International Journal of Application or Innovation in Engineering & Management (IJAIEM)
Web Site: www.ijaiem.org Email: editor@ijaiem.org
Volume 5, Issue 9, September 2016 **ISSN 2319 - 4847**

Image Mining for Mammogram Classification by Associative Classifier with Negative Rules Using Texture Features

Aswini kumar mohanty [1] Amalendu Bag [2] Swasati Sahoo [3] Arati Pradhan [4]

[1]KMBB Collegw of Engg & Tech,Daleiput,Khurda, Odisha

[2]KMBB Collegw of Engg & Tech,Daleiput,Khurda, Odisha

[3]Gandhi Engineering College,odisha

[4]UN Autonomous College, Adashapur,. Odisha,

ABSTRACT

Image mining is concerned with knowledge discovery in image databases. Image mining deals with the extraction of implicit knowledge, image data relationship, or other patterns not explicitly stored in the image databases. The focus of image mining is in extraction of patterns from large collection of images. Breast cancer is the leading cause of cancer death among women. Screening mammography is the only method currently available for the reliable detection of early and potentially curable breast cancer. Research indicates that the mortality rate could decrease by 30% if women age 50 and older have regular mammograms. The detection rate can be increased 5-15% by providing the radiologist with results from a computer-aided diagnosis (CAD) system acting as a second opinion.

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. The texture statistical second considered is spatial gray level dependence method, gray level run length method and gray level difference method. Features are extracted from the first-order statistical method and second-order statistical method and are combined. It is observed that the result of these combined features provides higher accuracy when compared with the features from the first-order statistical method and second-order statistical method alone.

Since mammography is considered as the most effective means for breast cancer diagnosis, this paper introduces multi dimensional genetic association rule mining for classification of mammograms. The purpose of our experiments is to explore the feasibility of image mining approach to extract patterns and whether that pattern will be helpful to diagnose breast cancer and tissue as well as increase the diagnostic accuracy of image processing and data mining techniques for optimum classification between normal and abnormalities in digital mammograms. Results shows very promising and the accuracy level is very high in compared to other techniques in case of image mining based on negative association rule mining. It is well known that data mining techniques are more suitable to larger databases than the one used for these preliminary tests. Computer-aided method using association rule could assist medical staff and improve the accuracy of mammogram detection. In particular, a Computer aided method based on association rules becomes more accurate with a larger dataset .Experimental results show that this new method can quickly and effectively mine potential association rules.

Keywords: Mammogram, Gray Level Co-occurrence Matrix features, Histogram Intensity, Region growing Classification, Genetic Algorithm; Association rule mining, Confusion matrix

1. INTRODUCTION

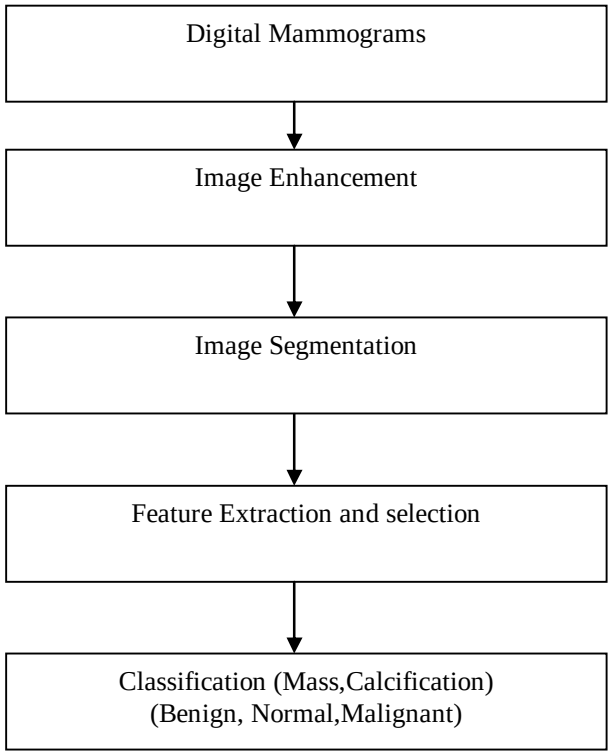
Breast Cancer is one of the most common cancers, leading to cause of death among women, especially in developed countries. There is no primary prevention since cause is still not understood. So, early detection of the stage of cancer allows treatment which could lead to high survival rate. Mammography is currently the most effective imaging modality for breast cancer screening. However, 10-30% of breast cancers are missed at mammography [1]. Mining information and knowledge from large database has been recognized by many researchers as a key research topic in database system and machine learning and researches that use data mining approach in image learning can be found in [2,3].

Data mining of medical images is used to collect effective models, relations, rules, abnormalities and patterns from large volume of data. This procedure can accelerate the diagnosis process and decision-making. Different methods of data mining have been used to detect and classify anomalies in mammogram images such as wavelets [4,5], statistical methods and most of them used feature extracted using image processing techniques [6].Some other methods are based

on fuzzy theory [7,8] and neural networks [9]. In this paper we have used classification method called a associative classifier using negative rule using texture features and it is proposed for negative rule construction. The result shows that the proposed rule-based approach reaches the classification accuracy over 95% and also demonstrates the use and effectiveness of association rule mining in image classification [10-12].

Segmentation is one important part is mammogram classification. It segregates the affected part of the breast instead of taking whole part of the image which enhances the computational cost and incurs overhead. The segmentation process can be manual or automated. The main idea behind the segmentation is to select the Region of Interest (ROI) rather than unwanted portion of the image. Manual segmentation to find ROI is possible for radiologist not for untrained people and hence automated segmentation using region going is used in our proposed method to find the ROI.

Classification process typically involves two phases: training phase and testing phase. In training phase the properties of typical image features are isolated and based on this training class is created .In the subsequent testing phase , these feature space partitions are used to classify the image. We have used supervised genetic association rule method by extracting low level image features for classification. The merits of this method are effective feature extraction, selection and efficient classification. The steps involved in processing mammograms for classification is shown in figure 1. The rest of the paper is organized as follows. Section II presents the pre-processing and section III presents the segmentation and section IV presents feature extraction phase. Section V discusses the proposed method of Feature selection and classification. In section VI the results are discussed and conclusion is presented in section VII.



. **Figure 1** Steps involved in diagnosing of Breast cancer

2. PREPROCESSING

The mammogram image for this study is taken from Mammography Image Analysis Society (MIAS)†, which is an UK research group organization related to the Breast cancer investigation [13]. As mammograms are difficult to interpret, preprocessing is necessary to improve the quality of image and make the feature extraction phase as an easier and reliable one. The calcification cluster/tumor is surrounded by breast tissue that masks the calcifications preventing accurate detection and shown in Figure.1. .A pre-processing; usually noise-reducing step is applied to improve image and calcification contrast figure 2. In this work [14] efficient filter referred to as the low pass filter was applied to the image that maintained calcifications while suppressing unimportant image features.

Figure 1.(a) shows original mammogram and figure 1.(b) shows output image after noise and artifact removal of the figure 1,(a) image cluster. By comparing the two images, we observe background mammography structures are removed while calcifications are preserved. This simplifies the further tumor detection step. Figure 1.(c) displays the mammogram image after filtering and in all images the calcification is preserved.

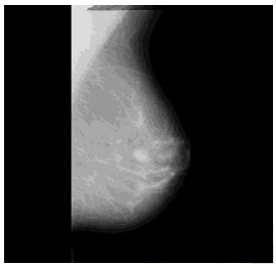


Figure 2a. Original mammogram (mdb 010).

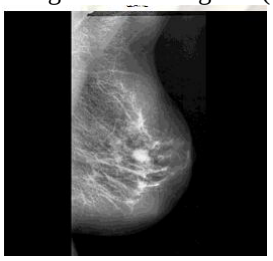


Figure 2b. mammogram after noise and artifact removal process.

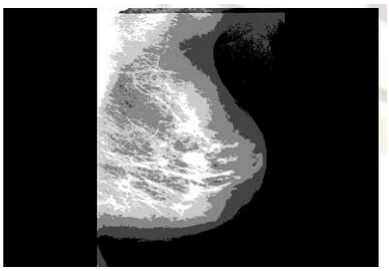


Figure 2c. Mammogram after contrast enhancement process.
ROI after Pre-processing Operation

2.1Histogram Equalization

Histogram equalization is a method in image processing of contrast adjustment using the image's histogram [15]. Through this adjustment, the intensities can be better distributed on the histogram. This allows for areas of lower local contrast to get better contrast. Histogram equalization accomplishes this by efficiently spreading out the most frequent intensity values. The method is useful in images with backgrounds and foregrounds that are both bright or both dark. In particular, the method can lead to better views of bone structure in x-ray images, and to better detail in photographs that are over or under-exposed. In mammogram images Histogram equalization is used to make contrast adjustment so that the image abnormalities will be better visible. Figure2.(a) shows the histogram of the breast image of figure 1.(b) The histogram equalization method forces image intensity levels to be redistributed with an equal probability of occurrence which is shown in Figure 2.(b) and intensity is redistributed by uniform probability density function .

† peipa.essex.ac.uk/info/mias.html

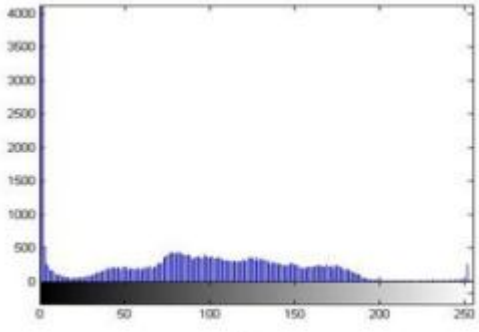


Figure 3.a Original Histogram of Mammogram

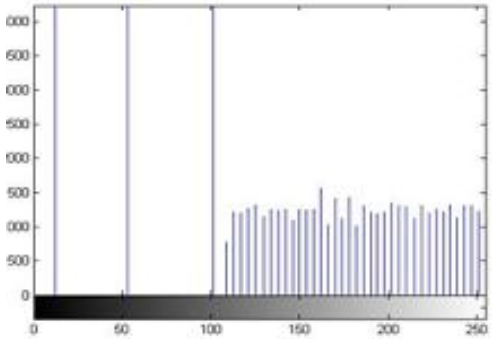


Figure 3.b Histogram Equalization of mammogram

3. SEGEMANTATION

The main objective of image segmentation is to extract various features of the images which can be merged or split in order to build objects of interest on which analysis and interpretation can be performed [36]. Image segmentation refers to the process of partitioning an image into groups of pixels which are homogeneous with respect to some criterion. The result of segmentation is the splitting up of the image into connected areas. Thus segment is concerned with dividing an image into meaningful regions. The image segmentation techniques such as thresholding, region growing, statistics models, active control modes and clustering have been used for image segmentation because of the complex intensity distribution in medical images, thresholding becomes a difficult task and often fails [16-17].

3.1Regiongrowing

It is a simple region-based image segmentation method. It is also classified as a pixel-based image segmentation method since it involves the selection of initial seed points [18.19].
This approach to segmentation examines neighbouring pixels of initial seed points and determines whether the pixel neighbours should be added to the region. The process is iterated on, in the same manner as general data clustering algorithms. A general discussion of the region growing algorithm is described below.
A simple approach to image segmentation is to start from some pixels (seeds) representing distinct image regions and to grow them, until they cover the entire image. For region growing we need a rule describing a growth mechanism and a rule checking the homogeneity of the regions after each growth step
The growth mechanism – at each stage k and for each region Ri(k), i = 1,…,N, we check if there are unclassified pixels in the 8-neighbourhood of each pixel of the region border Before assigning such a pixel x to a region Ri(k),we check if the region homogeneity: P(Ri(k) U {x}) = TRUE , is valid The arithmetic mean m and standard deviation sd of a class Ri having n pixels:

$$M = (1/n)(r,c)\text{€}R(i) \sum I(r,c)$$
$$s.d = \text{Square root}((1/n)(r,c)\text{€}R(i) \sum[I(r,c)-M]2)$$

Can be used to decide if the merging of the two regions R1,R2 is allowed, if

$M1 – M2| < (k)s.d(i)$, i = 1, 2 , two regions are merged

Homogeneity test: if the pixel intensity is close to the region mean value

$$|I(r,c) – M(i)| <= T(i)$$

Threshold Ti varies depending on the region Rn and the intensity of the pixel I(r,c).It can be chosen this way:

$$T(i) = \{ 1 – [s.d(i)/M(i)] \} T$$

The first step in region growing is to select a set of seed points. Seed point selection is based on some user criterion (for example, pixels in a certain gray scale range, pixels evenly spaced on a grid, etc.). The initial region begins as the exact location of these seeds [20].

The regions are then grown from these seed points to adjacent points depending on a region membership criterion. The criterion could be, for example, pixel intensity, gray scale texture, or colour.
Since the regions are grown on the basis of the criterion, the image information itself is important. For example, if the criterion were a pixel intensity threshold value, knowledge of the histogram of the image would be of use, as one could use it to determine a suitable threshold value for the region membership criterion.
There is a very simple example followed below. Here we use 4-connected neighbourhood to grow from the seed points. We can also choose 8-connected neighbourhood for our pixels adjacent relationship. And the criteria we make here is the same pixel value. That is, we keep examining the adjacent pixels of seed points. If they have the same intensity value with the seed points, we classify them into the seed points. It is an iterated process until there is no change in two successive iterative stages. Of course, we can make other criteria, but the main goal is to classify the similarity of the image into regions. Figure 3(a) shows mammograms after region generation process and figure 3(b) shows the final segmentation.

Figure 4(a). Mammogram after region generation process

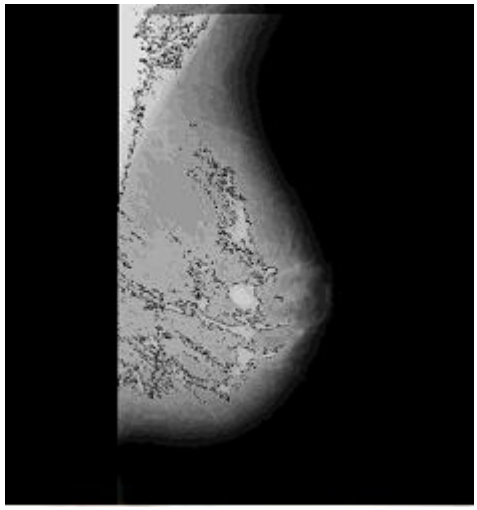


Figure 4.(b). Mammogram after final segmentation.

Features, characteristics [33,34,35]of the objects of interest, if selected carefully are representative of the maximum relevant information that the image has to offer for a complete characterization a lesion [21, 22]. Feature extraction methodologies analyze objects and images to extract the most prominent features that are representative of the various classes of objects. Features are used as inputs to classifiers that assign them to the class that they represent.
In texture analysis field, statistical texture is the most widely used method for quality grading or classification [21, 22]. Statistical texture methods can be classified in two categories. Firstly, the first order statistical methods are characterized by the pixel grey level distribution and organisation. Secondly, the second order statistical methods such as SGLDM, GLRLM and GLDM are considered. Texture image analysis procedure can be defined as a system in which input is an image and the output is a series of features provided by the analysing techniques implemented. Each image is then characterized by a vector of features.
In this Work intensity histogram features and Gray Level Co-Occurrence Matrix (GLCM) features are extracted.

4. TEXTURE FEATURE EXTRACTION

Features, characteristics [33,34,35]of the objects of interest, if selected carefully are representative of the maximum relevant information that the image has to offer for a complete characterization a lesion [21, 22]. Feature extraction methodologies analyze objects and images to extract the most prominent features that are representative of the various classes of objects. Features are used as inputs to classifiers that assign them to the class that they represent.
In texture analysis field, statistical texture is the most widely used method for quality grading or classification [21, 22]. Statistical texture methods can be classified in two categories. Firstly, the first order statistical methods are characterized by the pixel grey level distribution and organisation. Secondly, the second order statistical methods such as SGLDM, GLRLM and GLDM are considered. Texture image analysis procedure can be defined as a system in which

input is an image and the output is a series of features provided by the analysing techniques implemented. Each image is then characterized by a vector of features.

4.1Intensity Histogram Features

Intensity Histogram analysis has been extensively researched in the initial stages of development of this algorithm [23]. Prior studies have yielded the intensity histogram features like mean, variance, entropy etc. These are summarized in Table I Mean values characterize individual calcifications; Standard Deviations (SD) characterize the cluster. Table II summarizes the values for those features.

Table 1: Intensity histogram features

Feature Number assigned	Feature
1.	Mean
2.	Variance
3.	Skewness
4.	Kurtosis
5.	Entropy
6.	Energy

In this paper, the value obtained from our work for different type of image is given as follows:

Table 2: Intensity histogram features and their values

Image Type	Features					
	Mean	**Variance**	**Skewness**	**Kurtosis**	**Entropy**	**Energy**
normal	7.2534	1.6909	-1.4745	7.8097	0.2504	1.5152
malignant	6.8175	4.0981	-1.3672	4.7321	0.1904	1.5555
benign	5.6279	3.1830	-1.4769	4.9638	0.2682	1.5690

4.2GLCM features

It is a statistical method that considers the spatial relationship of pixels is the gray-level co-occurrence matrix (GLCM), also known as the gray-level spatial dependence matrix [24, 25]. By default, the spatial relationship is defined as the pixel of interest and the pixel to its immediate right (horizontally adjacent), but you can specify other spatial relationships between the two pixels. Each element (I, J) in the resultant GLCM is simply the sum of the number of times that the pixel with value I occurred in the specified spatial relationship to a pixel with value J in the input image. The formulae used for the metrics of the spatial gray level dependency matrix are as follows for the eleven features that are used in this study are as given below:

Contrast (CON):

$$CON = \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{\substack{i=1 \\ |i-j|=n}}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\} \quad (1)$$

Correlation (CORR):

$$CORR = \frac{\left[\sum_{i=1}^{N_g} \sum_{j=1}^{N_g} (i,j) p(i,j) \right] - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (2)$$

$$\mu_x = \sum_{i=1}^{N_g} \left[i \sum_{j=1}^{N_g} p(i,j) \right] \qquad \mu_y = \sum_{j=1}^{N_g} \left[j \sum_{i=1}^{N_g} p(i,j) \right]$$

$$\sigma_x = \sum_{i=1}^{N_g} \left[(i-\mu_x)^2 j \sum_{j=1}^{N_g} p(i,j) \right]$$

$$\sigma_y = \sum_{j=1}^{N_g} \left[(j-\mu_y)^2 i \sum_{i=1}^{N_g} p(i,j) \right]$$

Where μ_x ,μ_y are the mean values and σ_x ,σ_y are the standard deviations of P_X and P_y, respectively

Energy (ENER):

$$ENER = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g} P(i,j)^2 \quad (3)$$

Entropy (ENT):

$$ENT = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}[P(i,j)\log(P(i,j))] \quad (4)$$

Inverse difference moment (IDM):

$$IDM = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}\left[\frac{1}{1+(i-j)^2}P(i,j)\right] \quad (5)$$

Sum of Squares (SOS):

$$SOS = \sum_{i=1}^{N_g}\sum_{j=1}^{N_g}(i-\mu)^2\, p(i,j) \quad (6)$$

Sum average (SA):

$$SA = \sum_{i=2}^{2N_g}[i\ P_{x+y}(i)] \quad (7)$$

Sum Variance (SV):

$$SV = \sum_{i=2}^{2N_g}[(i-SA)^2 P_{x+y}(i)] \quad (8)$$

Sum Entropy (SE):

$$SE = -\sum_{i=2}^{2N_g}[P_{x+y}(i)\log[P_{x+y}(i)]] \quad (9)$$

Difference Variance (DV):

$$DV = \sum_{i=0}^{N_g-1}[(i-f')^2 P_{x-y}(i)] \quad (10)$$

Where

$$f' = \sum_{i=0}^{N_g-1}[i\,P_{x-y}(i)]$$

Difference Entropy (DE):

$$DE = -\sum_{i=2}^{N_g-1}[P_{x-y}(i)\log[P_{x-y}(i)]] \quad (11)$$

4.3Gray Level Run Length Method

Two kinds of methods are used for processing the grey level pixel-run length. In the first one, a vector considering pixel-runs is created from the function q(L, θ , T), in which L is length of the pixel-run (number of pixels in the pixel-run) while θ is direction of the pixel run and T, the threshold. Direction θ of pixel-run is defined similar to that in the GLCM method. Threshold value T for pixels to be merged into the pixel-run is given manually by the user. The procedure of constructing the pixel-runs is as follows: each pixel row of image at direction θ is scanned and the first pixel of the row is set to be the first pixel-run with length 1 and same grey value I as the first pixel; then the next pixel in the row is scanned; if *I I T n* | − | ≤ (In is the grey value of the next pixel), the next pixel is merged into the pixel-run, otherwise, a new pixel-run is created and the pointer is moved to the next pixel". This procedure is performed until the

scanning of the entire row is completed, and a new row is started [26]. Fig. 4(a) shows an image values and the pixel runs of similar values are build from an original image.

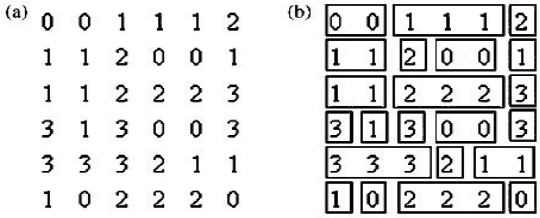


Fig. 4 Illustrations of building the pixel run lengths. (a) Initial input image; (b) Building the pixel run lengths with threshold T = 0 and direction θ = 0

In the GLRLM approach, the gray level runs are characterized by the gray tone of the run and the length of the run and the direction of the run [27]. "Let P(i, j) represent the run length matrix array. The matrix array consists of elements with the gray tone "i" has a run length "j". Textural features are calculated from the array elements that are used to study the nature of image textures. From the original run length matrix p(i, j), many numerical texture measures can be computed. The five original features of run length statistics derived by Galloway", [32] are as follows.

Short Run Emphasis (SRE):

$$SRE = \frac{1}{n_r} \sum_{j=1}^{N} \frac{P_r(j)}{j^2} \tag{12}$$

Long Run Emphasis (LRE):

Gray-Level Nonuniformity (GLN):

$$GLN = \frac{1}{n_r} \sum_{i=1}^{M} P_g(i)^2 \tag{13}$$

Run Percentage (RP):

$$RP = \frac{n_r}{n_p} \tag{14}$$

Run Length Nonuniformity (RLN):

$$RLN = \frac{1}{n_r} \sum_{j=1}^{N} P_r(i)^2 \tag{15}$$

Low Gray-Level Run Emphasis (LGRE):

$$LGRE = \frac{1}{n_r} \sum_{i=1}^{M} \frac{P_g(i)}{i^2} \tag{16}$$

High Gray-Level Run Emphasis (HGRE):

$$HGRE = \frac{1}{n_r} \sum_{i=1}^{M} P_g(i) . i^2 \tag{17}$$

"In the above equations, nr is the total number of runs and np is the number of pixels in the image. Based on the observation that most features are only functions of pr(j), without considering the gray level information contained in pg(i)", Chu et al. [28] proposed two new features, as follows, to extract gray level information in the matrix.

4.4Gray Level Difference Method

The run difference method is a generalized form of the GLDM, which is based on the estimation of the pdf of gray level differences in an image. GLDM seeks to extract texture features that describe the size and prominence of textural elements in an image. "Let I(x, y) be the image intensity function. For any given displacement δ = (, ΔY) let Iδ (x, y) = |I(x, y) - I(X + ΔX , Y + ΔY)|, and f(i|δ) be the probability density of Iδ (x, y). The value of f(i|δ) is obtained from the number of times Iδ (x, y) occurs for a given δ , i.e. f(i|δ) = P(Iδ (x, y) = i). If a texture is directional, the degree of spread of the values in f(i|δ) should vary with the direction of d, given that its magnitude is in the proper range. Thus, texture directionality can be analyzed by comparing spread measures of f(i|δ) for various directions of d. In the present

study, four possible forms of the vector d were considered: (0, d), (d, 0), (-d, d), and (-d, -d), with d being the inter pixel distance, each of which corresponds to a displacement in 00, 450, 900 and 1350 direction, respectively. From each of the density functions corresponding to one of the above-mentioned directions, five texture features were obtained" [29, 30, 31]:

$$I_{rgdif} = \sum_{\theta \in \Theta} I^{\theta}_{rgdif} \quad (18)$$

From which statistical measures are extracted from the distribution of gray level differences. Rather than extracting textural features directly from the matrix I, three characteristic vectors are calculated to define texture descriptors.
The distribution of gray level differences (DGD) vector is computed as follows:

$$DGD_j = \sum_{r=1}^{[S/2]} I_{grdif} \quad (19)$$

The distribution of the average gray level difference given r is represented by the DOD vector

$$DOD_r = \sum_{gdif=0}^{G-1} g_{dif} I_{rgdif} \quad (20)$$

and the distribution of the average distance given gdif is represented by the DAD vector

$$DAD_j = \sum_{r=1}^{[S/2]} rI_{rgdif} \quad (21)$$

Five features that describe the distribution of gray level differences are defined from these characteristic vectors:
Large difference emphasis (LDE), which measures the predominance of large gray level differences;

$$LDE = \sum_{j=0}^{n_g} DGD(j).\ln\left(\frac{K}{j}\right) \quad (22)$$

Where K is a constant

Sharpness (SHP), which measures the contrast and definition in an image;

$$SHP = \sum_{j=0}^{n_g} DGD(j).j^2 \quad (23)$$

SMG (Second Moment of DGD), which measures the variation of gray level differences;

$$SMG = \sum_{j=0}^{n_g} DGD(j)^2 \quad (24)$$

SMO (Second Moment of DOD), which measures the variation of average gray level differences;

$$SMO = \sum_{r=1}^{f_{max}} DOD(r)^2 \quad (25)$$

LDEL (long distance emphasis for large difference), which measures the prominence of large differences a long distance from each other.

$$LDEL = \sum_{j=0}^{n_g} DAD(j).j^2 \quad (26)$$

The Following GLCM, Gray Level Run Length Method , Gray Level Difference Method features were extracted in our research work:
Contrast, Correlation, Energy, Entropy, Inverse difference Moment, Sum of squares, Sum average, Sum variance, Sum entropy, Difference variance, Difference entropy. Short Run Emphasis, Long Run Emphasis, Gray level Nonuniformity, Run Percentage, Run length Nonuniformity, Low gray Level Run Emphasis, Low gray Level Run Emphasis , Large difference emphasis, Sharpness, Second Moment of DGD, Second Moment of DOD, Long distance emphasis for large difference. The value obtained for the above features from our work for a typical image is given in the following table III. Table IV. Table V.

Table 3: GLCM Features and values Extracted from Mammogram IMAGE (Malignant)

Feature No	Feature Name	Feature Values
1	Contrast	1.8927
2	Correlation	0.1592
3	Energy	0.1033
4	Entropy	2.6098

5	Inverse difference Moment	0.2863
6	Sum of squares	0.1973,
7	Sum average	44.9329
8	Sum variance	13.2626
9	Sum entropy	133.5676
10	Difference variance	1.8188
11	Difference entropy	1.8927

Table 4: Textural features calculated from the spatial gray level dependency matrices

Feature No	Feature Name	Feature Values
1	Short Run Emphasis	0.8989
2	Long Run Emphasis	159,4692
3	Gray level Nonuniformity	103/2133
4	Run Percentage	0.0409
5	Run length Nonuniformity	0.2863
6	Low gray Level Run Emphasis	157.7533,
7	Low gray Level Run Emphasis	48.9329

Table 5: Gray Level Difference Matrix Parameters

Feature No	Feature Name	Feature Values
1	Large difference emphasis	1.8927
2	Sharpness	15.9275
3	Second Moment of DGD	103.7837
4	Second Moment of DOD	260.9889
5	Long distance emphasis for large difference	286.7843

5. CLASSIFICATION

Associative classifier based on positive and negative association rules b. liu, w. hsu, and y. ma[37] proposed a framework, named associative classification, to integrate association rule mining and classification. The integration is done by focusing on mining a special subset of association rules whose consequent parts are restricted to the classification class labels, called "Class Association Rules" (CARs). This algorithm first generates all the association rules and then selects a small set of rules to form the classifiers. When predicting the class label for a coming sample, the best rule is chosen. It consists of two parts, a rule generator (called CBA-RG), which is based on algorithm Apriori for finding association rules and a classifier builder (called CBA-CB).

This classifier generates both positive and negative association rules and ranks them in terms of correlation coefficient. This set of rules is later used in the classification stage.

This categorizer is used to predict to which classes' new objects are attached. Given a new object, the classification process searches in this set of rules for those classes that are relevant to the object presented for classification. The set of positive and negative rules discovered are ordered by confidence and support.

ACN: Associative Classifier with Negative Rules

Some existing classifiers use negative rules for classification. They discover rules of the form a1^b1^c1->Yes and ~ (a1^b1^c1) -> Yes and a1^b1^c1->~Yes. Generally negative association rule mining is a difficult task and it's an ongoing research activity. In this classifier we consider a subset of rules that have at most one negated literal. So consider

a1^b1^c1->Y and a1^b1^~c1->Y but not a1^~b1^~c1->Y. Rules of this form are very important since it can express semantics like "If I have a playing partner and that partner

is not Robin, then I am going to enjoy sport" (because I have some problems with Robin),

Essence of our algorithm is that we only consider negated rules that arise naturally during

APriori rule mining process so that no extra overhead is needed. During APriori mining,

when we generate a Candidate A=a1^B=b1->Yes from two frequent ruleItems A=a1- >Yes and B=b1->Yes, we can generate two more ruleItems of the form A=a1^B=~b1->Yes and ~A=a1^B=b1->Yes which can have higher conf. &

sup than A=a1^B=b1->Yes. Support and confidence of the new 2 rules can easily be calculated based on already available information.supp (A=a1^~B=b1) = supp(A=a1) - supp(A=a1^B=b1)rulesup (A=a1^~B=b1) = rulesup(A=a1)-rulesup(A=a1^B=b1)conf (A=a1^~B=b1) = (rulesup(A=a1^~B=b1)/supp(A=a1^~B=b1))

Algorithm

P1=find_frequent_1p_itemsets(D)
N1=find_frequent_1n_itemsets(D)
For(k=2;Lk-1!=empty;k++)
PCk= candidates generated for level k
for each candidate generated for each literal on the candidate create a new negative rule by negating that literal
add this rule to NCk
calculate supports for each candidate of PCk
for each c in Ck update siblings of c in NCk
Lk=candidates in PCk that pass support threshold
Nk=candidates in NCk that pass support threshold
Return L=union of Lk union of Nk
Rule Ranking Criteria
• A rule ri is ranked higher than rj if
• Confidence(ri)>confidence(rj)
• Correlation(ri)>correlation(rj)
• Support(ri)>support(rj)
• Rulesize(ri)<rulesize(rj)
• If ri is positive & rj is negative
Database Coverage
Sort rules based on rule ranking criteria. For each rule taken in order if rule classifies at least one remaining training example correctly include that rule in classifier and delete those examples. If database in uncovered select majority class from remaining examples
Else select majority class from entire training set. Experimental Fact
Say, a rule A=a1^B=!b1 has confidence 80%. But a rule A=a1^B=b2 has confidence
100%. So this rule is selected and examples covered by this rule are removed. Now it can happen that the confidence of A=a1^B=!b1 has dropped so much that over remaining examples it is inaccurate because its previous high accuracy was largely due to the rule A=a1^B=b2. To remove this problem for negative rules, constraint has been adopted.
If a negative rule does not classify at least 55% of the remaining examples, it cannot be included.

More Pruning

All rules should be positively correlated.
So rules with correlation <0 are bad rules and they are pruned.
Rules with correlation greater than a threshold are good rules. We first try to cover database using these rules. But if database remains uncovered, then we take help of the rules that are positively correlated but correlation < threshold.
Experimentally set threshold = 0.2

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 [38, 39, 40, 41]. The results for the ten splits of the mammogram database are given in Table VI.

Table 6: Classification Accuracy For the Ten Splits with ANR

Splits	Classification Accuracy
1	90.95
2	97.89
3	96.56
4	97.76
5	95.98

6	94.59
7	94.78
8	92.94
9	95.09
10	97.69
Average	**95.47**

In this paper we used multi dimensional genetic association rule mining using image contents for the classification of mammograms. The average accuracy is 95.47 %. We have employed the freely available Machine Learning package, WEKA [42]. 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 V.

Table 7: Result obtained by the 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 VI.

Table 8: Confusion Matrix

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

7. CONCLUSIONS

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.

Classification of Microcalcification Clusters (MCs) is one the key to find the early sign of breast cancer. In this paper, we have proposed a novel association rule based system for classification of Microcalcification Clusters (MCs). Initially the MCs are segmented from the mammograms with region growing and the statistical GLCM, GLRLM, GLDM features are extracted. The proposed approach Classification by Associative Classifier with Negative Rules Using Texture Features is applied to construct the association rule to classify the images into three classes: normal, benign and malign. The result shows that this mrthod outperforms than the existing. In future, an efficient algorithm can be used to select the relevant features and the rules can be generated to improve the accuracy.

References

[1] Majid AS, de Paredes ES, Doherty RD, Sharma N Salvador X. “Missed breast carcinoma: pitfalls and Pearls”. Radiographics, pp.881-895, 2003.

[2] Osmar R. Zaïane,M-L. Antonie, A. Coman “Mammography Classification by Association Rule based Classifier,” MDM/KDD2002 International Workshop on Multimedia Data Mining ACM SIGKDD, pp.62-69,2002,

[3] Xie Xuanyang, Gong Yuchang, Wan Shouhong, Li Xi ,”Computer Aided Detection of SARS Based on Radiographs Data Mining “, Proceedings of the 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference Shanghai, China, pp7459 – 7462, 2005.

[4] C.Chen and G.Lee, "Image segmentation using multitiresolution wavelet analysis and Expectation Maximum(EM) algorithm for mammography" , International Journal of Imaging System and Technology, 8(5): pp491-504,1997.
[5] T.Wang and N.Karayaiannis, "Detection of microcalcification in digital mammograms using wavelets", IEEE Trans. Medical Imaging, 17(4):498-509, 1998.
[6] Jelena Bozek, Mario Mustra, Kresimir Delac, and Mislav Grgic "A Survey of Image Processing Algorithms in Digital mammography"Grgic et al. (Eds.): Rec. Advan. in Mult. Sig. Process. and Commun., SCI 231, pp. 631–657,2009
[7] Shuyan Wang, Mingquan Zhou and Guohua Geng, "Application of Fuzzy Cluster analysis for Medical Image Data Mining" Proceedings of the IEEE International Conference on Mechatronics & Automation Niagara Falls, Canada,pp. 36 – 41, 2005.
[8] Jensen, Qiang Shen, "Semantics Preserving Dimensionality Reduction: Rough and Fuzzy-Rough Based Approaches", IEEE Transactions on Knowledge and Data Engineering, pp. 1457-1471, 2004.
[9] I.Christiyanni et al ., "Fast detection of masses in computer aided mammography", IEEE Signal processing Magazine, pp:54- 64,2000.
[10] K. Thangavel , A. Kaja Mohideen "Classification of Microcalcifications Using Multi-Dimensional Genetic Association Rule Miner" International Journal of Recent Trends in Engineering, Vol 2, No. 2, pp. 233 – 235, 2009
[11] R.Agrawal, T. Imielinski, and A.Swami. Mining association rules between sets of items in large databases. In the Proc. of the ACM SIGMOD Int'l Conf. on Management of Data (ACM SIGMOD '93), Washington, USA, May 1993.
[12] Alex A. Freitas, "A Survey of Evolutionary Algorithms for Data Mining and Knowledge Discovery" Postgraduate Program in Computer Science, Pontificia Universidade Catolica do Parana Rua Imaculada Conceicao, 1155. Curitiba - PR. 80215-901. Brazil.
[13] Etta D. Pisano, Elodia B. Cole Bradley, M. Hemminger, Martin J. Yaffe, Stephen R. Aylward, Andrew D. A. Maidment, R. Eugene Johnston, Mark B. Williams,Loren T. Niklason, Emily F. Conant, Laurie L. Fajardo,Daniel B. Kopans, Marylee E. Brown • Stephen M. Pizer "Image Processing Algorithms for Digital Mammography: A Pictorial Essay" journal of Radio Graphics Volume 20,Number 5,sept.2000
[14] Pisano ED, Gatsonis C, Hendrick E et al. "Diagnostic performance of digital versus film mammography for breast-cancer screening". NEngl J Med 2005; 353(17):1773-83.
[15] Wanga X, Wong BS, Guan TC. 'Image enhancement for radiography inspection". International Conference on Experimental Mechanics. 2004: 462-8.
[16] Suzuki h.torkakij (1991) automatic segmentation of head MRI images by knowledge guided thresholding,computer medical imaging graphic15(4);233.
[17] Dr.samir kumar bandhyopathyway,tuhin utsab paul,segmentation of brian MRI imges research in computer science and software engineering,volume 2,issue 3,march2012,issn;2277 128x.
[18] M. Petrou and P. Bosdogianni, Image Processing the Fundamentals, Wiley, UK, 2004.
[19] R. C. Gonzalez and R.E. Woods, Digital Image Processing 2nd Edition, Prentice Hall, New Jersey, 2002.
[20] Jian-Jiun Ding, The class of "Advanced Digital Signal Processing", the Department of Electrical Engineering, National Taiwan University (NTU), Taipei, Taiwan, 2008.
[21] Patel, D., Hannah, I., & Davies, E. R, "Foreign object detection via texture analysis", 12th IAPR international conference on pattern recognition Proceeding: Vol. 1. Conference A: Computer vision and image processing, 1994.
[22] G. N. Srinivasan, and Shobha G, "Statistical Texture Analysis",Proceedings of World Academy of Science, Engineering and Technology Volume 36 December 2008 ISSN 2070-3740.
[23] Li Liu, Jian Wang and Kai He "Breast density classification using histogram moments of multiple resolution mammograms" Biomedical Engineering and Informatics (BMEI), 3rd International Conference, IEEE explore pp.146–149, DOI: November 2010, 10.1109/ BMEI.2010 .5639662
[24] Li Ke,Nannan Mu,Yan Kang Mass computer-aided diagnosis method in mammogram based on texture features, Biomedical Engineering and Informatics (BMEI), 3rd International Conference, IEEE Explore, pp.146 – 149, November 2010, DOI: 10.1109/ BMEI.2010.5639662,
[25] Azlindawaty Mohd Khuzi, R. Besar and W. M. D. Wan Zaki "Texture Features Selection for Masses Detection In Digital Mammogram" 4th Kuala Lumpur International Conference on Biomedical Engineering 2008 IFMBE Proceedings, 2008, Volume 21, Part 3, Part 8, 629-632, DOI: 10.1007/978-3-540-69139-6_157
[26] F. R. Renzetti, L. Zortea, "Use of a gray level co-occurrence matrix to characterize duplex stainless steel phases microstructure", Fratturaed Integrità Strutturale, 16 (2011) 43-51.
[27] Xiaoou Tang, "Texture Information in Run-Length Matrices", IEEE Transactions On Image Processing, Vol. 7, No. 11, November 1998.

[28] Chu, C. M. Sehgal, and J. F. Greenleaf, "Use of gray value distribution of run lengths for texture analysis", Pattern Recognit. Lett, Vol. 11, pp. 415–420. June 1990.
[29] Shoshana Rosskamm, "Computer Aided Diagnosis of Cystic Fibrosis and Pulmonary Sarcoidosis using Texture Descriptors Extracted from CT Images", thesis for the Master of Science degree of Applied Mathematics 2010.
[30] Stavroula G. Mougiakakou et al, "Differential diagnosis of CT focal liver lesions using texture features, feature selection and ensemble driven classifiers", Elsevier Artificial Intelligence in Medicine (2007) 41, 25—37.
[31] Wei-Ming Chen et al., "3-D Ultrasound Texture Classification using Run Difference Matrix", Elsevier Ultrasound in Med. & Biol., Vol. 31, No. 6, pp. 763–770, 2005.
[32] Galloway M,"Texture analysis using gray level run lengths", Comp Graph Im Proc 1975; 4:172-9.
[33] Yvan Saeys, Thomas Abeel, Yves Van de Peer "Towards robust feature selection techniques", www.bioinformatics.psb.ugent
[34] Gianluca Bontempi, Benjamin Haibe-Kains "Feature selection methods for mining bioinformatics data", http://www.ulb.ac.be/di/mlg
[35] Dougherty J, Kohavi R, Sahami M. "Supervised and unsupervised discretization of continuous features". In: Proceedings of the 12th international conference on machine learning.San Francisco:Morgan Kaufmann; pp 194–202, 1995
[36] D.Brazokovic and M.Nescovic, "Mammogram screening using multisolution based image segmentation", International journal of pattern recognition and Artificial Intelligence, 7(6): pp.1437-1460, 1993
[37] B. Liu, W. Hsu, and Y. Ma. Integrating classification and association rule mining. InKDD'98, pp. 80-86, New York, NY, Aug. 1998
[38] J Hipp, U Güntzer, and G Nakhaeizadeh, "Algorithms for association rule mining—a general survey and comparison", vol. 2, no. 1, 2000.
[39] Jiawei Han and Micheline Kamber, "Data Mining, Concepts and Techniques". Morgan Kaufmann, 2001.
[40] 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
[41] 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
[42] 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.