}\right)^{\frac{1}{(1-h)}}}{\sum_{p=1}^{Y}\left(\sum_{d \in NH_i} l_d a_{ipx}^g m_{dxy}\right)^{\frac{1}{(1-h)}}} \qquad (4)$$

The centroid of the cluster μ_{xy} is computed as follows:

$$\mu_{xy} = \frac{\sum_{i=1}^{N}\sum_{d \in NH_i} a_{ix}^g b_{ixy}^h K_C}{\sum_{i=1}^{N} a_{ix}^g b_{ixy}^h} \qquad (5)$$

Thus, the local weighted generalized mean that takes the spatial and cluster information into account, is computed. The altered membership and the sub-membership is given by,

$$a_{ix} = \frac{\sum_{d \in NH_i} l_d a_{dx}}{\sum_{p=1}^{Y}\sum_{d \in NH_i} l_d a_{dp}} \qquad (6)$$

$$b_{ixy} = \frac{\sum_{d \in NH_i} l_d a_{dxy}}{\sum_{p=1}^{Y}\sum_{d \in NH_i} l_d b_{dxp}} \qquad (7)$$

Thus, the mammograms are segmented by the GHFCM, which employs the fuzzy objective function that takes the hierarchical distance function and spatial constraints into account. This improves the robustness and efficiency of the clustering algorithm. As soon as the mammograms are clustered, the features are extracted from the images.

3.4 FEATURE EXTRACTION – PRELIMINARIES

3.4.1 CLBP:

The feature extraction is the heart of a CAD system, as this phase extracts meaningful features from the mammograms to identify the abnormalities. The features are extracted by the combination of CLBP and contourlet. CLBP is the enhancement of LBP and is proposed in [21] and is claimed as effective texture feature descriptor. CLBP represents a local region by means of a centre pixel and Local Difference Sign-Magnitude Transform (LDSMT). LDSMT is the subtraction of pixel values with the local central pixel from the magnitude. The central pixel is characterized by the binary global threshold map, which is named as $CLBP_{Centre}(CLBP_C)$. The CLMP consists of three significant entities and they are $CLBP_{Sign}(CLBP_S)$, $CLBP_{Magnitude}(CLBP_M)$ and $CLBP_{Centre}$ ($CLBP_C$).

$CLBP_S$ denotes the sign of the pixel value, which is computed by subtracting the central and the local pixel value. The sign can either be positive or negative. $CLBP_M$ indicates the difference of central and the local pixel value with respect to magnitude. Similarly, $CLBP_C$ is the subtracted value of local pixel value from the average central pixel. $CLBP_S$ is similar to the standard LBP and is given by,

$$CLBP_{S_{P,R}} = \sum_{p=0}^{p=1} s\left(g_p - g_c\right) 2^p \qquad (8)$$

where g_c is the grey level of the central pixel and g_p is the value of the neighbourhood pixels. P is the count of neighbourhood pixels and R is the radius of the neighbourhood window. The sign of the pixel value is determined by the following condition.

$$s\left(a\right) = \begin{cases} 1 & \text{if } a \geq 0 \\ 0 & \text{if } a < 0 \end{cases} \qquad (9)$$

$$CLBP_{M_{P,R}} = \sum_{p=0}^{p=1} t\left(m_p, c\right) 2^p \qquad (10)$$

$$t\left(x, c\right) = \begin{cases} 1 & \text{if } x \geq c \\ 0 & \text{if } x < c \end{cases} \qquad (11)$$

where, c is the threshold, which is a dynamic value and m_p is the magnitude. $CLBP_C$ is computed using Eq.(12).

$$CLBP_{C_{P,R}} = t\left(g_c, c_l\right) \qquad (12)$$

where, c_l is the average grey level of the entire image. All these three entities are clubbed either by combining all the three or by combining two entities followed by the other.

3.4.2 Contourlet:

The success of the feature extraction step depends on how the images are represented. Some of the noteworthy expectations for successful image representation are multiresolution, localization, directionality and anisotropy [22]. Multiresolution is the capability that allows the digital images to be represented in all

resolutions (course to fine). The image components have to be localized with respect to both spatial and frequency domain. The image representation technique must support multiple directional orientations and should not be limited. The image representation technique has to be able to trap the contours being present in the image, by means of different shapes with aspect ratio. From the above requirements, multiresolution and localization are supported by wavelets. However, all the requirements are required for successful image representation techniques.

Curvelets somehow satisfies the above mentioned requirements, however curvelets find it easier to represent continuous images and could not attain the required rate of image representation for discrete images. This issue is addressed by contourlets, which is a multidirectional multiresolution transform that functions exactly over the discrete domain. Contourlet transforms are fixed by nature and thus it is adaptable to any kind of image processing tasks. Though there are several transforms that provide multidirectional and multiresolution capabilities, these transforms do not allow several directions at a single scale. The contourlet do provide iterated filter banks and hence it is far efficient than the analogous transforms.

Contourlet combines Laplacian Pyramid (LP) and Directional Filter Bank (DFB) together. The DFB grabs the high frequency components, such that the low frequency components are eliminated before they are treated by DFB. Consider an input image $x_0[n]$ and is passed over the LP. The outcome of this step is the K bandpass images and it is represented as follows:

$$y_k\left[n\right]; k = \left(1, 2, ..., K\right) \text{ and } x_k\left[n\right] \qquad (13)$$

where $y_k\left[n\right]; k = \left(1, 2, ..., K\right)$ denotes the image in fine to coarse order and $x_k[n]$ is the lowpass image. This means that the kth level of LP subdivides the input image $x_{k-1}[n]$ into a coarse and a fine image, which are represented by $x_k[n]$ and $y_k[n]$, respectively. Every single bandpass image $y_k[n]$, is again subdivided to the degree of d_k to 2^{d_k} bandpass directional images and are represented by,

$$a_{k,l}^{(d_k)}\left[n\right]; k = \left(1, 2, ..., 2^{d_k} - 1\right) \qquad (14)$$

Thus, the contourlet is applied over the inputs for extracting the features.

3.4.3 Feature Extraction over Mammograms:

The main objective of this work is to attain maximum classification accuracy, such that the normal and abnormal mammograms can easily be distinguished. Besides this, it is important for any CAD system to show up least false positive and false negative rates. Taking the aforementioned statement as a challenge, this work has considerably reduced the false positive and false negative rates by framing the feature vector, which is the result of combining CLBP and contourlet.

To start up with feature extraction, this work extracts features from the inputs by CLBP. The values of P and R are varied and the texture features are computed. Based on the values of P and R, the feature count varies, which is because of the change in the window size. This is followed by the application of two level contourlet, from which the GLCM features are extracted from the sub-bands. The contourlet is applied over the segmented image through DFB in three scales and the coefficients are extracted. In

this work GLCM is employed, as it simply establishes the spatial relationship between pixels with respect to texture.

GLCM is a square matrix with size $GL \times GL$ and GL is the count of grey intensities in an image. Consider an element (a,b) of an image which is in the position (x,y) has the grey level intensity I_i. This indirectly notifies that the pixel at location $(x+dx,y+dy)$ has the grey level intensity I_j. The dx and dy represent different scale and orientation. The contourlet features are obtained by weighing the values of each and every co-occurrence matrix. Finally, these values are added together to form the feature set. The GLCM features being considered by this work are energy, contrast, correlation, autocorrelation, homogeneity, dissimilarity and inertia.

As the input image is applied with two level contourlet, the first level contourlet provides 2 subbands and the second level contourlet produces 4 subbands. The first level presents a matrix which contains 4×120 elements and the second level produces a matrix with 16×120 elements. These two matrices are combined together to form a resultant matrix with 20×120 elements. The co-occurrence of the three matrices with a single pixel as offset and the angles being considered are 0, 45 and 90 degrees. Hence for a single input image, three co-occurrence matrices are formed with the dimension of 8×8. The feature set being computed by this work contains 21 components, as 7 GLCM features alone are considered out of 14 properties, for 3 matrices. However, the feature matrix being obtained from CLBP varies with respect to the values of P and R. The classification process is initiated, as soon as the feature vectors are formed.

3.5 CLASSIFICATION WITH MULTICLASS ELM

ELM is proven to be the fastest and accurate classifier [23]. Initially the classifier is trained with training samples, in order to impart knowledge to it. The gained knowledge is applied over the test sample image and classification is done. Consider A training samples, which is denoted as (u_j, v_j), where $u_j = [u_{j1}, u_{j2}, \ldots, u_{jn}]^T \in D^n$ and u_j indicates the jth training sample with dimension n. $v_j = [v_{j1}, v_{j2}, \ldots, v_{jm}]^T \in D_m$ denotes the jth training label with dimension m, where m is the count of classes which are computed in the previous section. A Single hidden Layer Feed-Forward Neural Network (SLFN) is defined with a single activation function $a(x)$ and G neurons, is represented as follows,

$$\sum_{i=1}^{G} \gamma_i q\left(w_i \cdot u_j + p_i\right) = r_j; j = 1, 2, \ldots, n . \quad (15)$$

w_i is the weights represented as vectors $w_i = [w_{i1}, w_{i2}, \ldots, w_{in}]^T$, which is responsible for connecting the ith hidden neuron with the input neurons, where $i = [i_1, i_2, \ldots, i_m]^T$. This vector with weights interconnect the ith hidden neuron with the output neurons and the bias of the ith hidden neuron is given by bs_i. ELM doesn't require any prior knowledge of the input data, such that the w_i and bs_i can be allotted in a random fashion. Now, the SLFN can be written as

$$\sum_{i=1}^{G} \gamma_i a\left(w_i \cdot u_j + bs_i\right) = v_j; j = 1, 2, \ldots, n. \quad (16)$$

Let HL be the ELM's hidden layer output matrix, such that the ith column of HL denotes the ith hidden neurons output vector by considering the inputs $u_{j1}, u_{j2}, \ldots, u_{jn}$

$$HL = \begin{bmatrix} a\left(w_1 \cdot u_1 + bs_i\right) & \cdots & a\left(w_G \cdot u_1 + bs_G\right) \\ \vdots & \ddots & \vdots \\ a\left(w_1 \cdot u_n + bs_i\right) & \cdots & a\left(w_G \cdot u_n + bs_G\right) \end{bmatrix} \quad (17)$$

$$\gamma = \begin{bmatrix} \gamma_1^T \\ \vdots \\ \gamma_G^T \end{bmatrix} \quad (18)$$

$$V = \begin{bmatrix} v_1^T \\ \vdots \\ v_n^T \end{bmatrix} \quad (19)$$

The matrix form is represented as

$$HL\gamma = V \quad (20)$$

The weights of the outputs are computed by calculating the norm least-square solution, and the Eq.(20) is given as follows

$$\gamma = HL^\dagger V \quad (21)$$

where, $HL^\dagger$ is the HL's Moore-Penrose generalized inverse. The ELM training phase is carried out by passing the class count m, the activation function $a(x)$, count of hidden neurons G and ELM count in ensemble E. In the training phase, the ELM is passed with the training set $TS = \{(u_j, v_j) | u_j \in D^n, v_j \in D^m; j = 1, 2, \ldots, N\}$. The ELM is trained by computing γ for all TS using Eq.(21).

Consider, $TS' = \left\{u_j \middle| u_j \in D^n; j = 1, 2, \ldots, N'\right\}$ as the test set. Initially, the output matrix $v_{test}(k) = HL_{test}(k) \times \gamma_k$ with dimensionality $N' \times m$. All the output matrices are summed up and the maximal value in the row is marked. In this work, the value of k, which is the ensemble is fixed as 15. This is because, the optimal results are obtained when $k = 15$. The classification task is achieved by training the classifier with the feature vector being framed in the feature extraction phase. In the testing phase, as the input image is passed all the preliminary processes are done followed by the feature extraction process. The feature vector being computed is compared with the feature vector of the trained set and the result is returned.

4. EXPERIMENTAL ANALYSIS

The performance of the proposed work is analysed and the experimental results are presented in this section. This work employs MIAS dataset [24] for evaluating the performance of the proposed approach. The MIAS dataset consists of 322 images of size 1024×1024. The performance of the proposed work is evaluated by varying the classifiers and feature extraction techniques in terms of accuracy, sensitivity and specificity. In order to check the efficiency of the classifier, we compare the ELM with SVM and k-NN. The sample segmentation results of the proposed approach are presented in Fig.4.

(a)   (b)   (c)   (d)   (e)
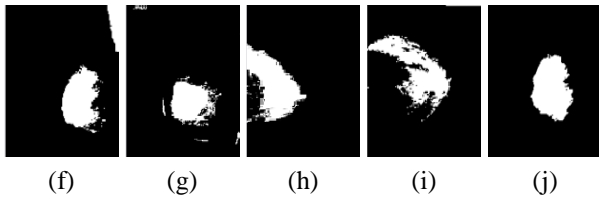
(f)   (g)   (h)   (i)   (j)
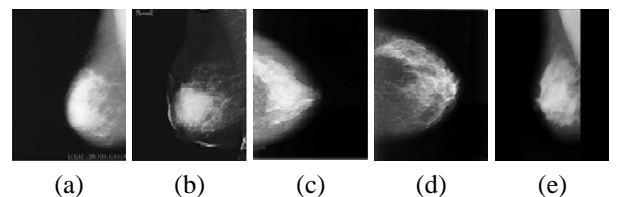
Fig.4(a)-(e). Original images, (f)-(j) are segmented images

The accuracy rate is the most important factor of a CAD system. Sensitivity rate is the rate of correctly classified images to the sum of images that are correctly and incorrectly classified. Specificity is the rate of images that are accurately classified as negative to the sum of images that are correctly classified as negative and the images that are wrongly classified as positive. The formulae for computing accuracy, sensitivity and specificity rates are presented below.

$$acc_r = \frac{T_p + T_n}{T_p + T_n + F_p + F_n} \times 100 \qquad (22)$$

$$Sen_r = \frac{T_p}{T_p + F_n} \times 100 \qquad (23)$$

$$Sp_r = \frac{T_n}{F_p + T_n} \times 100 \qquad (24)$$

where, acc_r is the accuracy rate, Sen_r is the sensitivity rate and Sp_r is the specificity rate. T_p, T_n, F_p, F_n are the True Positive, True Negative, False Positive and False Negative rates. Initially, the performance of ELM is compared with SVM and k-NN classifier in terms of accuracy, sensitivity and specificity. This is followed by the performance comparison of feature extractors with respect to accuracy, sensitivity and specificity. Finally, the P and R values of CLBP are varied to check the performance and the results are presented. The experimental results are presented as follows.

Initially, the performance of the proposed approach is evaluated by varying the classifier. The classifiers being utilized to check the performance are k-NN, SVM and proposed ELM. On analysis, it is observed that the performance of k-NN is the worst, when compared to SVM and ELM. The reason is that the k-NN classifier requires the pre-defined value for the parameter 'k'. As the classification is achieved by taking distance measures into account, it is necessary to study the performance of different distance measures such as Euclidean, Minkowski, Mahalanobis, Manhattan and so on. Additionally, the k-NN classifier consumes more time to classify between the entities. K-NN classifier suffers from time and computational complexity. Above all, the performance of k-NN solely depends on the distance measure being utilized.

As this proposed work employs ELM for classification, serious research on choosing distance measures has not been done. The distance measure used for k-NN classification is Euclidean distance. However, the k-NN classifier may prove better results for different distance measure. As this work focuses on ELM, the k-NN classifier is implemented with the standard Euclidean distance. The accuracy, sensitivity and specificity rates shown by k-NN classifier are 87.5, 87 and 91 percent respectively.

The results produced by SVM are far better than k-NN, Relevance Vector Machine (RVM) [26] but poorer than ELM.

RVM is proposed to address the issues faced by SVM. However, RVM suffers from the same pitfall as that of SVM. Though SVM is better than k-NN, it suffers from several drawbacks. SVM is slower when compared to ELM. The performance of SVM depends on the choice of kernel. The accuracy, sensitivity and specificity rates being shown by SVM are 92, 90 and 94 respectively. This work employs ELM for classification, owing to its simplicity, faster learning capability and the requirement of minimal human interference [25]. The performance of the ELM outperforms the comparative techniques by showing better accuracy, sensitivity and specificity rates.
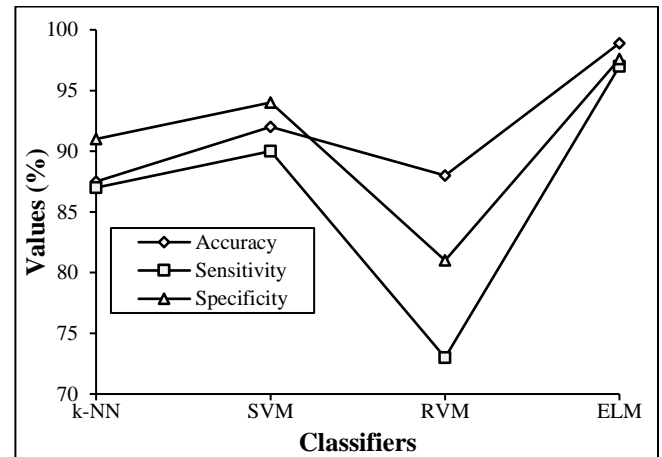

Fig.5. Comparative analysis w.r.t classifier

The classifiers are varied in order to check the performance and ELM presents remarkable performance, which is evident in the graphical representation. ELM outperforms k-NN and SVM with maximum accuracy, sensitivity and specificity rates. The experimental analysis with respect to feature extractors is presented in Fig.6. The performance of this work is again analysed by varying the feature extraction techniques. Actually, this work clubs both CLBP and contourlet together. However, it is quite interesting to check the performance of CLBP and contourlet when they are implemented alone. Initially, CLBP alone is utilized for extracting the features. The performance of CLBP is not satisfactory, as it shows only 92.2 as accuracy rate. The performance of CLBP depends on the total count of processing pixels and the size of the window. Though CLBP is an effective texture feature descriptor, it cannot do wonders alone. Secondly, contourlet alone is utilized for feature extraction process. Contourlet provides multidimensional and multidirectional capability, however the results are improved only when the CLBP is clubbed with the contourlet. The implementation of contourlet alone shows an accuracy rate of 98.2. Though the accuracy rate shown by the combination of CLBP and contourlet is 98.8, which is slightly greater than the accuracy rate proved by contourlet, the specificity rate shows much difference. The specificity rate of contourlet is 92.4, whereas the specificity rate of contourlet + CLBP is 94.3.

The Fig.6 illustrates the effectiveness of combining CLBP and contourlet. In order to prove the potentiality, initially the features are extracted with CLBP alone which is followed by contourlet alone. Finally, both these feature extractors are combined together and the performance is checked. The performance of CLBP + contourlet outperforms the CLBP and contourlet, when

implemented individually. Finally, the values of *P* and *R* are changed to check the performance and the results are presented in the following Fig.7. The main objective of this analysis is to show the optimal values of *P* and *R*. The performance of CLBP relies on the values of *P* and *R*. Thus, choosing optimal values of *P* and *R* are absolutely necessary. The values of *P* and *R* directly impacts over the accuracy, sensitivity and specificity rates. Hence, the better the optimality of *P* and *R*, the better is the outcome of the CLBP. It is noticed that the accuracy, sensitivity and specificity rates of the CLBP is better, when the *P* and *R* values are 24 and 3 respectively.
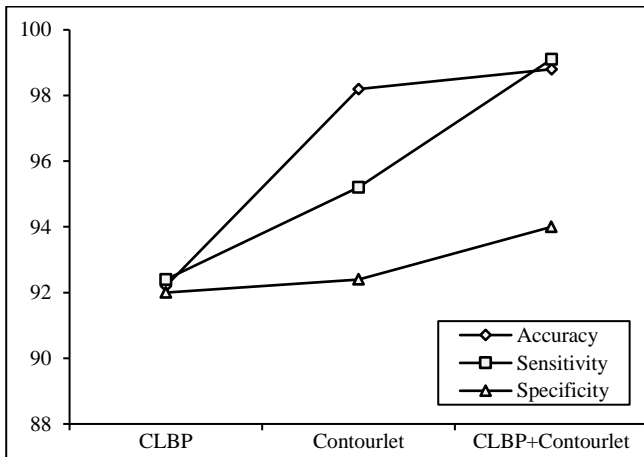

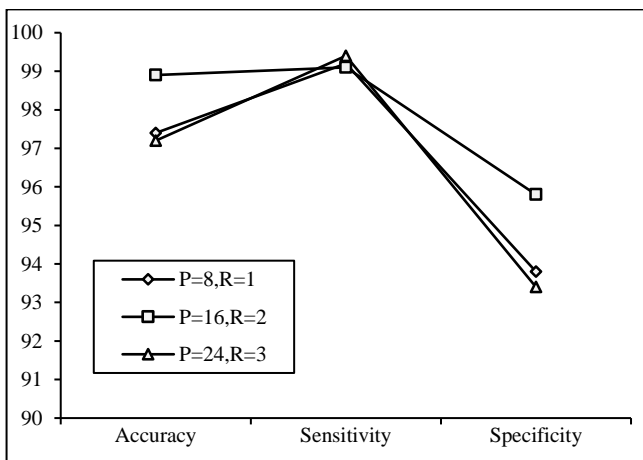Fig.6. Comparative analysis w.r.t feature extractors

Fig.7. Impact of *P* and *R* on the CLBP + Contourlet

The value of *P* and *R* are changed, so as to study the performance. From the experimental results, it is evident that the optimal value for *P* and *R* with respect to this dataset is 16 and 2 respectively. When *P*, *R* are fixed as 24 and 3, the performance starts to degrade. Thus, fixing the optimal values for *P* and *R* are necessary to attain best results. Thus, the reliable and efficient CAD system to detect microcalcification and mammogram classification is presented.

5. CONCLUSION

This paper presents a robust and efficient CAD system for mammogram classification, which can classify between the normal, malignant and benign cancer. The research goal is achieved by decomposing the work into four significant phases, which are mammogram pre-processing, segmentation, feature extraction and classification. The mammograms are pre-processed by adaptive median filter and segmented by GHFCM, which is proven to be robust and efficient. The segmented images are then treated with the feature extraction process, which is done by the combination of contourlet and CLBP to frame the feature set. Finally, ELM is employed to distinguish between the mammograms as normal, benign or malignant. The performance of the proposed approach is checked by varying the classifier such as k-NN, SVM, RVM, ELM and feature extractors. It is concluded that the performance of the combination of CLBP and contourlet works well with ELM classifier. The results shown by combination of CLBP and contourlet along with ELM are satisfactory in terms of accuracy, sensitivity and specificity rates. In future, the P and R values of CLBP are planned to be selected by introducing an optimal meta-heuristic algorithm.

REFERENCES

[1] Over 17 Lakh New Cancer Cases in India by 2020: ICMR, Available at: http://icmr.nic.in/icmrsql/archive/2016/7.pdf.

[2] S.M. Astley and F.J. Gilbert, "Computer-Aided Detection in Mammography", *Clinical Radiology*, Vol. 59, No. 5, pp. 390-399, 2004.

[3] Daniel C. Moura and Miguel A. Guevara Lopez, "An Evaluation of Image Descriptors combined with Clinical Data for Breast Cancer Diagnosis", *International Journal of Computer Assisted Radiology and Surgery*, Vol. 8, No. 4, pp. 561-574, 2013.

[4] Suman Shrestha, "Image Denoising using New Adaptive based Median Filter", *Signal and Image Processing : An International Journal*, Vol. 5, No. 4, pp. 1-13, 2014.

[5] Yuhui Zheng, Byeungwoo Jeon, Danhua Xu, Q.M. Jonathan Wu and Hui Zhang, "Image Segmentation by Generalized Hierarchical Fuzzy C-Means Algorithm", *Journal of Intelligent and Fuzzy Systems*, Vol. 28, No. 2, pp. 961-973, 2015.

[6] M. Mirmehdi, X. Xie and J. Suri, "*Handbook of Texture Analysis*", Imperial College Press, 2009.

[7] Mohamed Meselhy Eltoukhy, Ibrahima Faye and Brahim Belhaouari Samir, "A Statistical based Feature Extraction Method for Breast Cancer Diagnosis in Digital Mammogram using Multiresolution Representation", *Computers in Biology and Medicine*, Vol. 42, No. 1, pp. 123-128, 2012.

[8] I. Christoyianni, A. Koutras, E. Dermatas and G. Kokkinakis, "Computer Aided Diagnosis of Breast Cancer in Digitized Mammograms", *Computerized Medical Imaging and Graphics*, Vol. 26, No. 5, pp. 309-319, 2002.

[9] A. Karahaliou, S. Skiadopoulos, I. Boniatis, P. Sakellaropoulos, E. Likaki, G. Panayiotakis and L Costaridou, "Texture Analysis of Tissue Surrounding Microcalcifications on Mammograms for Breast Cancer Diagnosis", *The British Journal of Radiology*, Vol. 80, No. 956, pp. 648-656, 2007.

[10] S.J.S. Gardezi, I. Faye and M.M. Eltoukhy, "Analysis of Mammogram Images based on Texture Features of Curvelet Sub-Bands", *Proceedings of 5th International Conference on Graphic and Image Processing*, pp. 906924-906926, 2014.

[11] A. Oliver, X. Llado, R.Marti, J. Freixenet and R. Zwiggelaar, "Classifying Mammograms using Texture Information", *Proceedings of International Conference on Medical Image Understanding and Analysis*, pp. 223-227, 2007.

[12] A. Oliver, X. Llad, J. Freixenet and J. Mart, "False Positive Reduction in Mammographic Mass Detection using Local Binary Patterns", *Proceedings of International Conference on Medical Image Computing and Computer-Assisted Intervention*, pp. 286-293, 2007.

[13] Sophie Paquerault, Nicholas Petrick, Heang-Ping Chan, Berkman Sahiner and Mark A. Helvie, "Improvement of Computerized Mass Detection on Mammograms: Fusion of Two-View Information", *Medical Physics*, Vol. 29, No. 2, pp. 238-247, 2002.

[14] Y.A.S. Duarte, M.Z. Nascimento and D.L.L. Oliveira, "Classification of Mammographic Lesion based in Completed Local Binary Pattern and using Multiresolution Representation", *Journal of Physics: Conference Series*, Vol. 490, pp. 1-5, 2013.

[15] M.M. Eltoukhy, I. Faye, and B.B. Samir, "Curvelet based Feature Extraction Method for Breast Cancer Diagnosis in Digital Mammogram", *Proceedings of International Conference on Intelligent and Advanced Systems*, pp. 1-5, 2010.

[16] Jae Young Choi and Yong Man Ro, "Multiresolution Local Binary Pattern Texture Analysis combined with Variable Selection for Application to False-Positive Reduction in Computer-Aided Detection of Breast Masses on Mammograms", *Physics in Medicine and Biology*, Vol. 57, No. 21, pp. 7029-7034, 2012.

[17] S. Chen and D. Zhang, "Robust Image Segmentation using FCM with Spatial Constraints based on New Kernel-Induced Distance Measure", *IEEE Transactions on Systems, Man and Cybernetics*, Vol. 34, No. 4, pp. 1907-1916, 2004.

[18] D. Pham and J.L. Prince, "An Adaptive Fuzzy C-Means Algorithm for Image Segmentation in the Presence of Intensity Inhomogeneities", *Pattern Recognition Letters*, Vol. 20, No. 1, pp. 57-68, 1999.

[19] A. Pedrycz and M. Reformat, "Hierarchical FCM in a Stepwise Discovery of Structure in Data", *Soft Computing*, Vol. 10, No. 3, pp. 244-256, 2006.

[20] Nicolaos B. Karayiannis, "Generalized Fuzzy C-Means Algorithms", *Proceedings of the 5th IEEE International Conference on Fuzzy Systems*, Vol. 2, pp. 1036-1042, 1996.

[21] Zhenhua Guo, Lei Zhang and David Zhang, "A Completed Modeling of Local Binary Pattern Operator for Texture Classification", *IEEE Transactions on Image Processing*, Vol. 19, No. 6, pp. 1657-1663, 2010.

[22] Minh N. Do and Martin Vetterli, "The Contourlet Transform: An Efficient Directional Multiresolution Image Representation", *IEEE Transactions on Image Processing*, Vol. 14, No. 12, pp. 2091-2106, 2005.

[23] Guang-Bin Huang, Hongming Zhou, Xiaojian Ding and Rui Zhang, "Extreme Learning Machine for Regression and Multiclass Classification", *IEEE Transactions on Systems, Man and Cybernetics-Part B*, Vol. 42, No. 2, pp. 513-529, 2012.

[24] The new MIAS Database. Available: http://www.wiau.man.ac.uk/services/MIAS/MIASfaq.html.

[25] Erik Cambria and Guang-Bin Huang, "Extreme Learning Machines", *IEEE Intelligent Systems*, pp. 30-59, 2013.

[26] L. Wei, Y. Yang, R.M. Nishikawa, M.N. Wernick and A. Edwards, "Relevance Vector Machine for Automatic Detection of Clustered Microcalcifications", *IEEE Transactions on Medical Imaging*, Vol. 24, No. 10, pp. 1278-1285, 2005.