

Feature Extraction of Mammograms

Monika Sharma¹, R. B. Dubey², Sujata³, S. K. Gupta⁴

ECE dept., Hindu College of Engineering, Sonapat, India^{1,2}, EE Department, DCRUST, Murthal, Sonapat, India³
Vaish College of Engineering, Rohtak, India⁴

Abstract

Breast cancer is the second leading cause of cancer deaths in women today. Early detection of the cancer can reduce mortality rate. Studies have shown that radiologists can miss the detection of a significant proportion of abnormalities in addition to having high rates of false positives. Pattern recognition in image processing requires the extraction of features from regions of the image and the processing of these features with a pattern recognition algorithm. We consider the feature extraction part of this processing; with a focus on the problem of micro calcification detection in digital mammography. For every pattern classification problem, the most important stage is feature extraction. The accuracy of the classification depends on the feature extraction stage. We have extracted textural, statistical and structural features which show promising results than most of the existing technology.

Keywords

Mammograms, features extraction, GLCM, Textural, Statistical features.

1. Introduction

Breast cancer is one of the main causes of death for woman between ages of 30 to 35. This mortality can be reduced by only early detection and prevention of breast cancer. As per opinion of radiologist micro calcification can be the only mammographic sign of non-palpable breast disease. Mammography is an effective tool for early detection because in many cases it can detect abnormalities such as masses, calcification and other suspicious anomalies up to two years before they are palpable [2].

Karssemeijer [11–13] introduced a statistical method for features extraction in digital mammograms. The method is based on the use of statistical models and the general framework of Bayesian image analysis. Chan et al. [14–16] proposed a computer-based method for the detection of micro calcification in digital mammograms. The method is based on a

difference image technique in which a signal suppressed image is subtracted from a signal enhanced image to remove structured background in the mammogram.

F. Andrew et al. suggested an approach for mammographic feature analysis by over complete multi resolution representations. Mammograms are reconstructed from wavelet coefficients modified at one or more levels by local and global nonlinear operators. In each case, edges and gain parameters are identified adaptively by a measure of energy within each level of scale-space.

Timp et al present a fully automated computer-aided diagnosis program to detect temporal changes in mammographic masses between two consecutive screenings rounds. The goal was to improve the characterization of mass lesions by adding information about the tumour behaviour over time. Jiang et al. [4] have introduced an automated computer scheme that can classify clustered micro-calcifications (MCs) more accurately than radiologists. When it was used as a diagnostic aid, this scheme could lead to significant improvement in radiologist's performance in distinguishing between malignant and benign clustered MCs.

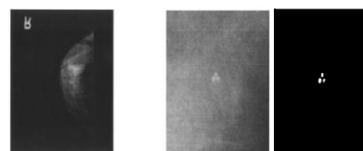


Fig.1: (a) An original mammogram image (b) A cluster of micro calcifications in (a). (c) The ground truth image of (b).

Wu et al. and Lo et al. have independently investigated CADx schemes that employ artificial neural networks where the image features were extracted by the radiologist. Linear discriminant analysis was used by Chan et al. for classification of MCs [6, 11, 12]. Dhawan developed a radial basis function neural network for classifying hard-to-diagnose cases of MCs. Neural networks; Bayes

classifiers and nearest-neighbour method were used for classifying MCs. [13-15].

The rest of this paper is organized as follows. In Section 2, pre-processing stage is formulated. Section 3 deals with the various features. Section 4 describe segmentation and feature extraction. The experimental results are evaluated and discussed in Section 5. Finally, the concluding remarks are given in Section 6.

2. Pre-processing stage

The pre-processing phase of digital mammograms refers to the enhancement of mammograms intensity and contrast manipulation, noise reduction, background removal, edges sharpening, filtering, etc.

2.1. Enhancing image contrast

Histogram equalization technique is used for enhancing the contrast of an image. Its basic idea lies on mapping the gray levels based on the probability distribution of the input gray levels. It flattens and stretches the dynamic range of image histogram and resulting in overall contrast improvement.

2.2. Edge filtering

In diagnosis, the radiologists mainly use their eyes to observe abnormalities occur on the mammograms. However, in many cases, cancer is not easily detected by the eyes because of bad imaging conditions. In order to improve the correct diagnosis rate of cancer, image enhancement technology is often used to enhance the image quality and assist the radiologists. Here, firstly we modify the brightness and contrast to enhance images for better visual quality. Second, the image enhancement is convoluted with mask of edge detectors and thus it can increase value of edge pixel in image building edge image for continued process. Third, the best edge image among several edge images produced by the used edge detectors in this system is automatically visualized based on single parameter.

Edge detection is a kind of method of image segmentation based on range non-continuity. Edge always appears in two neighbouring areas having different grey level. It exists between object and background, object and object, region and region, and between element and element. We have used edge detectors such as Differential operator, Roberts's operator, Sobel operator for convolution the mammograms.

2.2.1. Differential operator: Suppose that the image is $f(x, y)$, and its derivative operator is the first order partial derivative $\partial f/\partial x, \partial f/\partial y$. The vector with direction and modulus from $f(x, y)$ is called as the gradient of the function, that is

$$F(x,y) = \frac{\delta f}{\delta x} = G_x/G_y \quad (1)$$

$$\text{where } G_x = \frac{\delta F(x,y)}{\delta x} = \frac{\{f(x+\Delta x,y)-f(x,y)\}}{\Delta x} \quad (2)$$

$$G_y = \frac{\delta F(x,y)}{\delta y} = \frac{\{f(x+\Delta x,y)-f(x,y)\}}{\Delta y}$$

The gradient modulus operator is defined as

$$G [f(x, y)] = \sqrt{(G_x)^2 + (G_y)^2} = G_x + G_y \quad (3)$$

The direction derivative of function $f(x, y)$, has a maximum at a certain point and the direction of this point is

$$\alpha(X, y) = \tan^{-1}(G_x/G_y) \quad (4)$$

Here, Differential operator mostly includes Roberts's operator and Sobel operator.

2.2.2. Robert operator: Roberts's operator makes use of partial difference operator to look for edge.

$$R_+ (x, y) = f(x+1, y+1) - f(x, y)$$

$$R_- (x, y) = f(x, y+1) - f(x, y)$$

$$R_+ = \begin{bmatrix} 1 & 0 \\ 0 & -1 \end{bmatrix}$$

$$R_- = \begin{bmatrix} 0 & 1 \\ -1 & 0 \end{bmatrix}$$

The direction of this point is

$$\alpha(x, y) = \frac{\pi}{4} + \tan^{-1}(R_-/R_+) \quad (5)$$

The gradient modulus operator is given by eqn. (6).

$$G [f(x, y)] = \sqrt{(R_x)^2 + (R_y)^2} = R_x + R_y \quad (6)$$

2.2.3. Sobel operator: The Sobel operator counts difference using weighted for 4 neighbourhoods on the basis of the Prewitt operator. The Sobel operator has the similar function as the Prewitt operator, but the edge detected by the Sobel operator is wider. Sobel operator can process those images with lots of noises and gray gradient well.

$$S_x = (a_2 + ca_3 + a_4) - (a_0 + ca_7 + a_6)$$

$$S_y = (a_0 + ca_1 + a_2) - (a_6 + ca_5 + a_4)$$

With $c=2$, S_x and S_y can be declared as

$$S_x = \begin{bmatrix} -1 & 0 & 1 \\ -2 & 0 & 2 \\ -1 & 0 & 1 \end{bmatrix}$$

$$S_y = \begin{pmatrix} 1 & 2 & 1 \\ 0 & 0 & 0 \\ -1 & -2 & -1 \end{pmatrix}$$

$$M = \sqrt{(S_x)^2 + (S_y)^2}$$

The direction of this point is
 $\alpha(x, y) = \tan^{-1}(S_x / S_y)$ (11)

2.3. Image Orientation

The image is rotated and reflected to adjust the pectoral muscle on the left side of the image [14].

2.4. Normalization of local contrast

Local contrast c_i at site i can be defined by

$$c_i = y_j - 1/N \sum_{j \in \partial i} y_j$$

(12)
 With the pixel value at site i and ∂i a neighborhood or window at i of size N . The standard deviation of local contrast σ_c is determined as a function of gray level from the image at hand and from this information the normalization is performed. After determination of local $\sigma_c(y)$, the contrast c_i is normalized by

$$c'_j = c_j / \sigma_c(y_j)$$

(13)
 Where c_i represents normalized local contrast at site i . We used bins of variable size where the term size is the number of pixels in a bin. Bin size is large at lower brightness whereas at highest brightness values decreases rapidly [2].

The statistical model is based on the use of Bayesian techniques applications of a Markov random field model, where the latter models the fact that micro calcifications occur in clusters. During the detection process pixel labels x_i are iteratively updated by maximizing their probability, given the image data in a small neighborhood $y_{\partial i}$ of site i and given the current estimate of the rest of the labeling.

$$X'_{j=\max} [p(x_j = l | y_{\partial j}, X_{\bar{j}})]$$

(14) Where $l = 1, 2, 3, 4$ represent four pixel classes such as background, micro calcifications, lines/edge and film emulsion errors. The probability to be maximized can be written as

$$x_j = l / y_{\partial j}, X_{\bar{j}} \quad x_j = l, X_{S/i} p \left(x_j = \frac{l}{X_{\bar{j}}} \right)$$

(15)
 Where Θ_i is a vector denoting the values of the three features at particular site. The a priori probability $p(x_i | X_{\bar{i}})$ of the labels represents the Markov random field and models spatial relations [1].

3. Different features

3.1. Gray level co-occurrence matrix (GLCM)

Mathematically, a co-occurrence matrix C is defined over an $n \times m$ image I , parameterized by an offset $(\Delta x, \Delta y)$, as

$$C_{\Delta x, \Delta y}(i, j) = \sum_{p=1}^n \sum_{q=1}^m \begin{cases} 1, & \text{if } I(p, q) = i \text{ and } I(p + \Delta x, q + \Delta y) = j \\ 0, & \text{otherwise} \end{cases}$$

(16)
 The co-occurrence matrix is often formed using a set of offsets sweeping through 180 degrees (i.e. 0, 45, 90 and 135 degrees) at the same distance to achieve a degree of rotational invariance. After making the GLCM symmetrical, there is still one step to take before texture measures can be calculated. The measures require that each GLCM cell contain not a count, but rather a probability. The normalization equation is

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j} V_{i,j}}$$

(17)
 where i is the row number and j is the column number. i and j keep track of cells by their horizontal and vertical coordinates.

Calculation of texture measures: Generally texture calculations are weighted averages of the normalized GLCM cell contents. A weighted average multiplies each value to be used by a factor before summing and dividing by the number of values.

Creating a texture image: The result of a texture calculation is a single number representing the entire window. This number is put in the place of the centre pixel of the window, then the window is moved one pixel and the process is repeated of calculating a new GLCM and a new texture measure. In this way an entire image is built up of texture values.

Edge of image problems: Each cell in a window must sit over an occupied image cell. This means that the centre pixel of the window cannot be an edge pixel of the image. Image edge pixels usually represent a very small fraction of total image pixels, so this is only a minor problem. However, if the image is very small or the window is very large, the image edge effect should be remembered when examining the texture image. Edge effects can be a problem in classification.

Measures related to contrast: Values on the GLCM diagonal show no contrast, and contrast increases away from the diagonal. So, create a weight that

increases as distance from the diagonal increases. Contrast is also called sum of squares variance.

$$\sum_{i,j=0}^{N-1} P_{i,j} (i-j)^2 \quad (18)$$

When i and j are equal, the cell is on the diagonal and (i-j)=0.

Dissimilarity: In the contrast measure, weights increase exponentially (0, 1, 4, 9, etc.) as one moves away from the diagonal. However in the dissimilarity measure weights increase linearly (0, 1, 2, 3 etc.). Dissimilarity equation is

$$\sum_{i,j=0}^{N-1} P_{i,j} |i-j| \quad (19)$$

Homogeneity: It is also called the inverse difference moment. If weights decrease away from the diagonal, the result will be larger for windows with little contrast. Homogeneity equation is

$$\sum_{i,j=0}^{N-1} \frac{P_{i,j}}{1+(i-j)^2} \quad (20)$$

Angular second moment (ASM) and energy

ASM and Energy use each P_{ij} as a weight for itself. High values of ASM or Energy occur when the window is very orderly. ASM equation is

$$\sum_{i,j=0}^{N-1} P_{i,j}^2 \quad (21)$$

The square root of the ASM is sometimes used as a texture measure, and is called energy. Energy equation is

$$\text{Energy} = \sqrt{\text{ASM}} \quad (22)$$

The various implemented GLCM texture measure are shown in Fig. 2.

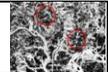
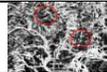
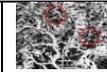
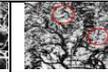
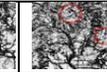
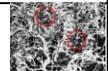
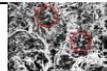
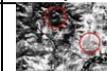
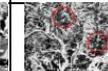
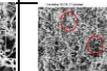
				
Contrast	Dissimilarity using GLCM	GLDV Mean, which is the same as GLCM Dissimilarity	Homogeneity	ASM
				
GLCM Entropy	GLDV Entropy	GLCM Mean	GLCM Standard Deviation	GLCM Correlation

Fig 2: GLCM textures implementation

3.2. Textural, Statistical and Structural features

1. Mean: The mean is calculated using the formula:

$$M = 1/MN \sum_{i=1}^M \sum_{j=1}^N (P(i,j) - \mu)^2 \quad (23)$$

Where p (i, j) is the pixel value at point (i, j) of an image of size M x N.

2. Standard Deviation: The standard deviation, σ is the estimate of the mean square deviation of gray pixel value p (i, j) from its mean value. It is determined using the formula:

$$\sigma = \sqrt{1/MN \sum_{i=1}^M \sum_{j=1}^N (p(i,j) - \mu)^2} \quad (24)$$

3. Smoothness: Relative smoothness, R is a measure of gray level contrast that can be used to establish descriptors of relative smoothness. The smoothness is determined by

$$R = 1 - 1/1 + \sigma^2 \quad (25)$$

Where, σ is the standard deviation of the image.

4. Entropy: Entropy is a measure of the uncertainty associated with a random variable. Entropy in an information sense is a measure of unpredictability. It is given by

$$f_s = - \sum_i \sum_j p(i,j) \log(p(i,j)) \quad (26)$$

5. Skewness: Skewness, S characterizes the degree of asymmetry of a pixel distribution in the specified window around its mean. Skewness is a pure number that characterizes only the shape of the distribution.

The formula for finding skewness is given in the below

$$S = 1/MN \sum_{i=1}^M \sum_{j=1}^N [p(i,j) - \mu]^3 / \sigma^3 \quad (27)$$

Where, p (i, j) is the pixel value at point (i, j), μ and σ are the mean and standard deviation respectively.

6. Kurtosis: Kurtosis, K measures the peakness or flatness of a distribution relative to a normal distribution. The conventional definition of kurtosis is

$$k = 1/MN \sum_{i=1}^M \sum_{j=1}^N [p(i,j) - \mu]^4 / \sigma^4 - 3 \quad (28)$$

Where, p (i, j) is the pixel value at point (i, j), μ and σ are the mean and standard deviation respectively.

7. Root Mean Square (RMS): It computes the RMS value of each row or column of the input, along vectors of a specified dimension of the input, or of the entire input. The RMS value of the jth column of an MxN input matrix u is given by

$$\sqrt{\sum_{i=1}^M |\mu_{ij}|^2} / M \quad (29)$$

8. Inverse Difference Moment (IDM): It is a measure of image texture. IDM ranges from 0.0 for an image that is highly textured to 1.0 for an image that is untextured. The formula for finding the IDM is

$$H = \sum_{i,j} \frac{p(i,j)}{1+|i-j|} \quad (30)$$

9. Energy: Energy is used to describe a measure of information when formulating an operation under a probability framework such as MAP (maximum a priori) estimation in conjunction with Markov Random Fields [45]. Sometimes the energy can be a negative measure to be minimized and sometimes it is a positive measure to be maximized. It is given by

$$f_7 = \sum_i \sum_j p(i, j)^2 \quad (31)$$

10. Contrast: Contrast is the difference between the light and dark areas of a picture, such as a photograph or video image [41]. Contrast also affects our ability to see details in an image. It is given by

$$f_2 \sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i, j) |i-j|=n \right\} \quad (32)$$

11. Correlation: Correlation is basic operation that we will perform to extract information from images [42]. It is given by

$$f_3 = \frac{\sum_i \sum_j (ij) p(i, j) - \mu_x \mu_y}{\sigma_x \sigma_y} \quad (33)$$

12. Homogeneity: Its formula is

$$f_9 = \sum_i \sum_j \frac{1}{1+(i-j)^2} p(i, j) \quad (34)$$

13. Variance: Variance map of an image is calculated by taking a square window of a set size around a centre pixel and is given by

$$f_{11} = \sum_i \sum_j (i-j)^2 p(i, j) \quad (35)$$

4. Segmentation and Feature Extraction

4.1. Segmentation

Shape and contrast features of micro calcification are often used in schemes for automated differentiation between true positive and false positive detected micro calcifications. In determining such features, segmentation plays an important role and influences classification and thereby detection performance. By segmentation we mean here determination of the precise outline of micro calcification. To perform an accurate segmentation on detected micro calcifications a background trend correction is applied. We assume that a detection step has been carried out so that the positions of the micro calcifications are known. To define a background area a disc with a diameter of 1 mm is use. Micro calcifications will be covered by this disc. The centre of the disc coincides with the centre of the data field. The pixel values in the disc are area R1 replaced by new pixel values interpolated from the surrounding

background. Pixel value y_i with $i \in R$ is replaced by y_i according to the following weight function

$$y'_j = \frac{\sum_{k \in B} \frac{y_k}{d_{ik}}}{\sum_{k \in B} \frac{1}{d_{ik}}} \quad (36)$$

Where d_{ik} is the Euclidean distance between site i and site k with k a pixel on the boundary L of R . The background image is low-pass filtered using a 5 x 5 uniform kernel and then subtracted from the original image. In the second stage, the segmentation is performed on the resulting image. This image is also used for calculating contrast features of the object. A Gaussian model is used for representing the fluctuation of gray levels due to noise.

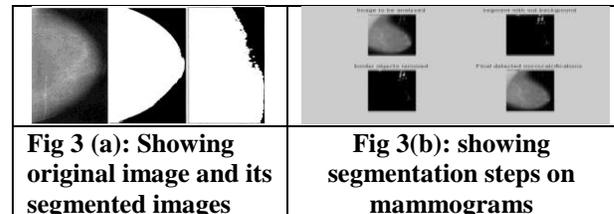


Fig 3 (a): Showing original image and its segmented images

Fig 3(b): showing segmentation steps on mammograms

4.2. Features Extraction

Features are nothing but observable patterns in the image which gives information about the image. The accuracy of the classification depends on the feature extraction stage. Here, we will calculate texture, statistical and structural features. Feature is used to denote a piece of information which is relevant for solving the computational task related to a certain application. More specifically, features can refer to the result of a general neighbourhood operation applied to the image, specific structures in the image itself, ranging from simple structures such as points or edges to more complex structures such as objects. Many features have been extracted for the abnormalities of mammograms. The extraction methods of texture feature play very important role in detecting abnormalities of mammograms because of the nature of mammograms. Texture features have been proven to be useful in differentiating masses and normal breast tissues. Texture features are able to isolate normal and abnormal lesion with masses and micro calcification. The features that are extracted are as under.

4.2.1. Individual micro calcification features

The following local features are considered.

Perimeter, defined as the number of pixel sides that touch a background pixel.

Area represented by the number of micro calcification pixels.

Compactness, defined as $c = (\text{perimeter})^2 / 4\pi \cdot \text{area}$.

Eccentricity, defined as

$$e = \frac{l_{xx} + l_{yy} - \sqrt{(l_{xx} - l_{yy})^2 + 4l_{xy}^2}}{l_{xx} + l_{yy} + \sqrt{(l_{xx} - l_{yy})^2 + 4l_{xy}^2}} \quad (37)$$

where l_{xx} , l_{xy} and l_{yy} are the moments of inertia.

Thickness, calculated as the width of the best fitting rectangle.

Orientation defined as the angle of axis of the least moment of inertia with respect to the xy-plane.

Direction, calculated as the relative direction in which the micro calcification is located viewed from its cluster's gravity centre.

Line, the mean of the output of the line/edge detector in a detected micro calcification.

Background, the mean intensity level of the background.

Foreground, the mean intensity of the detected micro calcification.

Distance, the distance to the closest neighbour calcification.

The contrast measure for a micro calcification pixel at site i depends on micro calcification thickness d_{mc} an estimate of the linear attenuation coefficient of micro calcification μ_{mc} and background tissue μ_b , the film curve gradient and the digitization constant c_0 . The expressions of the micro calcification pixels contrast is given by [10]:

$$C_i^{mc} = y_j - y_b = c_0 c_1 (\log e) d_{mc} (\mu_{mc} - \mu_b) \text{ for } y_j \in R \quad (38)$$

from equation it is clear that using proper scaling, micro calcification pixel contrast is approximately independent of breast thickness, and exposure level.

4.2.2. Cluster features

We only calculated the standard deviations for the orientation and direction features. These angles depend on the orientation of the breast in the xy-plane. The standard deviation angles, however gives information that is invariant for breast positioning.

4.2.3. GLCM for feature extraction

GLCM calculates the probability of a pixel with the gray level i occurring in a specific spatial relationship to a pixel with the value j [9]. The number of gray levels in the image determines the size of the GLCM. GLCM calculated in 4 angles (0, 45, 90, 135) and 4 distances (1, 2, 3, 4). The 18 descriptors extracted from GLCM texture measurement including autocorrelation, contrast, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, maximum probability, sum of squares: variance, sum average, sum variance, sum entropy, difference entropy, information measure of correlation 1, information measure of correlation 2 and inverse difference moment normalized. The

features for normal and abnormal patterns are shown in table 1.

5. Results

The results obtained after the feature extraction algorithms for 8 mammogram samples are presented here. The comparison tables have been made to categorise the mammograms into defected and non-defected stages. The various features extracted are energy, entropy, contrast, variance, homogeneity, correlation sum average, sum entropy, sum variance, difference variance, difference entropy, correlation, autocorrelation, cluster shade, cluster prominence and dissimilarity.

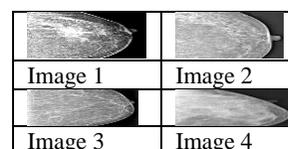


Fig 4: Image read from database

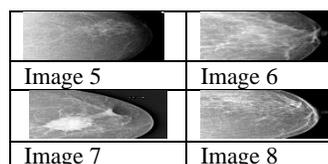


Fig. 5: Image read from database

Table 1

Feature no	GLCM Features	Normal	Abnormal
1	Autocorrelation	9.1504	7.4782
2	Contrast	0.7890	0.1363
3	Correlation	0.5613	0.9642
4	Cluster Prominence	22.9289	69.4536
5	Cluster Shade	-2.9112	3.0820
6	Dissimilarity	0.4118	0.0998
7	Energy	0.1890	0.3118
8	Entropy	2.0868	1.5162
9	Homogeneity	0.8471	0.9544
10	Maximum probability	0.3368	0.4557
11	Sum of squares: Variance	9.4439	7.5398
12	Sum average	5.8807	4.7554

13	Sum variance	20.8125	18.5144
14	Sum entropy	1.6375	1.4269
15	Difference entropy	0.8260	0.3289
16	Information measure of correlation	-0.3423	-0.7440
17	Information measure of correlation	0.7660	0.9150
18	Inverse difference moment normalized	0.9887	0.9980

Table 2: Extracted features

features	Image 1	Image 2	Image 3	Image 4
autocorrelation	5.0707	6.001	6.049	6.363
contrast	3.873	1.398	4.500	3.121
correlation	9.760	7.109	9.069	2.257
correlation probability	9.760	7.109	9.069	2.257
cluster prominence	3.143	1.127	1.487	1.809
cluster shade	-2.592	-9.50	-1.12	-2.37
dissimilarity	5.534	1.997	6.429	4.459
Energy	6.627	8.735	8.921	9.871
Entropy	5.541	3.092	2.491	4.385
homogeneity	9.930	9.750	9.919	9.944
homogeneity probability	9.922	9.720	9.909	9.937
maximum probability	7.881	9.336	9.433	9.935
Sum of squares: Variance	5.084	6.048	6.047	6.354
sum average	1.308	1.527	1.527	1.595
sum variance	1.891	2.327	2.350	2.536
sum entropy	5.501	2.894	2.428	3.943
difference variance	3.873	1.398	4.500	3.121
difference entropy	4.613	1.296	5.222	3.856
Information measure of correlation-1	-9.102	-4.88	-7.81	-4.60
Information measure of correlation-2	7.798	4.257	5.232	1.428
difference entropy	9.963	9.8666	9.957	9.970
Information measure of correlation-1	9.965	9.876	9.960	9.972

Table 3: Extracted features

Features	Image 5	Image 6	Image 7	Image 8
autocorrelation	5.548	5.558	6.293	5.062
contrast	2.439	2.727	6.357	3.572
correlation	9.785	9.757	4.437	9.780
correlation probability	9.785	9.757	4.437	9.780
cluster prominence	2.845	2.826	2.230	3.148
cluster shade	-2.28	-2.26	-1.82	-2.59
Dissimilarity	3.485	3.896	9.08	5.102
Energy	7.640	7.661	9.638	6.616
Entropy	4.209	4.202	1.11	5.537
homogeneity	9.956	9.951	9.886	9.936
homogeneity probability	9.951	9.945	9.872	9.928
maximum probability	8.642	8.658	9.817	7.869
Sum of squares: Variance	5.548	5.56	6.301	5.072
sum average	1.413	1.416	1.583	1.306
sum variance	2.105	2.109	2.491	1.887
sum entropy	4.185	4.174	1.023	5.497
difference variance	2.439	2.727	6.357	3.572
difference entropy	3.136	3.445	6.926	4.313
Information measure of correlation-1	-9.20	-9.12	-2.63	-9.16
Information measure of correlation-2	7.192	7.149	1.824	7.819
inverse differential information measure	9.976	9.974	9.939	9.965
inverse difference moment normalized	9.978	9.975	9.943	9.968

6. Conclusions

A frame work for detection of micro calcification based on prior knowledge on the GLCM has been presented. Extractions of features from different domains are studied. Image derived features such as measures of spectral and spatial features provide useful information for micro calcification detection. Spatial and spectral features are not independent; rather these features exist simultaneously in the image. The algorithm extracts the micro calcification effectively and gives the best diagnosis. We have

extracted textural, statistical and structural features which show promising results than most of the existing technology.

References

- [1] American Cancer Society, "Breast cancer facts and figures 2007-2008", Atlanta, Georgia: American Cancer Society, Inc.2007.
- [2] American Cancer Society, "Cancer prevention and Early Detection Facts and Figures", Atlanta, GA: American Cancer Society, 2003.
- [3] X. Liu and D. Wang, "Image and Texture Segmentation Using Local Histograms", IEEE Trans. Med. Img., vol.15, pp. 3066-3076, 2006.
- [4] V. Oktem and I. Jouny, "Automatic Detection of Malignant Tumours in Mammograms," Proceedings of the 26th Annual International Conference of the IEEE EMBS, pp. 1770-1773, 2004.
- [5] G. N. Srinivasan, and G. Shobha, "Statistical Texture Analysis", Proc. World Academy of Science, Engineering and technology, vol.. 36, pp. 1264-1268, 2008.
- [6] S. Timp, C. Varela and N. Karssemeijer, "Temporal Change Analysis for Characterization of Mass Lesions in Mammograms," IEEE Trans. Med. Img., vol.26, pp. 945-953, 2007.
- [7] M. P. Sampat, M. K. Markey and A. C. Bovik, "Computer- Aided Detection and Diagnosis in Mammography," Elsevier Academic Press, 2005.
- [8] C-M Tiu, T-L Jong and C-W Hsieh, "Self-organizing map neural network with fuzzy screening for micro-calcifications detection on mammograms", IEEE Conference on Soft Computing in Industrial Applications, pp. 421-425, May 2009.
- [9] T. A. Doccusse, A. S. Pereira and N. Marranghello, "Micro calcification border characterization", IEEE Journal on Engineering in Medicine and Biology, vol. 28, Issue 5, pp. 41-43, September 2010.
- [10] A. Sultana, M. Ciuc, L. Florea and C. Florea, "Detection of Mammographic Micro calcifications using a Refined-Region Growing approach", IEEE International Symposium on Signals, Circuits and Systems, pp. 1-4, July 2010.
- [11] J. Tang, R. M. Rangayyan, J. Xu, El Naga and I. Y. Yang, "Computer-Aided Detection and Diagnosis of Breast Cancer with Mammography: Recent Advances", IEEE Transactions on Information Technology in Biomedicine, vol. 13, no. 2, March 2011.
- [12] M. Sameti, R. K. Ward, J. Morgan -Parkes and B. Palcic, "Image Feature Extraction in the Last Screening Mammography Prior to Detection of Breast Cancer", IEEE Journal of Selected Topics in Signal Processing, vol. 3, Issue 1, pp. 46-52, February 2011.
- [13] R. Hupes and N. Karssemeijer, "Use of Normal Tissue Context in Computer-Aided Detection of Masses in Mammograms", IEEE Trans. on Med. Img, vol. 28, Issue. 12, pp. 2033-2041, December 2010.
- [14] A. Lambrou, H. Papadopoulos and A. Gammerman, "Evolutionary Conformal Prediction for Breast Cancer Diagnosis", IEEE 9th International Conference on Information Technology and applications in Biomedicine, pp. 1-4, January 2011.
- [15] X. Gao, Y. Wang, X. Li and D. Tao, "On Combining Morphological Component Analysis and Concentric Morphology Model for Mammographic Mass Detection", IEEE Transactions on Information Technology in Biomedicine, vol. 14, no. 2, pp. 266-273, March 2011.



Rash Bihari Dubey was born in India on 10th November 1961. He received the M. Sc. degree in Physics with specialization in Electronics in 1984 from Agra University Agra, India, the M. Tech. degree in Instrumentation from R.E.C. Kurukshetra, India in 1989 and the Ph.D. degree in Electronics Enggg., from M. D. University, Rohtak, India in 2011. He is at present Professor and Head in the Department of Electronics and Communication Engineering at Hindu College of Engineering, Sonapat, India. He has well over 40 publications in both conferences and journals to his credit His research interest are in the areas of Medical Imaging, Digital Signal Processing, Digital Image Processing, Biomedical Signal Analysis, and Industrial Real Time Applications.



Monika Sharma was born on 13th November 1988. She received her B. Tech. degree in Electronics and Communication Engineering from M. D. U., Rohtak in 2010 and M. Tech. Degree in Electronics and Communication Engineering from M. D. U., Rohtak in 2012.



Sujata Bhatia was born on 21st October 1987. She received her B. Tec. degree in biomedical engineering from Guru Jambheshwer University of science and technology, India 2010 and pursuing M. Tech. in Instrumentation and Control from DCRUST, Murthal, Sonapat, India.



S. K. Gupta was born in India on 20th May 1967. He received the M. Sc. degree in Mathematical Statistics in 1989 and Ph. D. Degree in Reliability Engineering in 1996 from M.D.U., Rohtak., India. He has 15 research publications of national and international journal of repute. He has published many books and various research papers have been accepted for publication. He organized various seminars / conferences. He is active member of the panel of M.D.U., Rohtak inspection team for grant of permanent affiliation to the engineering Institutes. He is at present working as a Principal of the Vaish College of Engineering, Rohtak, India. His research interests are in the areas of reliability engineering, mathematical modeling and simulation.