An Effective Data Mining Technique For Classification Of Mammograms To Diagnose Breast Cancer

Manmohan Sahoo
Kmbb college of engg
Bhubaneswar,Odisha

Suryamani Biswal
Regional College of
Management, Bhubaneswar
Odisha

Amalendu Bag
Regional College of
Management,Bhubaneswar,
Odisha

Aswini kumar mohanty
The Techno School
Bhubaneswar,Odisha

Abstract:-Breast cancer is the leading cause of cancer death among women. Screening x-ray mammography is the first and regular method currently available for the reliable early detection and treatment can be started for hopeful cure of breast cancer. Research indicates that the mortality rate could decrease by 35% if women age around 45-50years and older have regular mammograms. The detection rate can be confirmed 15-30% by providing the radiologist with results from a computer-aided diagnosis (CAD) system acting as a second opinion. However, among screening mammograms routinely interpreted by radiologists, very few (approximately 10%) cases actually have breast cancer. It would be beneficial if an accurate CAD system existed to identify normal mammograms and thus allowing the radiologist to focus on suspicious cases. This strategy could reduce the radiologist's workload and improve screening performance.

Since mammography is considered as the most effective means for breast cancer diagnosis, this paper introduces two separate techniques for mass and micro-calcification segmentation in digital mammograms. There are two different segmentation techniques used for masses and microcalcifications. For masses there are three steps- background subtraction, fuzzy texture representation and entropic theresholding. Similarly micro-calcifications are also segmented in three stages – background subtraction, Laplacian of Gaussian filtering and contrast estimation followed by thresholding. GLCM texture feature extractions are known to be the most common and powerful techniques for texture analysis including histogram statistical features. Experiments have been taken for a data set of 322 images taken from MIAS of different types with the aim of improving the accuracy of Association Rule classifier by generating minimum no. of rules to cover more patterns. In this study, an incremental updating technique is applied to associative classification for constructing classification system when a new training dataset is appended to an old training dataset. The proposed algorithm, called Incremental Classification Based on Association Rules (ICBA). ICBA has 2 phases which are rule generator phase (ICBA-RG) and classifier building phase (ICBA-CB). In order to reduce the execution time, we applied the concept of Fast Update algorithm (FUP) algorithm to both phases of our algorithm. The experiment results show that the proposed algorithm has execution time better than CBA algorithm.

Keywords: Breast cancer, fuzzy texture, entropy, fuzzy segmentation, Entropic thresholding, microcalcification, Laplacian of Gaussian filter. GLCM texture features, Association rule mining

1. INTRODUCTION

Breast cancer is one of the most dangerous diseases prevalent among the women. It is estimated that 182 000 new cases are being reported every year and 46 000 die of this cancer [1]. Mammography is best suited for early diagnosis of breast cancer. Computer aided diagnosis helps as a second opinion for the radiologist to decide if the cancer is present. Masses and micro-calcifications are the two kinds of abnormalities present in the mammogram. Much of the work on computer aided diagnosis is done on micro-calcifications. This is because microcalcifications show early sign of breast cancer.

Micro-calcifications are high contrast rods or spots present on the mammogram. Only the regions where these micro-calcifications appear as clusters within a radius of 1cm are considered suspicious. The morphologic appearance of micro-calcifications has been emphasized in the last decades.

Mass segmentation plays an important role in this diagnosis. One of the approaches for segmentation is thresholding. Effective thresholding techniques that lead to the extraction of masses are proposed [2,3]. Some of these techniques are entropy based [4]-[5]. Templates are used to code the geometric structures of specific types of masses and then matched those template patterns from a mammogram. The step involved in classification of mammogram is depicted in figure 1.

Texture analysis has been used by some researchers to separate out the differently textured regions [6-10]. Laws developed a pixel based method of texture analysis particularly applicable to radiographic images [11]. The Laws' method uses filter masks to extract secondary features from natural micro-structure characteristics of the image (level, edge, spot and ripple) which can then be used for segmentation or classification. Other methods include thresholding [12], edge detection and gradient-based [13]. Region growing [14] is used for segmentation of suspicious portions from the mammograms. If pixel intensity falls in a specific range, the region-growing algorithm is applied and the intensity gradient is computed to test whether the candidate pixel satisfies the mean and variance criteria. A problem with this algorithm is that it requires many user-input variables. These variables are actually image dependent and should be determined automatically. Some of the hierarchical schemes have been extended to the case of multiple features and applied to the most difficult problem of segmentation based on texture. A natural generalization of applying thresholds to a single gray level characteristic gives rise to segmentation based on clustering. Several edges oriented methods are introduced and these generally attempt to locate texture edges based on the computation of a multi-feature gradient-like operator. Hierarchical approaches using pyramid node linking [15] or applying the split-and-merge algorithm to the co occurrence matrix [16] also exist in the literature.

The rest of the paper is organized as follows. In section 2 we propose segmentation schemes for masses as well as micro-calcifications. The region of interest containing mass is initially the background subtracted and we apply fuzzy texture representation method for capturing the texture information from the neighborhood. Then Entropic thresholding is used for segmentation. The region of interest containing the micro-calcifications is background subtracted and Laplacian of Gaussian filter is applied. This is followed by local contrast estimation and thresholding. Section 3 have presents the result of segmentation followed by conclusion and future work in section 4.
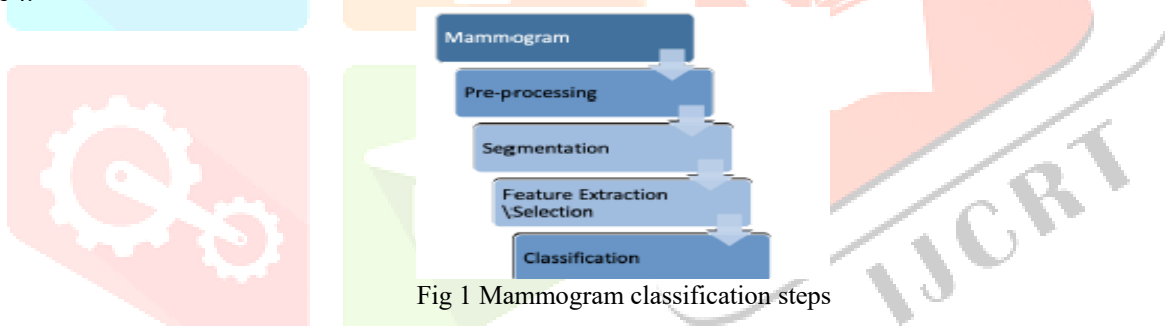

Fig 1 Mammogram classification steps

2. SEGMENTATION OF MASSES AND MICROCALCIFICATIONS

Two separate methods are proposed to segment the masses and micro-calcifications in mammograms. The digital mammography images serve as the input to the proposed algorithm and no other information is required.

2.1. Segmentation of Micro-calcification

The method suggested for the detection of microcalcifications and suspicious masses from mammogram images is tested over several images taken from mini-MIAS (Mammogram Image Analysis Society, UK) database [17]

Local Background Subtraction

At first the mammogram is considered as a three dimensional plot with the third axis (z) corresponding to the intensity of each pixel. The whole image is split up into 30 x 30 sub-regions. We calculate one value for each sub-region since the microcalcifications fall in the top 30% gray level range. So we consider the mean of the gray levels of the pixels that are in the other part of the range as the value for the sub-region. We use bi-cubic interpolation to obtain the second plot representing the intensity level of the local background. The interpolated image is subtracted from the original mammogram producing a third image with each pixel value providing the difference between the original and local background pixel values. The image is normalized to range 0-255 and subjected to filtering.

Filtering using Laplacian of Gaussian Filter

The Laplacian is a 2-D isotropic measure of the second spatial derivative of an image. The Laplacian of an image highlights regions of rapid intensity change and is therefore often used for edge detection. The Laplacian is often applied to an image that has first been smoothed with an approximate Gaussian smoothing filter in order to reduce its sensitivity to noise. The Laplacian operator normally takes a single gray level image as the input and produces another gray level image as the output. The Laplacian $L(x, y)$ of an image with pixel intensity values $I(x, y)$ is given by:

$$L(x,y) = \frac{\partial^2 I}{\partial x^2} + \frac{\partial^2 I}{\partial y^2} \qquad (1)$$

Using an appropriate convolution kernel, the Laplacian can be evaluated using the standard convolution methods. The kernel approximates a second derivative measurement on the image, which is very sensitive to noise. To counter this, the image is often Gaussian smoothed before applying the Laplacian filter. This pre-processing step reduces the high frequency noise components prior to the differentiation step [29].Since the convolution operation is associative, we convolve the Gaussian smoothing filter with the Laplacian filter first, and then convolve this hybrid filter with the image to achieve the required result. The 2-D LoG function centered on zero with the Gaussian standard deviation σ shown in Fig. 2 has the form:

$$LoG(x,y) = \frac{1}{\pi\sigma^4}\left(1 - \frac{(x^2 + y^2)}{2\sigma^2}\right)e^{\frac{(x^2+y^2)}{2\sigma^2}} \qquad (2)$$

The discrete version of the filter suitable to be applied on digital images is of the form:

$$LoG_\sigma(i,j) = \frac{(i^2 + j^2 - 2\sigma^2)h_g(i,j)}{2\Pi\sigma^6 \sum_i\sum_j h_g} \qquad (3)$$

A discrete kernel that approximates the Gaussian function with σ = 1.7 having a size of 51 x 51 pixels is chosen in our case
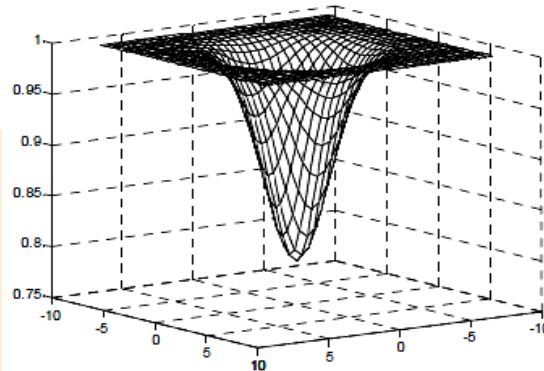
Fig. 2 The 2-D Laplacian of Gaussian (LoG) function.
The x and y axes are marked in terms of standard deviations (σ).

In the proposed approach the input signal is fed to LoG filter set to the standard deviation of 1.7

$$G = LoG_{\sigma=1.7} * I \qquad (4)$$

where, G is the response of the filtering. The effect of this operation is the increase in the contrast of the micro-calcifications.

Local Contrast Estimation and Thresholding
The local contrast measure C of each pixel is estimated and compared with a threshold value CT and if C of the pixel is found greater than CT then the pixel of interest is labeled as microcalcification pixel.
The local contrast is defined as the difference between the maximum value of the window surrounding the pixel and the pixel of interest. The value of CT used is 55. The resulting image is a binary image showing the white regions as the micro calcifications.

2.2 Segmentation of Mass

The literature is ripe with so many approaches for characterizing the texture. Most of the approaches are complex and time consuming. There is hardly any attempt to characterize texture by fuzzy logic. As the texture is a region concept, it is hoped that suitable local information will represent the underlying texture. By incorporating neighborhood information in the fuzzification function, it is possible to represent texture by means of fuzzy logic techniques. The ambiguities in texture arising due to fuzzy nature of the image function are a motivating factor to devise fuzzy texture technique. Since texture is region based, we consider arrangement of image functions (i.e., intensities) of pixels in a local region, say, a window, in order to characterize the texture using the sigmoid type membership function.

Background Subtraction
Since the background in the Region of Interest (ROI) is non-uniform, background subtraction technique in [25] is adopted here. This technique estimates the low-frequency background by using the grey level of a band of pixels around the perimeter of the ROI. This method is based on visual comparison of the background-corrected images with the original image so that the background is leveled and no artifact is present. Running average of the pixel values along the perimeter of the ROI is calculated using a box filter of a 32 x 16 kernel, of which the long dimension is parallel to the edge of the ROI. For the perimeter pixels that are within 16 pixels of one of the four corners of the ROI, the long dimension of the box filter kernel is reduced on the side that is limited by the ROI edge. For example, the average pixel value at a corner of the ROI is obtained by a 16 x 16 box filter, with one apex of the box filter kernel coinciding with the corner pixel. The gray level, G(i,j), of a given pixel (i, j) in the estimated background image of the ROI is then calculated as

$$G(i,j) = \left[\sum_1^4 g_k/d_k\right] \Big/ \left[\sum_1^4 1/d_k\right] \qquad (5)$$

where, gk is the gray level of the pixel at the intersection between one side of the lowpass-filtered ROI perimeter and the normal from the pixel (i, j) to that side, and dk is the distance from the pixel (i, j) to the intersection; k ranges from 1 to 4, denoting the four sides of the ROI. More than four pixels around the ROI perimeter might be used in the weighted sum for the background estimation and a low-pass filter might be applied to the interpolated background image to provide further smoothing. We use (5) for the interpolation and apply a low-pass box filter of a 32 x 32 kernel to the interpolated background image in this study. The background image is then subtracted from the original ROI, thus reducing the background to near zero.

Fuzzy texture representation

To convert the spatial domain image into the fuzzy domain, we consider the spatial arrangement of gray levels of pixels over a window. The fuzzy property can be expressed in terms of a membership function. A membership function to this effect is defined by the sigmoid function.

$$\mu_{(kl)}(i,j) = \frac{1}{1+\exp\left[-\{(x(i,j)-x(k,l))/\tau\}^2\right]} \qquad (6)$$

where, x (i, j) is the gray level of the pixel at the (i,j)th position and τ is the fuzzifier which is taken as the window size. In our experiment, we take the value of τ as 5. We note that

$$\mu_{(i,j)}(k,l) = 0.5 \qquad if \; x(i,j) = x(k,l) \qquad (7)$$

To consider the response from the neighboring pixels, we obtain the cumulative response of (i, j)th pixel as follows

$$y(i,j) = \frac{\sum_{k,l} \mu_{(k,l)}(i,j)*x(k,l)}{\sum_{k,l} \mu_{(k,l)}(i,j)} \qquad (8)$$

This is the defuzzified response of the (i,j)th pixel over the window of size 5. This process is repeated for all pixels in the image and this yields a texture image consisting of all defuzzified values. For convenience of notation, we designate the matrix formed by y(i,j) as the response matrix. The image y is subjected to application of entropic thresholding.

Mass segmentation by entropic thresholding

The defuzzified response which is subjected to entropic thresholding is used for mass segmentation to extract the suspicious tumor from the background. The procedure for calculation of the threshold for the image is as follows. Assume that the gray level range is G= {0, 1,..., L-1}. Also let t be the threshold partitioning the gray levels G into the two regions G0={0,1,2,...,t} and G1={t,t+1,...,L-1}. The co-occurrence matrix W thresholded by t can be divided into four quadrants which represent ackground to background (BB), background to foreground (BF), foreground to background (FB) and foreground toforeground (FF) respectively provided that pixels with gray levels above t are assumed to be foreground pixels and otherwise are background pixels. Assume that the probabilities of the four quadrants are pBB(i,j), pBF(i,j), pFB(i,j) and pFF(i,j). The probabilities of the four quadrants are obtained by the following formulae:

$$P_{BB}(i,j) = \frac{T(i,j)}{\sum_{i=1}^{T}\sum_{j=1}^{T}T(i,j)} \qquad (9)$$

$$P_{BF}(i,j) = \frac{T(i,j)}{\sum_{i=1}^{T}\sum_{j=T+1}^{L}T(i,j)} \qquad (10)$$

$$P_{FB}(i,j) = \frac{T(i,j)}{\sum_{i=T+1}^{L}\sum_{j=1}^{T}T(i,j)} \qquad (11)$$

$$P_{FF}(i,j) = \frac{T(i,j)}{\sum_{i=T+1}^{L}\sum_{j=T+1}^{L}T(i,j)} \qquad (12)$$

Then, the entropies associated with the four quadrants are computed from:

$$H_{BB}(t) = -\sum_{i=0}^{T}\sum_{j=0}^{T}p_{BF}(i,j)\log p_{BB}(i,j) \qquad (13)$$

$$H_{BF}(t) = -\sum_{i=1}^{T}\sum_{j=T+1}^{L}p_{BF}(i,j)\log p_{BF}(i,j) \qquad (14)$$

$$H_{FB}(t) = -\sum_{i=T+1}^{L}\sum_{j=1}^{T}p_{FB}(i,j)\log p_{FB}(i,j) \qquad (15)$$

$$H_{FF}(t) = -\sum_{i=T+1}^{L}\sum_{j=T+1}^{L}p_{FF}(i,j)\log p_{FF}(i,j) \qquad (16)$$

Since BB and FF depict the gray level transitions between background to background and between foreground to foreground respectively, they represent the local properties. Using Eqns. (13) and (16) a local entropy is defined as

$$H(t) = H_{BB}(t) + H_{FF}(t) \qquad (17)$$

Therefore an optimal local threshold tLE is the one that maximizes Local Entropy. More precisely,

$$t_{LE} = \arg\{\max, H_{LE}(t)\} \qquad (18)$$

This value tLE is used to segment the image.

3. RESULTS OF SEGMENTATION

The test images are taken from Mammographic Image analysis Society database (MIAS) [17]. The results using the proposed approach for microcalcification detection are shown in Figures. In the resulting figures the white portion corresponds to the microcalcifications. Figure 1 shows the input image considered. Figure 2 shows the result of application of the image to the Laplacian of Gaussian filter after subjecting to background subtraction. Figure 3 shows the final binary image showing the microcalcifications. Figure 4 shows another input image of mammogram considered for segmentation.

Its results are shown in Figs. 5 and 6. It is observed that the proposed algorithm performs quite well and is capable of segmenting the microcalcifications from the mammogram. As a second part of the study, we adopt the background subtraction technique and then the texture representation of masses in mammographic images taken from MIAS database. The defuzzified response is calculated for each pixel over the window of size 5x5. The size of window is such that it sufficiently captures the pattern of textures present in the natural image. We apply Entropic thresholding to segment the images. The masses are shown in Fig. 6 and the segmented images of the mammograms are shown in the Fig. 7. Figure 8 shows the background subtraction image.
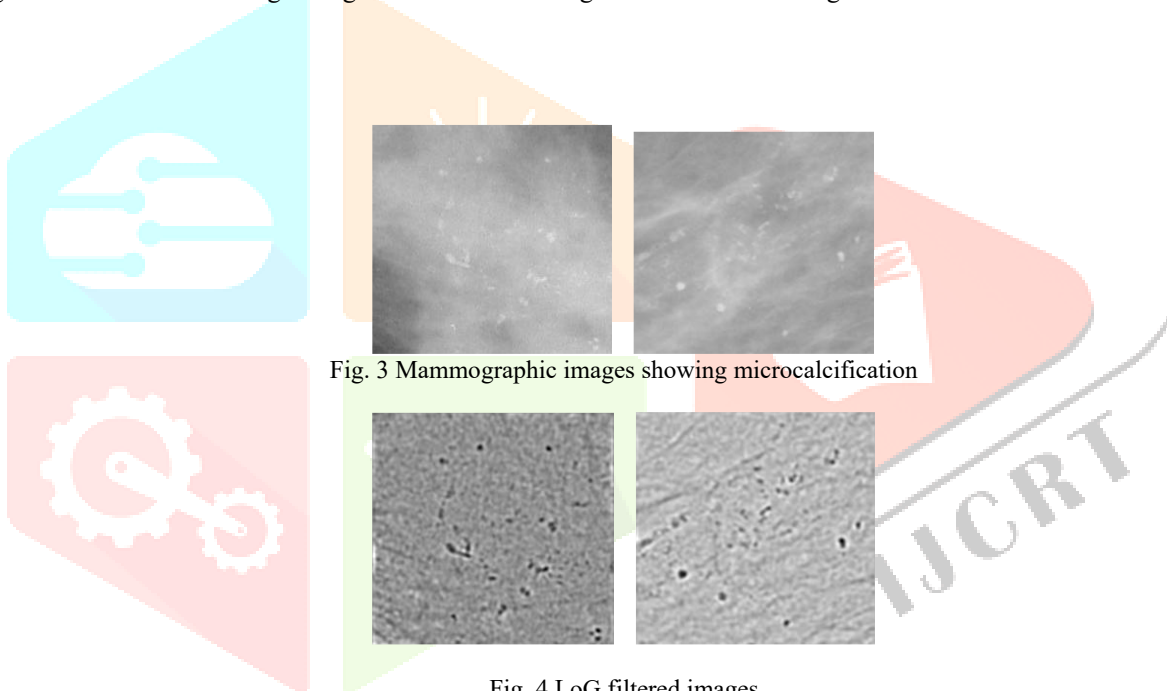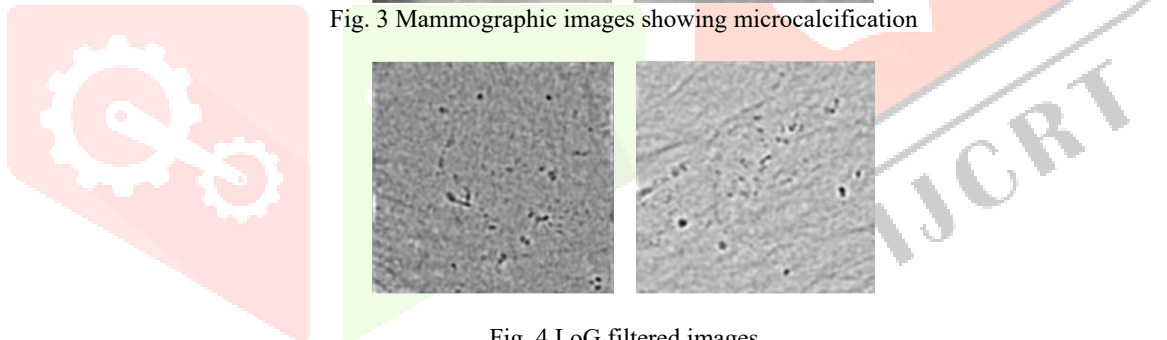

Fig. 3 Mammographic images showing microcalcification
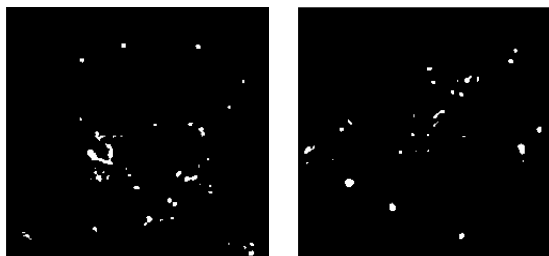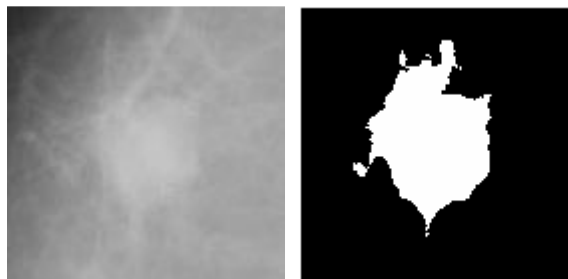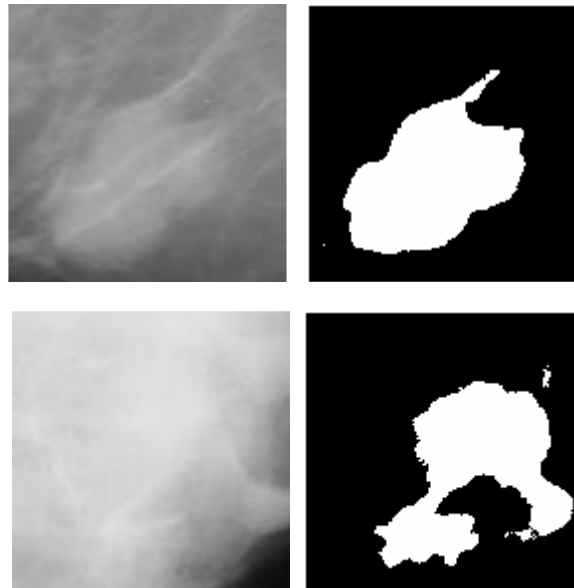

Fig. 4 LoG filtered images

Fig. 5 Segmented images

Fig. 6. Mammographic images showing masses
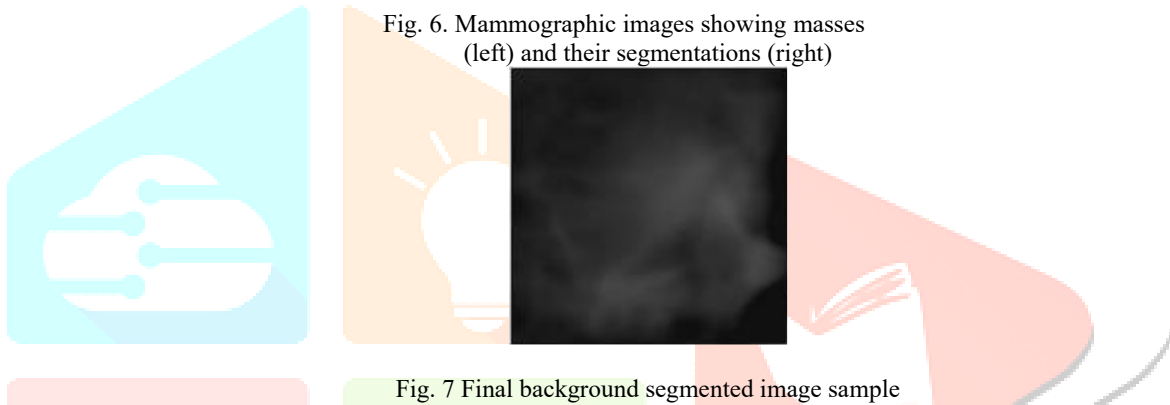(left) and their segmentations (right)

Fig. 7 Final background segmented image sample

4. TEXTURE FEATURE EXTRACTION

As the tissues present in brain are difficult to classify using shape or intensity level of information, the texture feature extraction is found to be very important for further classification [25,26, and 27]. The analysis and characterization of texture present in the medical images can be done using several approaches like run length encoding, fractal dimension, and discrete wavelet transform and co-occurrence matrices [28, 29].

Though many texture features have been used in the medical image classification, Spatial Gray Level Dependent Features (SGLDF) can be used to calculate the inter sample distance for better diagnosis [30, 31]. In order to detect the abnormalities in medical images association rule mining is built using texture information [32, 33]. This information can be categorized by the spatial arrangement of pixel intensities. In order to capture the spatial distribution of the gray levels within the neighborhood, two dimensional co-occurrence matrices can be applied to calculate the global level features and pixel level features.

The following ten descriptors can be used for extracting texture features.

$$\text{Entropy} = -\sum_{i}^{M}\sum_{j}^{N} P[i,j]\log P[i,j]$$

$$\text{Energy} = \sum_{i}^{M}\sum_{j}^{N} P^2[i,j]$$

$$\text{Contrast} = \sum_{i}^{M}\sum_{j}^{N} (i-j)^2 P[i,j]$$

$$\text{Homogeneity} = \sum_{i}^{M}\sum_{j}^{N} \frac{P[i,j]}{1+|i-j|}$$

$$\text{SumMean} = \frac{1}{2}\sum_{i}^{M}\sum_{j}^{N} (i*P[i,j] + j*P[i,j])$$

$$\text{Variance} = \frac{1}{2}\sum_{i}^{M}\sum_{j}^{N}\left((i-\mu)^2 P[i,j] + (j-\mu)^2 P[i.j]\right)$$

$$\text{Maximum\_Probability} = \underset{i,j}{\text{Max}}\, P[i,j]^{M,N}$$

$$\text{Inverse\_Difference\_Moment} = \sum_{i}^{M}\sum_{j}^{N}\frac{P[i,j]}{|i-j|^k}$$

$$\text{Cluster\_Tendency} = \sum_{i}^{M}\sum_{j}^{N}(i+j-2\mu)^k P[i,j]$$

$$\text{Correlation} = \sum_{i}^{M}\sum_{j}^{N}\frac{(i-\mu)(j-\mu)P[i,j]}{\sigma^2}$$

where P is the normalized co-occurrence matrix, (i, j) is the pair of gray level intensities and M by N is the size of the co-occurrence matrix. The inter sample distance is estimated based on estimation of the second-order joint conditional probability density function for the pixel (i, j), P[i, j | d, θ] for θ = 0o ,45o ,90o and135o . The function is the probability that two pixels which are located with an inter sample distance d and a direction θ.The estimated joint conditional probability density functions are defined as

$$P\left[i,j\big|d,0^0\right] = \#\left\{\begin{array}{l}((k,l),(m,n))\in[L_x\times L_y]\times[L_x\times L_y]: \\ k=m\,|l-n|=d, S(k,l)=i, S(m,n)=j\end{array}\right\} / T(d,0^0)$$

$$P\left[i,j\big|d,45^0\right] = \#\left\{\begin{array}{l}((k,l),(m,n))\in[L_x\times L_y]\times[L_x\times L_y]: \\ (k-m=d,l-n=-d)\text{or} \\ (k-m=-d,l-n=d) \\ S(k,l)=i, S(m,n)=j\end{array}\right\} / T(d,45^0)$$

$$P\left[i,j\big|d,90^0\right] = \#\left\{\begin{array}{l}((k,l),(m,n))\in[L_x\times L_y]\times[L_x\times L_y]: \\ (k-m=d,l=n), S(k,l)=i, \\ S(m,n)=j\end{array}\right\} / T(d,90^0)$$

$$P\left[i,j\big|d,135^0\right] = \#\left\{\begin{array}{l}((k,l),(m,n))\in[L_x\times L_y]\times[L_x\times L_y]: \\ (|k-m|=d,l-n=-d), \\ S(k,l)=i, S(m,n)=j\end{array}\right\} / T(d,135^0)$$

where # denotes the number of elements in the set, S(x, y) is the image intensity at the point (x, y) and T(d,θ) stands for the total number of pixel pairs within the image which has the inter sample distance d and direction θ.

Co-occurrence matrices can be calculated for the directions 0 ,45 ,90 ,135 degrees and their respective pixels are denoted as 1, 2, 3 and 4. Once the co-occurrence matrix is calculated around each pixel, the features such as entropy, energy, variance, homogeneity and inverse variance can be obtained for each matrix with respect to the inter sample distance. From the co-occurrence matrices the feature vectors can be calculated and stored in the transaction database. Next, the continuous valued features are discretized into intervals, where each interval represents an item in the process of mining association rules [34].

5. CLASSIFICATION

Associative Classification [35] is a framework that integrates classification and association rule mining [36,37]. The goal of associative classification is to build a model that uses association rules for classification to predict future data objects for which the class label is unknown.

Model Construction generally consists of two major phases: rule generation and classifier building. Firstly, the rule generation is discovering the set of class association rules (CARs) which satisfy the user specified constraints denoted respectively by minimum support and minimum confidence thresholds. Secondly, a classifier is built by choosing a subset of the generated class association rules (CARs). Many studies have shown that Associative Classification is often more accurate than do traditional classification techniques.

When a new training dataset is appended to an old training dataset, the classifier that uses association rules may need to be changed in order to reflect any change in the new training dataset. As a brute force technique to deal with this situation, both old training dataset and a new training dataset are merged into an updated training dataset. Then, the model construction process starts building a classifier based on the updated training dataset. This brute force technique is time consuming and inefficiency.

5.1 ASSOCIATIVE CLASSIFICATION (AC)

Associative Classification is considered as a new approach for classification. The framework of associative classification is integration of classification and association rule mining. The first associative classification algorithm is called Classification Based on Association Rules (CBA) [35].The algorithm has two major phases:

- CBA – Rule Generator (CBA-RG)
- CBA – Classifier Building (CBA-CB).

CBA-RG algorithm generates a complete set of class association rules (CARs) that satisfy the minimum support and minimum confidence thresholds. To generate the set of class association, CBA-RG algorithm finds all large ruleitem by making multiple pass over   data similar to Apiori algorithm. Ruleitems are large ruleitem if their supports are greater than or equal to minimum support. For all ruleitems with the same condset, the ruleitem that have the highest confidence is chosen as possible rule. The result of this step is the set of CARs.

CBA-CB algorithm sorts the set of CARs according to the precedence relation $(>)$. The rule ranking is defined as follows: Given two rule r_i and r_j ; $r_i > r_j$ (r_i has higher precedence over r_j), if one of the following holds good:
 1. The confidence of r_i is greater than that of r_j
2. Their confidence are the same but support of r_i is greater than that of r_j
3. Both confidences and supports of r_i and r_j are the same, but r_i is generated before r_j

After rule ranking, each training instance is covered by a rule having the highest precedence among the rules that can cover the case. The rules that do not cover any training instances are removed. Then, training instances that do not fall into any of the observed classed are added to a default class. Finally, rules that do not improve the accuracy of the classifier are discarded. The remaining rules and the default class of the last rule are formed as associative classifier.

5.2 AN INCREMENTAL UPDATING TECHNIQUE

When new transactions are added to the database shown in figure 8, association rules may be changed. For dynamic databases, several incremental updating techniques have been developed for mining association rule. An Incremental Updating Technique [38,39] is proposed for dynamic database which new transactions are appended.

The concept of incremental updating technique is to reuse large itemset of previous mining to obtain the update large itemset of an incremental database. Fast Update algorithm (FUP) was first introduced in [40]. The algorithm handles database with transactions insertion only. An efficient algorithm FUP is presented for computing the large itemset in the updated database. It is shown that the information from the old large itemset can be reused.
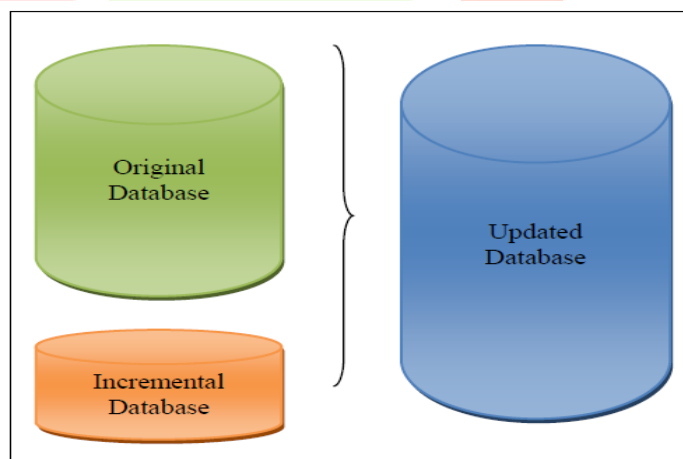

Fig. 8. Incremental Database

5.3 INCREMENTAL CLASSIFICATION BASED ON ASSOCIATION RULE ALGORITHM

When a new training dataset is appended to an old training dataset, an associative classifier may need to be changed in order to reflect any change in the new training dataset. However, when the training dataset changed, the existed Associative Classification algorithm always scans the changed training dataset in order to reflect the changes done. So far, there are rarely researches on the incremental learning of associative classification but there had been some studies on incremental association rule discovery algorithms that we can use their ideas for reference to solve the incremental associative classification problem. We propose an efficient incremental associative classification algorithm;   a new algorithm is proposed to update an associative classifier when a new training dataset is appended to an old training dataset. The algorithm called Incremental Classification Based on Association Rules (ICBA) algorithm [38], is based on the concept of FUP algorithm to solve this problem.
The algorithm is divided into two parts. As the first part, ICBA-RG algorithm shown in figure 9 discovers the set of class association rule (CARs′) in updated training datasets. Then, the second part is building a classifier for an updated training datasets. The algorithm for the second part shown in figure10 is called ICBA-CB algorithm.

According to Figure 9, the steps of rule generation part are outlined as follows. An incremental training dataset (t) is scanned to determine large 1-ruleitem of an updated training dataset (UT) shown at line 1-16. If a candidate 1-ruleitem is a member of previous large 1-ruleitem, its support is updated. On the other hand, if a candidate 1-ruleitem is not a member of old large 1-ruleitemset, our algorithm checks the support of the ruleitem in an incremental training datasets (t). If the support of the ruleitem in an incremental training datasets is equal or above minimum support of incremental training datasets i.e., support $\geq$ (s×d) when s is minimum support and d is size of incremental training dataset, then the algorithm scans original training datasets (T) to update support count of ruleitem (X.supportUT). In this paper, we called ruleitem which its support is greater than minimum support "winner" and ruleitem which its support is lower than minimum support "loser". As shown at line 17-33, the large k-ruleitems of updated training datasets are determined when k is greater than or equal to 2. Candidate k-ruleitems are generated by applying candidateGen function shown in line 18, this function is joining step similar to Apriori algorithm (see this function details in [38]. At k-th iteration, loser in L_k will be filtered out in a scan of t. The filtering is done by two steps. Firstly, a large k-ruleitems in L_k containing the ruleitem that cannot be the winner in the k-th iteration will be filtered out by ruleSubset function shown in line 19.

And the second, ICBA-RG filtered out loser in L_k without checking it against t. The set of losers $Y = L_k - L_{k-1}$ have been identified in line 21. Therefore, any sets of $X \in L_k$, which have subset Y such that $Y \in L_k - L_{k-1}$, cannot be large ruleitem and are filtered out from L_k. Then, if a ruleitem is a member of L_k and its support is equal or above $s \times (D+d)$ it becomes large k-itemset of update training datasets (L'_k). On the other hand, if a ruleitem is a member of C'_k and its support less than s×d, the item will be removed from candidate k-ruleitem if the support of ruleitem is equal or above (s× (D+d), the ruleitem is inserted into L'_k which will be generated to class association rules (CARs) by genRule function[3.] at line 14 and 32. Finally, pruneRule function shown at line 15 and 33 is prune CARs by minimum confidence same as Apriori algorithm. All rules in CARs′ which have their confidence less than minimum confidence ICBA-RG will be filter out.

```
Input : UT =  The updated training datasets
        T   =  The original training datasets with the total number of
               transactions, equal to D.
        t   =  The incremental training datasets with the total of
               number transaction equal to d.
        L'k  =  The set of large-ruleitems in UT
        W   =  L1 (large 1-ruleitem of T)
        s   =  minimum support

Output : CARs′ = The set of class association rules of updated
                  training datasets.

1.   for all X ∈ t do
2.       if  X ∈ W do
3.         scan t to update X.supportUT then
4.             if  X.supportUT ≥ s × (D + d)  do
5.                 if  X.supportUT ≥ s × (D + d)  do
6.                     insert X at the end of L'1
7.       else
8.          if  X ∉ W do
9.              for all X ∉ W do
10.                 if  X.supportUT ≥ (s × d)  do
11.                     scan UT to update X.supportUT then
12.                         if  X.supportUT ≥ s × (D + d)  do
13.                             insert X at the end of L'1
14.  CARs1 = genRules(L'1)
15.  prCARs1′ = pruneRules(CARs1)
16.  end
17.  for (k=2; Lk-1 ≠∅; k++) do
18.      C'k = candidateGen (L'k-1) - Lk
19.      C's  = ruleSubset (C'k)
20.              for all k-1 ruleitem in C's  do
21.                  Y = Lk-1 - L'k-1  do
22.                      if  Y ⊆ X then W = W-{X}
23.          for all X ∈ C's do
24.              if  X.supportUT ≥ s × (T + t)  do
25.                  scan UT to update X.supportUT then
26.                      insert X at the end of L'k
27.          for X ∈ W do
28.              scan UT to update X.supportUT then
29.                  if  X.supportUD ≥ s × (T + t)  do
30.                      insert X at the end of L'k
31.          end
32.  CARs'k = genRules(L'k)
33.  prCARs'k = pruneRules(CARs'k)
34.  end
```
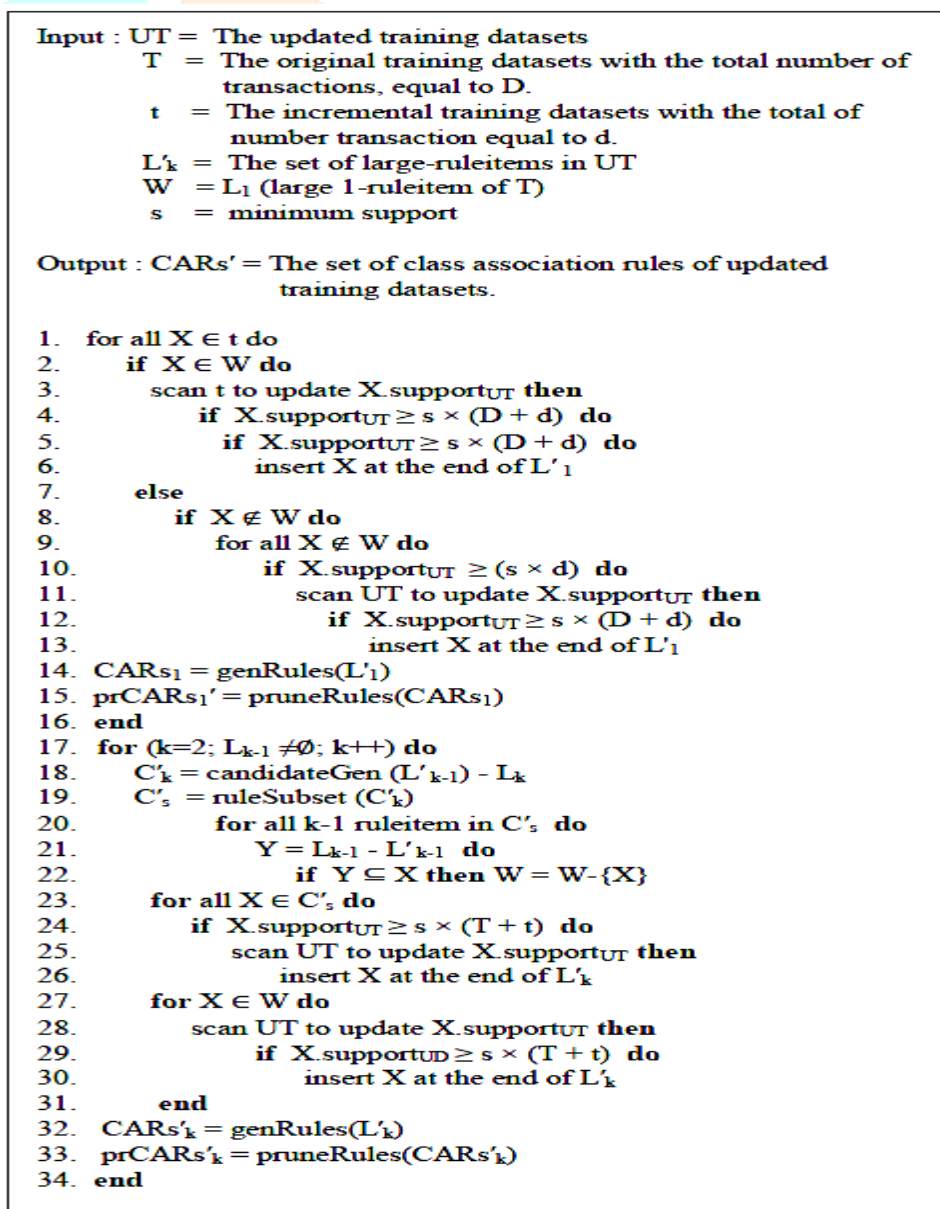
Fig. 9 ICBA-RG algorithm

For the last phase, ICBA-CB algorithm shown in figure 10 builds a classifier using CARs′. In this phase, in order to reduce the execution time we try to scan transactions of training datasets as less as we can. ICBA-CB has three steps to build the classifier. The first step which is at line 1 is sorting the set of CARs′ according to the relation "$>$"

Then, the second step at line 2-24 is selecting rules for classifier following the sorted sequence. For each rule which is member of CARs, we go through t to find those cases covered by the rule. If selected rules are not member of CARs, our algorithm goes through UT to find those covered case instead. We mark selected rules if they correctly classify a case (line 6 and line 17). If selected rules can correctly classify at least one case, they may be our potential rule in our classifier. The rules that do not cover any case are removed and the cases that do not fall into any of the observed classes are added to a default class (in case we stop selecting more rule for our classifier (C′)). The algorithm computes and records the total number of error made by classifier and default class. When there is no rule of training case left, the rule selection process is completed.

```
Input : CARs  = Set of class association rules of original training
                 datasets
          R′    = Class association rule is CARs′
          UT   = Updated training datasets
          T    = Original training datasets
          t    = Incremental training datasets
Output :   C′   = Classifier

1.   R′ = sort (R′)
2.   for each rule  r′ ∈ R′ in sequence do
3.       if r′ ∈ CARs then
4.         for each case  X in t do
5.             if X satisfies the condition of r′ do
6.                 store X.id in temp and mark r′ if it
                   correctly classifies t ;
7.         if r′ is marked then
8.         insert r′ at the end of C′ (our classifier)
9.            delete all the ruleitem with the id in temp from db
10.           selecting a default class for the current C′
11.           compute the total number of errors of C′
12.      end
13.   else
14.     if r′ ∉ CARs then
15.         for each case  X in UT do
16.             if X satisfies the condition of r′ do
17.                 store X.id in temp and mark r′ if it
                    correctly classifies X;
18.         if r′ is marked then
19.         insert r′ at the end of C′ (our classifier)
20.           delete all the ruleitem with the id in temp from UT
21.           selecting a default class for the current C′
22.           compute the total number of errors of C′
23.   end
24. end
25. Find the first rule p′ in C′ with the lowest total
     number of errors and drop all the rule after p′ in C′
26. Add the default class associated with p′ to end of C′
     and return C′
```

Fig. 10 ICBA-CB algorithm

The third step at line 25-26 is discarding those rules in the classifier that do not improve the accuracy. The cutoff rule is the first rule at which there is the least number of errors recorded on UT. The remaining rules and the default class of the last rule in the classifier form our classifier.

6. EXPERIMENTAL RESULTS

The digital mammograms used in our experiments were taken from the Mammographic Image Analysis Society (MIAS). The database consists of 322 images, which belong to three categories: normal, benign and malign (ftp://peipa.essex.ac.uk). There are 208 normal images, 63 benign and 51 malign, which are considered abnormal.

The proposed method is evaluated based on ten-fold cross validation method. The following table presents the rule accuracy of the proposed classification system compared with other association rule based system proposed in [42,43]. The results for the ten splits of the mammogram database are given in Table I.

TABLE I

CLASSIFICATION ACCURACY FOR THE

TEN SPLITS WITH MGARM

Splits	Classification Accuracy
1	98.65
2	95.89
3	97.76
4	93.12
5	92.23
6	98.66
7	94.41
8	98.54
9	98.05
10	95.89
Average	**96.32**

In this paper we used multi dimensional genetic association rule mining using image contents for the classification of mammograms. The average accuracy is 96.32 %. We have employed the freely available Machine Learning package, WEKA [41]. Out of 322 images in the dataset, 230 were used for training and the remaining 92 for testing purposes and the result is shown in Table 2.

TABLE 2

RESULTS OBTAINED BY PROPOSED METHOD

Normal	100%
Malignant	88. 23%
Benign	97.11%

The confusion matrix has been obtained from the testing part .In this case for example out of 51 actual malignant images 06 images was classified as normal. In case of benign all images are correctly classified and in case of normal images 6 images are classified as malignant. The confusion matrix is given in Table 3.

TABLE 3

CONFUSION MATRIX

Actual	Predicted class		
	Benign	**Malignant**	**Normal**
Benign	63	0	0
Malignant	51	45	06
Normal	208	6	202

7. CONCLUSIONS AND FUTURE WORK

Automated breast cancer detection has been studied for more than two decades Mammography is one of the best methods in breast cancer detection, but in some cases radiologists face difficulty in directing the tumors. We have described a comprehensive of methods in a uniform terminology, to define general properties and requirements of local techniques, to enable the readers to select the efficient method that is optimal for the specific application in detection of micro calcifications in mammogram images. The CAD mammography systems for micro calcifications detection have gone from crude tools in the research laboratory to commercial systems. Mammogram image analysis society database is standard test set but defining different standard test set (database) and better evaluation criteria are still very important. With some rigorous evaluations, and objective and fair comparison could determine the relative merit of competing algorithms and facilitate the development of better and robust systems. The methods like one presented in this paper could assist the medical staff and improve the accuracy of detection.

A new approach is presented for the detection of micro-calcifications and masses. The detection of individual micro-calcifications is based on the LoG filter response of a background subtracted mammogram. The identification of candidate pixels is done by estimating the local contrast measure of the pixel. We have tested our system in the MIAS mammographic databases with satisfactory results. The masses are generally recognizable by texture. In this paper, we attempt to segment mammogram containing masses by its texture. A fuzzy technique is explored in this work. A membership function to this effect is defined by the sigmoid function. The defuzzified image is segmented using Entropic thresholding. In the present work, a computer-aided system for

segmentation of masses is presented. Our major concern is to identify tumor in an image and it has been observed that the tumor is clearly recognizable by its texture information. Results of segmentation obtained for both the cases reveal that they match those observed by an experienced radiologists indicating thereby that the proposed fuzzy technique performs well in terms of reduction of false positive rates. In further research we shall extract the features from the identified masses and will focus on improving feature selection performance with different data mining classifiers as well as shall classify the tumor masses to normal, benign and malignant class.

In this study, we propose an improved   new classification based on association rules   algorithm called Incremental classification based on association rules (ICBA). The experiment results show that our algorithm is more efficient than CBA algorithm. In the future, further researches and experiments on the proposed algorithm will be presented. In this paper, a new method for association rule mining is proposed. The main features of this method are that it only scans the transaction database once, it does not produce candidate jtemsets, and it adopts the Boolean vector "relational calculus" to discover frequent itemsets. In addition, it stores all transaction data in binary form, so it needs less memory space and can be applied to mining large databases.

The methods like one presented in this paper could assist the medical staff and improve the accuracy of detection. Our method can reduce the computation cost of mammogram image analysis and can be applied to other image analysis applications. The algorithm uses simple statistical techniques in collaboration to develop a novel feature selection technique for medical image analysis.

References

[1]   P. A. Wingo, T. Tong and S. Bolden, "Cancer statistics," CA. J. Clin.,vol. 45, pp. 8–30, 1995.

[2]   W. K. Pratt, Digital Image Processing: 3rd Edition, Wiley-Interscience; 2001.

[3]   R. C. Gonzalez and Richard E. Woods Digital Image Processing, Prentice Hall; 2nd Edition, 2002.

[4]   J.N. Kapur, P.K. Sahoo and A.K.C. Wong, "A new method of gray level picture thresholding using entropy of the histogram", Computer Vision & Graphics Image processing, Vol. 29, 1985, pp. 273-285.

[5]   Li E.H., Lee C.K. "Minimum cross-entropy thresholding", Pattern Recognition, pp. 617- 625,1992.

[6]   A.S. Abutaleb, "Automatic thresholding of gray level pictures using two dimensional entropy", Computer Vision and Graphics Image Processing, Vol. 47, 1989, pp. 22-32.

[7]   M. Hanmandlu, V.K. Madasu and S. Vasikarla, "A fuzzy Approach to texture segmentation", Proc. International Conference on Information Technology: Coding and Computing, vol-1, pp.636-642, 2004.

[8]   V. Chalana and Y. Kim, "A Methodology for Evaluation of Boundary algorithms on Medical Images," IEEE Trans. on Medical Imaging, vol. 16(5), pp. 642-52, 1997.

[9]   R Gupta et al., "The use of texture analysis to delineate suspicious masses in mammography", Phys. Med. Biol., vol 40 835- 855, 1995.

[10]   N.R. Pal and S.K. Pal, "Entropic Thresholding", Signal Processing, vol.16, pp.97-108,1989.

[11]   R. Gupta and P.E. Undrill, "The use of texture analysis to identify suspicious masses inmammography", Phys. Med. Bio., vol 15. 835-855, 1997.

[12]   J. S. Weska et al., "A survey of threshold selection techniques", Computer Graphics &Image Processing, vol. 7. pp. 259-265, 1978.

[13]   W. A. Perkins, "Area segmentation of images using edge points." IEEE Trans. on Pattern Analysis & Machine Intelligence, vol. PAMI-2, pp. 8-15. 1980.

[14]   S. Zucker. "Region growing: Childhood and adolescence", Computer Graphics and ImageProcessing, vol. 5. pp. 382-399. 1976.

[15]   M. Pietikainen and A. Rosenfeld. "Image segmentation by texture using pyramid nodelinking," IEEE Trans. on Systems, Man & Cybernetics, vol. 11, pp. 822-825. 1981.

[16]   P. C. Chen and T. Pavlidis, "Segmentation by texture using a co occurrence matrix,"
Computer Graphics and Image Processing,vol. 10. pp. 172-182. 1979

[17]   Mammographic Image Analysis Society database: http://peipa.essex.ac.uk/info/mias.html

[18]   http://homepages.inf.ed.ac.uk/rbf/HIPR2

[19]   S-C. Yang, C-M. Wang, Y-N. Chung, G-C.Hsu, S-K. Lee P-C. Chung and C-I. Chang,"A Computer-Aided System for MassDetection and Classification in Digitized Mammograms", Biomedical Engineering applications: Basis & Communications, Vol. 17 No. 5 October 2005 pp. 215-229

[20]   G. Gunther and D. Ivo, Physica, Berlin (2000) pp. 545–565.

[21]   L. Arbach, N. AI-Mudarres and M. Ammar, "Computerized detection of breasts from mammograms",   The Thirty Seventh Week of Science, Damascus. Syria. June 1997.

[22]   Z. R. Yang and R. G. Harrison, "Detecting false benign in breast cancer diagnosis", IEEETrans. On Neural Networks, vol. 3, pp. 655458. July 2000.

[23]   N. Pandey and J. Sivaswamy, "Fuzzy logic based micro-calcification detection", IEEEConference on Neural Networks, vol. 2, pp. 662-671, December 2000..

[24]   Chan et al., "Computer-aided classification of mammographic masses and normal tissue: linear discriminant analysis in texture feature space", Phys. Med. Biol., vol. 40, 857-876.

[25]   R. Abraham, J.B. Simha,and S.S Iyengar, "Medical datamining with a new algorithm for Feature   Selection and Naive Bayesian classifier.," In Proc: 10th International Conference on Information Technology (ICIT), 2007, pp. 44-49.

[26]   C. Christophe, S.G. Jean, L.M. Gael, and K. Michel, "Efficient Data Structures and Parallel Algorithms for Association Rules Discovery," In Proc: Fifth Mexican International Conference in Computer Science (ENC), 2004, pp. 399-406.

[27]   P.G. Foschi, D.Kolippakkam, H. Liu, and A. Mandvikar , " Feature extraction for image mining," In Proc: 8th Int. Workshop Multimedia Inf. Syst, Tempe, AZ, 2002, pp. 103-109.

[28] R.M. Haralick, K.Shanmugam, and I. Distein, "Textural features for image classification.," IEEE Trans. Syst, Man, Cybern, vol. SMC-3, 1973, pp. 610–621.

[29] A. Ranjit, B.S. Jay, and S.S. Iyengar, "Medical Data mining with a New Algorithm for Feature Selection and Naive Bayesian Classifier," In Proc: 10th International Conference on Information Technology (ICIT), 2007, pp.44-49.

[30] L. Hui, W. Hanhu, C. Mei, and W. Ten , "Clustering Ensemble Technique Applied in the Discovery and Diagnosis of Brain Lesions," In Proc: Sixth International Conference on Intelligent Systems Design and Applications (ISDA) , vol. 2: 2006, pp. 512-520.

[31] C.F. Joaquim, X.R. Marcela, P.M.S. Elaine, J.M.T. Agma, and T.J. Caetano, "Effective shape-based retrieval and classification of mammograms," In Proc: ACM symposium on Applied computing, 2006, pp. 250 – 255.

[32] N.R. Mudigonda, R.M. Rangayyan , "Detection of breast masses in mammograms by density slicing and texture flow-field analysis," IEEE Trans. Med. Imag. 20(12), 2001, pp. 1215–1227.

[33] K. Murat, I.M. Cevdet,"An expert system for detection of breast cancer based on association rules and neural network," An International Journal Expert Systems with Applications 36: 2009, pp. 3465–3469.

[34] J. Dougherty, R. Kohavi, and M. sahami, "Supervised and Unsupervised Discretization of continuous features," In Proc: 12th International Conference Machine Learning, 1995, pp.56-69.

[35] Bing Liu., Wynne Hsu., Yimming Ma. "Integrating Classification and Association Rule Mining", In Proc. Of the Fourth International Conference on Knowledge Discovery and Data Mining, New York, NY, pp.80-86, 1998

[36] R. Agrawal., T. lmielinski, and A. Swami, " Mining association rule between sets of items in large database", In Proceeding of the ACM SIGMOD Int'l Conf. on Management of Data, Washington, USA, May 1993, pp. 207-216.

[37] R. Agrawal, and R. Srikant, "Fast Algorithm for Mining Association Rules," Proceedings of the International Conference on Very Large Database, Santiago, Chile, 1944, pp. 487-499

[38] D. Cheung, S.D. Lee, and B. Kao. "A General Incremental Technique for Maintaining Discovered Association Rules," Proceedings of the 5th International Conference on Database System for Advanced Applications, Melbourne, Australia, 1997, pp. 185-194.

[39] F. Thabtah. "Challenges and Interesting Research Directions in Associative Classification," Proceeding of the Sixth IEEE International Conference on Data Mining Workshops, 2006, pp.785-792.

[40] D. Cheung, J. Han, V. Ng, and C. Y. Wong. "Maintenance of Discovered Association Rules in Lange Database: An Incremental Updating Technique," Proceedings of the 12th IEEE International Conference on Data Engineering, 1996, pp. 106-114.

[41] Holmes, G., Donkin, A., Witten, I.H.: WEKA: a machine learning workbench. In: Proceedings Second Australia and New Zealand Conference on Intelligent Information Systems, Brisbane, Australia, pp. 357-361, 1994.

[42] ML Antonie, OR. Zaiane, and A Coman, "Application of data mining techniques for medical image classification". In Proc. Of Second Intl. Workshop on Multimedia Data Mining (MDM/KDD'2001) in conjunction with Seventh ACM SIGKDD, pp 94–101, San Francisco, USA, 2001.

[43]]Deepa S. Deshpande "ASSOCIATION RULE MINING BASED ON IMAGE CONTENT" International Journal of Information Technology and Knowledge Management January-June 2011, Volume 4, No. 1, pp. 143-146