

A Study on Morphological and Textural Features for Classifying Breast Lesion in Ultrasound Images

P. K. Saranya¹, Dr. E. S. Samundeeswari²

Research Scholar, Department of Computer Science, Vellalar College for Women, Tamil Nadu, India¹

Associate Professor, Department of Computer Science, Vellalar College for Women, Tamil Nadu, India²

ABSTRACT: Developing a Computer Aided Diagnosis (CAD) system for Breast Ultrasound (BUS) is a high exigent work in medical imaging analysis systems. BUS detects abnormalities or suspicious regions accurately that is misspelled by mammogram, and helps to find the varying degrees of malignancy. Extracting discriminating features in CAD is a most prominent task that results in good diagnosis. This study analyzed and investigated thirty three quantitative Morphological features and thirty eight Textural features (six Histogram, fourteen GLCM, eleven GLRLM, Fractal Dimension and six Tamura) of breast lesions on ultrasound image for prediction and classification of benign and malignant tumor through Computer Aided Diagnosis System.

KEYWORDS: Breast Ultrasound, Morphological Features, Textural Features, Feature Extraction, Computer Aided Diagnosis

I. INTRODUCTION

Early diagnosis is the effective way to increase the survivability of women affected by breast cancer. Mammogram is the recommended tool to visualize small and non-palpable tumor, but it produces ionization and low negative value in screening pregnant women and women with dense breast [17]. Breast Ultrasound (BUS) is a painless, non invasive, less expensive and less time consuming modality employed as an adjunct to mammography to overcome these drawbacks. Physicians evaluate the lesion and tissue disparity in BUS image which is captured by internal echoes of US waves. Since ultrasound image interpretation is operator dependent, it is difficult to extract the accurate lesion portion due to inherited speckle noise and tissue related textures. To overcome these difficulties, researchers come forward to develop a Computer Aided Diagnosis (CAD) system for BUS to help physicians to perform effective diagnosis [9, 16, 21]. CAD system is designed with preprocessing, segmentation, feature extraction and classification stages to detect and extract abnormalities or suspicious portion for identifying varying degrees of malignancy [5, 19, 24, 34, 35]. Feature extraction is the most prominent stage after segmentation where significant features are extracted from lesion portion, and the corresponding numerical values are stored as feature vector [17]. During classification stage, highly discriminated features from feature vector are finally used for discriminating benign and malignant tumor appropriately.

In this paper, the morphological, textural, model-based and descriptor features extracted and defined by many researchers for differentiating benign from malignant tumor are studied, and analyzed for future work. This study also describes the computer based quantitative and qualitative Morphological and Textural features extracted for effective classification of benign from malignant lesion.

II. RELATED WORK

Segmentation is the crucial task in segmenting Breast Ultrasound Image precisely. Segmentation can be done by any one of the methods like threshold, histogram, merging, image growing, clustering, soft computing and others methods. The delineated and the segmented sample image is shown in Fig. 1. Properly segmented Region of Interest will help to

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

extract features accurately from ultrasound image to differentiate benign and malignant tumor. Many features may be either redundant or irrelevant. Hence there is a need to study and analyze the features to form an optimum feature subset for CAD system [7]. The features of breast ultrasound are categorized as Morphological, Textural, Model-based and Descriptor features [26, 41].

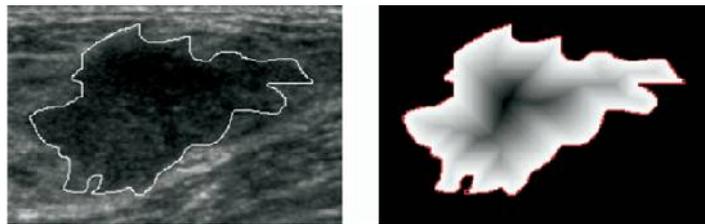


Fig. 1 A delineated and segmented sample BUS image

A. Morphological Features

The features extracted from shape and contour properties of lesion are defined as Morphological features [13]. Many researchers designed various statistical parameters for shape and margin characteristics to differentiate benign and malignant tumor [6, 17, 19, 39, 45, 46, 48]. The descriptions of important geometrical features widely used are described in Table 1.

Table 1. Morphological Features based on Geometrical structure

S.No	Shape Features	Descriptions
M1	Perimeter ^{[17] [46] [48]}	It represents the length of the circumference of tumor portion. Malignant tumors have irregular shapes with large tumor perimeter.
M2	Area ^{[17] [46] [48]}	It is the area of the breast tumor. Malignant tumors have a large area compared with benign tumors.
M3	NSPD (Number of Substantial Protuberances and Depressions) ^[17]	It is used to calculate the number of concave points that defines the level of boundary irregularity. If P_i is a point in the contour, the k-curve angle of P_i , i.e. θ_i , can be obtained by P_i , $P_i + k$ and $P_i - k$. If θ_i is $\leq 40^\circ$ then P_i is defined as a convex point and as a concave point if θ_i is $> 40^\circ$. The sample points of malignant tumor are shown in Fig. 2(a). $NSPD = 2 \times n$, where n is the number of concave points
M4	LI (lobulation index) ^[17]	The lobe region is a region enclosed by a lesion contour (i.e. a line connected by any two adjacent concave points). A_{max} and A_{min} denote the sizes of the largest and the smallest lobe regions, and $A_{average}$ denotes the average size of all lobe regions. $LI = \frac{A_{max} - A_{min}}{A_{average}}$
M5	ENC (elliptic-normalized circumference) ^[17]	$ENC = \frac{\text{Equivalent_Ellipse_Perimeter}}{\text{Perimeter}}$ where Equivalent_Ellipse_Perimeter is the perimeter of the ellipse drawn around the tumor region.
M6	ENS (elliptic-normalized skeleton) ^[17]	If a set S is defined as skeleton points of a tumor region, then ENS is defined as the sum of the skeleton points in S. A malignant lesion always has a twisted boundary and generates a large ENS. The sample skeleton of malignant tumor is shown in Fig. 2(b).
M7	LS_Ratio (long axis to short axis ratio) ^[17]	It is the length of the ratio of the major (long) axis to minor (short) axis of the equivalent ellipse.
M8	Aspect_Ratio ^{[6] [17] [39] [45]}	It is the ratio of a tumor's depth to width. $\text{Aspect_Ratio} = \frac{\text{Max_Diameter}}{\text{Min_Diameter}}$ If it is greater than 1, the tumor has a high probability of being malignant.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

M9	Form_Factor ^{[6][17][45]}	It defines to what extent the shape of tumor differ from circle or elliptical shape. Form_Factor = $\frac{4\pi * \text{Area}}{\text{Perimeter}^2}$
M10	Roundness ^{[6][17][45]}	Roundness = $\frac{4 * \text{Area}}{\pi * \text{Max_Diameter}^2}$ where Max_Diameter is the length of the major axis from equivalent ellipse of a tumor. If roundness is close to 1, then the tumor shape is round.
M11	Solidity ^{[6][17][39][45][46][48]}	Solidity = $\frac{\text{Area}}{\text{Convex_Area}}$ where Convex_Area is the area of convex hull of a tumor. If solidity is close to 0, the tumor has high possibility of malignant.
M12	Convexity ^{[6][17][45]}	Convexity = $\frac{\text{Convex_Perimeter}}{\text{Perimeter}}$ where Convex_Perimeter is the perimeter of convex hull of a tumor.
M13	Extent ^{[6][17][45][46][48]}	It specifies the ratio of pixels in the region to pixels in the total bounding box. Extent = $\frac{\text{Area}}{\text{Bounding_Rectangle}}$ where the Bounding_Rectangle is the smallest rectangle containing tumor.
M14	TCA_Ratio ^[17]	TCA_Ratio = $\frac{\text{Area}}{\text{Convex_Area}}$
M15	TEP_Ratio (tumor perimeter to ellipse perimeter ratio) ^[17]	It is the ratio of tumor perimeter to its corresponding ellipse perimeter. TEP_Ratio = $\frac{\text{Perimeter}}{\text{Ellipse_Perimeter}}$
M16	TEP_Difference (difference between tumor perimeter and ellipse perimeter) ^[17]	It is defined as the difference between tumor perimeter and the corresponding ellipse perimeter. TEP_Difference = $\text{Perimeter} - \text{Ellipse_Perimeter}$
M17	TCP_Ratio (tumor perimeter to circle perimeter ratio) ^[17]	It is the ratio of perimeter of a tumor to the corresponding circle, if the corresponding circle has the same perimeter as the perimeter of tumor then it must be benign. TCP_Ratio = $\frac{\text{Perimeter}}{\text{Circle_Perimeter}}$
M18	TCP_Difference (difference between tumor perimeter and circle perimeter) ^[17]	It is defined as the difference between the tumor perimeter and its corresponding circle. TEP_Difference = $\text{Perimeter} - \text{Circle_Perimeter}$ If the corresponding circle is having the same area as the tumor then it has high possibility of benign character.
M19	AP_Ratio (Area to Perimeter ratio) ^[17]	It is the ratio of the area and the perimeter of a tumor. AP_Ratio = $\frac{\text{Area}}{\text{Perimeter}}$
M20	Volume (VOL) ^[19]	If p is the pixel size, t is the thickness and r is the individual radial length, then VOL = $\sum_{\forall x, \forall y, \forall z} F_{ROI}(x, y, z) * p^2 * t$ where F _{ROI} is the pixels in Region of Interest (RoI)
M21	Surface Area (SURF) ^[19]	SURF = $\sum_{\forall x, \forall y, \forall z} S_{ROI}(x, y, z) * p * t$ where S _{ROI} is the pixels along the boundary of RoI

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

M22	Compactness (COMP) ^{[19][46][48]}	It is defined as the ratio of the square of surface area to the total volume of the lesion. A spiculated mass (benign) will have a high compactness index i.e. ≈ 1 . $COMP = \frac{Surf^2}{Vol}$
M23	NRL (Normalized Radial Length) mean ^{[19][8]}	It is defined as the mean of the Euclidean distance from the object's center (Center of Mass) to each of its contour pixels of the lesion. $NRL\ mean = \mu_{NRL} = \frac{1}{N} \sum_{j=1..N} r_j$ where r_j is the individual radial length from centroid to contour points and N is the number of contour pixels.
M24	Sphericity ^[19]	It is the ratio of the mean of NRL to standard deviation of NRL $Sphericity = \frac{\mu_{NRL}}{\sigma_{NRL}}$ where σ_{NRL} is the SD of NRL
M25	NRL entropy ^[19]	$NRL\ entropy = E_{NRL} = - \sum_{j=1..N} prob_j \log_2(prob_j)$ where $prob_j = \frac{r_j}{\sum r_j}$
M26	NRL ratio ^[19]	$NRL\ ratio = R_{NRL} = \frac{1}{N * \mu_{NRL}} \sum_{j=1..N} (r_j - \mu_{NRL}) \quad where \quad r_j > \mu_{NRL}$
M27	Roughness. ^{[19][8]}	$Roughness = \frac{\sqrt{\frac{1}{N} \sum_{j=1..N} (r_j - \mu_{NRL})^4} - \sqrt{\frac{1}{N} \sum_{j=1..N} (r_j - \mu_{NRL})^2}}{\mu_{NRL}}$
M28	MajorAxisLength ^{[46][48]}	It is the length (in pixels) of the major axis of the ellipse that has the same normalized second central moments
M29	MinorAxisLength ^{[46][48]}	It is the length (in pixels) of the minor axis of the ellipse that has the same normalized second central moments
M30	Eccentricity ^{[46][48]}	It is the ratio of the distance between the centroid of an ellipse and its major axis length. The values ranges between 0 and 1, and 0 represents circle.
M31	EquivDiameter ^{[46][48]}	The diameter of a circle with the same area as region. $Equivdiameter = \sqrt{4 * Area / pi}$
M32	Smoothness ^{[46][48]}	It is quantified by measuring the difference between length of radial line and the mean length of the lines surrounding it.
M33	Speculation ^[40]	It calculates the mean of the variation of boundary pixels from the centre of the image. To extract speculation, the radial distances from the center of the image for n number of adjacent boundary pixels are calculated. If the difference between the consecutive radial distances varies higher then increment the malignant count. If the value of count is high then it is indicates a sign of malignant tumor.

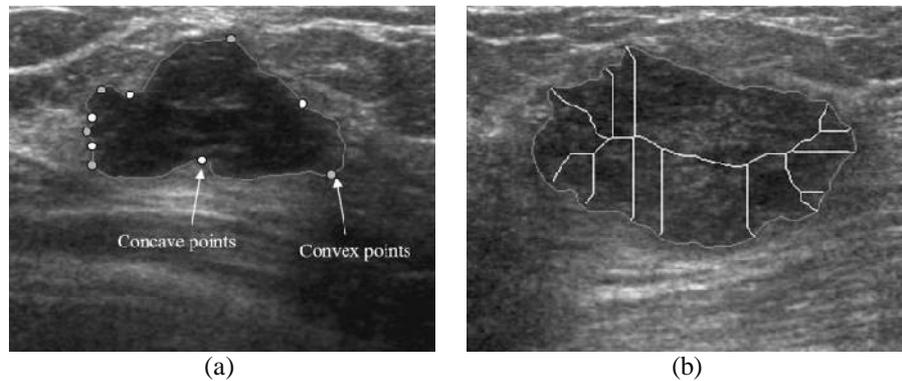


Fig. 2 (a) Convex and Concave Points, and (b) Skeleton (internal lines)
- of a malignant breast tumor.

These Morphological features can be extracted and analyzed for effective cancer diagnosis in CAD systems. Chang et al.(2005) and Wen-Jie Wu et al.(2012) applied six Morphological features such as Form_Factor, Roundness, Aspect_Ratio, Solidity, Convexity and Extent (M8-M13 in Table 1) for distinguishing breast masses in breast sonographic image CAD system and yielded 77.1% classification accuracy [6, 45]. Huang et al.(2008) applied nineteen geometrical shape features M1 – M19 in Table 1 including the features extracted by Chang et al. and yielded 82.2% classification accuracy in distinguishing benign breast lesion and malignant lesion [17]. Ki Nei et al. (2009) used eight geometrical features M20 – M27 in Table 1 for diagnostic prediction in Breast MRI and achieved 80% AUC (area under curve) [19]. The features M20-M22 (Volume, Surface Area and Compactness) are mostly applied for 3D image properties. Chou et al. (2001) applied three shape features- circularity, NRL and Roughness for breast cancer diagnosis and achieved 97.2% sensitivity and 80.0% specificity [8]. Yasmeen Mourice George et al. (2014) used ten shape based features (i.e. M1, M2, M11, M13, M22 and M28-M32 in Table 1) for detecting and diagnosing tumor using cytological images in Breast CAD system and yielded 99.7% sensitivity and specificity [46,48]. Shirley Selvan et al. (2010) defined five strain and shape features including Area difference, Solidity, Perimeter difference, Contour difference, Width-height ratio for characterizing the breast lesions in ultrasound echography and elastography [39]. Beherans et al. (2007) extracted features based on the morphological descriptors described in Breast Imaging-Reporting and Data System (BI-RADS) developed by the American College of Radiology [3, 4, 28]. From these analyses, it is found out that the Morphological features are widely used in CAD system in discriminating benign from malignant and it is also observed that malignant tumor has larger values than benign tumor for all the geometrical shape features.

B. Textural Features

Pictographically, the texture or texel is termed as a repeating pattern of local variations in grey level intensity. Statistical methods are used to characterize the texture indirectly according to the gray level intensities of an image in particular area [10, 25, 30]. Sliding window concept is used to find and measure the changes of texture in the particular region of an image. The windows of larger size estimate the feature value and smaller size estimate the location precisely. This concept analyzes the spatial distribution of gray values by computing the local features at each point (central point) in a window to derive a set of statistics. The Textural feature extraction method is carried out by first order (estimate properties of one pixel), second order (pair of pixels) and higher (more pixels) order statistics. The first order statistics estimates the average and variance properties that are derived from histogram and the second order statistical features for texture analysis are mostly derived from the co-occurrence matrix. The higher order textural features are derived from the Laws' texture energy measures, Gray Level Run Length Matrices (GLRLM), and Fourier analysis. In this study, Histogram, Gray Level Co-occurrence Matrix (GLCM), Gray Level Run Length Matrix (GLRLM), fractal dimension and Tamura Textural features are described.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

B.1. Histogram Features

Histogram based feature extraction is a technique for texture analysis of an image based on the first and second order statistical properties [27]. The six histogram textural features are calculated using the probability distribution of the intensity levels in the histogram bins of all blocks [11]. Let $M \times N$ be a size of an image, G be the graylevels and $p(i)$ is the gray level intensity histogram, then the descriptions of the six features are shown in Table 2.

Table 2. Histogram Features based on Gray Level Intensity Histogram

S. No	Textural Features-Histogram	Descriptions
H1	Mean ^{[11][27]}	It measures the average value of the intensity values. $\text{Mean} = \mu = \frac{\sum_{i=0}^{G-1} i * p(i)}{\sum_{i=0}^{G-1} p(i)} = \sum_{i=0}^{G-1} i * P(i) \text{ where } P(i) = \frac{p(i)}{\sum_{i=0}^{G-1} p(i)}$
H2	Variance ^{[11][27][47]}	It measures the region roughness. $\text{Variance} = \sigma^2 = \sum_{i=0}^{G-1} (i - \mu)^2 * P(i)$
H3	Skewness ^{[11][27]}	It refers to how far Textural intensities (darker / lighter pixels) vary from average. $\text{Skewness} = \sum_{i=0}^{G-1} (i - \mu)^3 * P(i)$
H4	Kurtosis ^{[11][27]}	It refers to how far the greylevel distribution is uniform. $\text{Kurtosis} = \sum_{i=0}^{G-1} (i - \mu)^4 * P(i)$
H5	Energy ^{[11][27]}	It refers to how uniform/non-uniform is the greylevel distribution. $\text{Entropy} = - \sum_{i=0}^{G-1} P(i) * \log_2 * P(i)$
H6	Entropy ^{[11][27] [47]}	It refers to the intensity variation in the region. $\text{Energy} = \sum_{i=0}^{G-1} [P(i)]^2$

Recent studies have showed that Mehdi Hassan et al. (2012) considered histogram and few GLCM measures to extract features from carotid artery ultrasound images for segmentation [27]. Fazal Malik et al. (2013) applied histogram features to retrieve image detail and obtained 79% recall [11]. It is found that the histogram features can be assessed individually or combined with other Textural features to identify the difference in classification accuracy.

B.2. GLCM Features

Haralick et al.(1973) defined a second order statistical method to compute fourteen grey-level co-occurrence matrix (GLCM) enhancement features [14]. GLCM is also known as gray-level spatial dependence matrix (GLSDM) that defines the spatial (textural) relationship between two pixels, the first pixel is a reference and the second is a neighbor pixel. If $\{I(x,y), 0 \leq x \leq N_x-1, 0 \leq y \leq N_y-1\}$ is an image of size $N_x \times N_y$ with G gray levels then the $G \times G$ is the gray level matrix for a displacement vector $d=(dx,dy)$ and direction θ . The gray level co-occurrence matrix P_d^θ is calculated by how often a pixel with the intensity (gray-level) value i occurs in a specific spatial relationship to a pixel with the value j . Each element (i,j) in the resultant P_d^θ is calculated by the sum of number of times the pixel value i occurred in specific spatial relationship to the pixel value j at a given distance d and direction θ [i.e. averaging the four values of θ -

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

0°, 45°, 90°, and 135°] [1, 42, 43]. The horizontal, diagonal, vertical and anti-diagonal direction of θ defining 0°, 45°, 90°, and 135° of a gray level matrix is shown in Fig. 3(a). The gray level co-occurrence matrix P_d^θ is calculated as

$$P_d^\theta(i, j) = \#\{((r, s), (t, v)) : I(r, s) = i, I(t, v) = j\}$$

where $(r, s), (t, v) \in N_x \times N_y$ i.e. i and j is the pair of pixels at the position (r, s) and (t, v) in the gray level co-occurrence matrix, and the displacement vector from one pixel position (t, v) to another pixel position (r, s) is calculated as $(t, v) = (r + dx, s + dy)$.

Fig. 3(b) depicts a sample 5x5 GLCM matrix calculation for $\theta = 0^\circ$ and $d=2$ and the number of occurrences of pairs pixels between same and different gray level values.

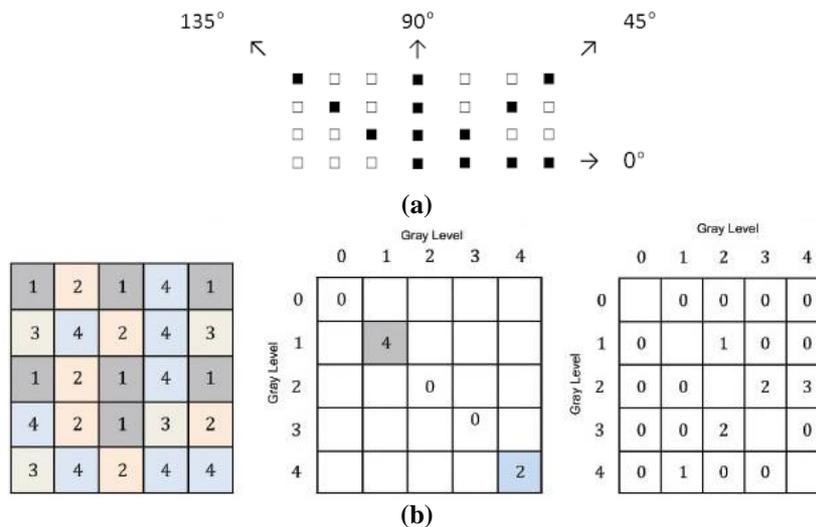


Fig. 3 (a) Various θ directions in gray level matrix and
(b) 5x5 GLCM matrix calculation for $\theta = 0^\circ$ and $d=2$

For $N \times N$ is an image with $W \times W$ window size, 14 GLCM features extracted in recent research works are shown in Table 3.

Table 3. GLCM Textural Features based on Co-occurrence Matrix

S. No	Textural Features- GLCM	Descriptions
T1	Uniformity / Energy / Angular Second Moment (ASM) ^{[10][12][42]}	It is a measure of homogeneity of an image Angular Second Moment (ASM) = $\sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{P_{d,\theta}(i, j)\}^2$ ASM will be zero if they are homogeneous.
T2	Entropy (E) ^{[10][12][42] [27]}	Entropy = $E = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j) \times \log(P_{d,\theta}(i, j))$
T3	Contrast (CTR) / Inertia (INR) ^{[10][12][42] [27]}	Contrast measures the local intensity variation between the referred pixel and its neighbor. CTR = $\sum_{n=0}^{N_g-1} n^2 \left\{ \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j) \right\}$, where $ i - j = n$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

		$INR = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i-j\}^2 * P_{d,\theta}(i, j)$
T4	Correlation (COR) ^{[10][12][42] [27]}	<p>It is a measure of gray level linear dependence between the pixels at the specified positions relative to each other. If 0 then it is uncorrelated and 1 it is perfectly correlated.</p> $COR = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j) \frac{(i-\mu_i)(j-\mu_j)}{\sigma_i \sigma_j}$
T5	Homogeneity / Inverse Difference moment (IDM) ^{[10][12][27]}	<p>It measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal</p> $IDM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \frac{1}{1+(i-j)^2} P_{d,\theta}(i, j)$ <p>IDM will result in one along the diagonals if the elements are closely distributed.</p>
T6	Cluster Shade (SHD) ^[27]	$SHD = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i+j-\mu_x - \mu_y\}^3 * P_{d,\theta}(i, j)$
T7	Cluster Prominence (PRM) ^[27]	$PRM = \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} \{i+j-\mu_x - \mu_y\}^4 * P_{d,\theta}(i, j)$
T8	Sum of Squares : Variance ^{[10][27][42]}	$\text{Square, Variance} = \sum_{i=0}^{2(N_g-1)} (i - SAV)^2 P_{x+y}(i)$
T9	Sum Average (SAV) ^{[10][27][42]}	$SAV = \sum_{i=0}^{2(N_g-1)} i * P_{x+y}(i)$
T10	Sum Entropy (SEN) ^{[27][42]}	$SEN = \sum_{i=0}^{2(N_g-1)} P_{x+y}(i) \log(P_{x+y}(i))$
T11	Difference variance ^{[10][27][42]}	$DVR = \sum_{i=0}^{N_g-1} (i-f)^2 P_{x-y}(i) \text{ where } f = \sum_{i=0}^{N_g-1} i * P_{x-y}(i)$
T12	Difference entropy (DEN) ^{[10][27][42]}	$DEN = \sum_{i=0}^{N_g-1} P_{x-y}(i) \log(P_{x-y}(i))$
T13	Information measures of correlation (1) (IMC-1) ^{[27][42]}	$IMC-1 = \frac{HXY - HXY1}{\max(HX, HY)}$
T14	Information measures of correlation (2) (IMC-2) ^{[27][42]}	<p>IMC-2 = $(1 - \exp[-2(HXY2 - HXY)])^{1/2}$ where</p> $P_x(i) = \sum_{j=0}^{N_g-1} P_{d,\theta}(i, j)$ $P_x(j) = \sum_{i=0}^{N_g-1} P_{d,\theta}(i, j)$ $HX = - \sum_{i=0}^{N_g-1} P_x(i) \log(P_x(i))$ $HY = - \sum_{i=0}^{N_g-1} P_y(i) \log(P_y(i))$

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

		$HXY = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i,j) \log(P_{d,\theta}(i,j))$ $HXY1 = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_{d,\theta}(i,j) \log(P_x(i)P_y(j))$ $HXY2 = - \sum_{i=0}^{N_g-1} \sum_{j=0}^{N_g-1} P_x(i)P_y(j) \log(P_x(i)P_y(j))$
--	--	--

Many conventional approaches used to study texture have concentrated on using 2D techniques to compute features relating to image texture. Dustin Newell and Ke Nie et al. (2010) combined ten GLCM features with few shape features and yielded 89% accuracy in classification of Benign and Malignant one [10]. Tingting Mu et al. (2008) considered 14 GLCM with 5 shape features and yielded $A_z=93$ (Receiver Operating Characteristic curve) [42]. Hadi Tadayyon et al. (2014), Tadayyon H et al. (2014), Mehdi Hassan et al. (2012), Liao Y-Y et al. (2011) and Alvarenga et al. (2007) had used the GLCM method for textural analysis of cancer category to design a quantitative ultrasound diagnosis and prediction system. Each GLCM feature exhibits the maximum and minimum gray value to differentiate benign and malignant lesion.

B.3. GLRLM Features

Grey-Level Run-Length Matrix (GLRLM) is a higher order method for extracting Textural features where a pattern of grey intensity pixel in a particular direction from a reference pixel is calculated and analyzed. Run Length is calculated by counting the number of adjacent pixels that has same intensity of grey value in a particular direction. GLRLM is a two dimensional matrix used for calculating the number of co-occurrence gray level value from grey level i to run length j in the particular direction θ . The horizontal, diagonal, vertical and anti-diagonal of θ in GLRLM will take the direction as $0^\circ, 45^\circ, 90^\circ$, and 135° and is shown in Figure 3(a). The sample 5x5 GLRLM matrix calculations for $\theta (0^\circ)$ and its occurrences of grey levels for run length of 3 is shown in Fig. 4.

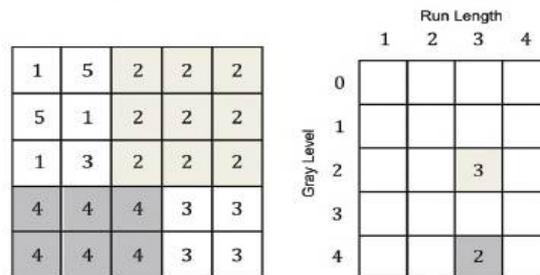


Fig. 4 5x5 GLRLM matrix calculation for $\theta = 0^\circ$ and run length=3

The 11 features extracted from segmented image using GLRLM are Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray-Level Nonuniformity (GLN), Run Length Nonuniformity (RLN), Run Percentage (RP), Low Gray-Level Run Emphasis (LGRE), High Gray-Level Run Emphasis (HGRE), Short Run Low Gray-Level Emphasis (SRLGE), Short Run High Gray-Level Emphasis (SRHGE), Long Run Low Gray-Level Emphasis (LRLGE) and Long Run High Gray-Level Emphasis (LRHGE).

Aswini kumar mohanty et al.(2011) extracted 5 features (Energy, Inertia, Entropy, Maxprob, Inverse) from GLCM matrix and 7 features (SRE, LRE, GLN, RLN, LGRE, HGRE, and SRLGE) from GLRLM matrix to predict and distinguish malignant mass image and benign mass with a level of accuracy of 92.3% [2]. Wei Q and Hu Y(2009) had combined GLCM and GLRLM for identifying fissure regions from isotropic CT image stacks and obtained 87% accuracy [44]. Raghesh Krishnan and Sudhakar(2013) had extracted the features using Gray Level Run Length matrix Features and performed classification using Support Vector Machines for liver ultrasound images [36].

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

B.4. Fractal Dimension

Fractals provide a succinct and accurate method for describing the quality of roughness and self-similarity of objects like sphere, cylinders, curves and cubes at different scales. But fractals are difficult for measuring irregular and complicated structures. Mandelbrot (1977) was the first researcher who described the fractal dimension for complicate structures which were not simply characterized by Euclidean geometry measure. The fractal dimension is the method of measuring the degree of irregularity or roughness of image texture. A texture in image is seen as a repetition of similar patterns that are randomly distributed and fractal dimension measures the patterns by translation, rotation and scaling [37]. The structure with rougher or more irregular character has greater fractal dimension [32, 33].

Box-counting and Korcak are the two methods used for calculating fractal dimension of an image. Box counting remains the most practical method to estimate the fractal dimension and is a simplification of the Hausdorff dimension [22, 29]. In Box counting, the image is covered with a grid of square cells with cell size(number of pixels) [15, 20, 23]. Given a fractal structure **A** in **d** dimension, the box counting method will partition the structure space with **d** dimensional fixed grid of square boxes of equal size **r**. Then the number of nonempty boxes **N(r)** of size **r** to cover fractal structure is

$$N(r) \sim r^{-D} \text{ where } -D = \lim_{r \rightarrow 0} \frac{\log N(r)}{\log r}$$

In Korcak method, the fractal dimension is calculated by plotting the intensity at each x and y position in the Z plane (textural surface) [38]. Later, Xiangjun Shi et al.(2006) designed and extracted five fractal dimensions of ultrasound breast image using above Hausdroff dimension equation [47]. The five features are

fd1 = fractal dimension of the ROI.

fd2 = fractal dimension of the suspicious area.

fd3 = fractal dimension of the ROI including the suspicious area.

fd4 = fd2 / fd1.

fd5 = fractal dimension of the edge of the suspicious area.

Fiz et al. (2014) uses fractal dimension as a main feature to characterize and analysis the inner structure of hilar and mediastinal nodes and to differentiate the benign endobronchial ultrasound nodes from malignant one [18].

B.5. Tamura Features

Mori Tamura and Yamawaki (1978) proposed six higher order Textural features- Coarseness, Contrast, Directionality, Line-Likeness, Regularity and Roughness. The six Tamura features extracted are shown in Table 4.

Table 4. Tamura Textural Features

S.No	Textural Features-Tamura	Descriptions
Ta1	Coarseness ^[31]	It relates to the distance in gray levels of spatial variations and identifies the texture that is mostly occurred (i.e. texture pattern that exists in large size) $A_k(x,y) = \sum_{i=x-2k-1}^{x+2k-1} \sum_{j=y-2k-1}^{y+2k-1} \frac{f(i,j)}{2^{2k}}$ where, $2^k * 2^k$ size is the average of neighborhood. $E_{k,h}(x,y) = A_k(x + 2_{k-1},y) - A_k(x - 2_{k-1},y) $ $E_{k,h}(x,y)$ calculates the difference between pair of averages corresponding to non-overlapping neighborhoods.
Ta2	Contrast ^[31]	It measures to what extent distribution of gray levels varies in an image subject to black and white. $\text{Contrast} = \sigma / (\alpha_4)$ $\alpha_4 = \mu_4 / \sigma^4$ where, μ_4 is the fourth moment about the mean and σ^2 is the variance. The second order and normalized fourth-order central moments of the gray levels are used to define the contrast and its closest value according to Tamura.
Ta3	Directionality ^[31]	It measures the total degree of directionality in an image or frequency distribution of oriented local

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

		edges against their directional angles. $\text{Directionality} = 1 - r \cdot n_{\text{peaks}} \sum_{p=1}^{n_{\text{peaks}}} \sum_{a \in w_p} (a - a_p)^2 H_{\text{directionality}}$ <p>where, n_{peaks} is the number of peaks, a_p is the position of the peak, w_p is the range of the angles attributed to the P^{th} peak, r denotes a normalizing factor related to quantizing levels of the angles a, and a denotes quantized directional angle, $H_{\text{Directionality}}$ is the histogram of quantized direction values, a is constructed by counting number of the edge pixels with the corresponding directional angles.</p>
Ta4	Line-Likeness ^[31]	Line-Likeness in an image is average coincidence of direction of edges that co-occurred in the pairs of pixels separated by a distance along the edge direction in every pixel.
Ta5	Regularity ^[31]	It measures a regular pattern or similar pixels in an image.
Ta6	Roughness ^[31]	It is the summation of contrast and coarseness of an image.

Neelima Bagri and Punit Kumar Johari (2015) combined Tamura texture and shape invariant features for content based image retrieval system and yielded 0.78% precision and 0.87% recall than the GLCM Textural and shape moment invariant features [31]. Yu et al.(2015) used Tamura features to identify the needle insertion point precisely in ultrasound images of lumbar spine in an automatic manner to facilitate the anesthetists' work [50]. Ying Liu (2015) has used Tamura feature extraction method for image retrieval and found Contrast-Tamura feature suits more for their research work since it describes the statistical distribution of the brightness in the entire image [49]. Madheswaran and Anto Sahaya Dhas (2015) extracted Texture and Tamura features using GSDM and Tamura method, and used Genetic algorithm for feature selection. Support Vector Machine is used as classifier and yielded 98.83% accuracy [26]. This study identified six Histogram, fourteen GLCM, eleven GLRLM, Fractal Dimension and six Tamura second order statistical measures that can be extracted from ROI to classify the benign and malignant tumor more accurately in CAD system.

C. Model Based and Descriptor Features

Model base Features are extracted from the backscattered echo from the breast tissue [7]. The resulting echo's are modeled and the parameters of the model are used to distinguish benign and malignant tumor. Since the background of the model is complex, the estimation of parameter is very complex. PSNL mode, Nakagami model based, GS model, K distribution model, SNR, normalized spectral power, margin strength and speckle factor are the various feature that belong to Model based features [7]. Descriptor features are easy to understand since there is no empirical classification and numerical expressions [7]. Some of the features that mostly referred to benign characteristics are speculated margins, presence of calcifications, posterior shadow, echogenicity, duct extension, disortation and microlobulation. The posterior shadow refers to the region posterior to lesion portion that has darker gray value than the surroundings. The shape and bilateral refraction sign features indicate the high possibilities of malignant character. In shape feature, oval or round shape indicates the benign tumor and other irregular shape indicate malignant category.

III. CONCLUSION

The identification of tumor pathological characteristics is an important part of breast cancer diagnosis, prognosis, and treatment planning. Designing a Breast US CAD system is a painless screening tool in diagnosing the abnormal tissues effectively. Extracting discriminate features from lesion will lead to proper characterization of breast tumors in CAD system. This study analyzes 33 morphological and 38 textural features and reviews the individual and combined performance of these features in differentiating benign and malignant lesion.

REFERENCES

- [1] Alvarenga A.V., Pereira W.C.A., Infantsi A.F.C., and Azevedo C.M., "Complexity Curve and Grey Level Co-occurrence Matrix in the Texture Evaluation Of Breast Tumor On Ultrasound Images", Med Phys 34(2), 379-387, 2007.
- [2] Aswini kumar mohanty, Swapnasikta beberta, Saroj kumar lenka, "Classifying Benign and Malignant Mass using GLCM and GLRLM based Texture Features from Mammogram", International Journal of Engineering Research and Applications (IJERA), Vol. 1, No 3, pp.687-693, 2011.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

- [3] Behrens S., Laue H. H., Althaus M., et al, "Computer Assistance for MR Based Diagnosis of Breast Cancer: Present and future challenges", *Comput Med Imaging Graph.* Vol.31(4-5):236-247, 2007.
- [4] Breast Imaging Reporting and Data System (BI-RADS) Breast Imaging Atlas. American College of Radiology, 2003.
- [5] Chan H.P., Sahiner B., Petrick N., et al, "Computerized Classification Of Malignant And Benign Microcalcifications On Mammograms: Texture Analysis Using An Artificial Neural Network", *Phys Med Biol.* Vol.42(3):549-567, 1997.
- [6] Chang R.F., Wu W.J., Moon W.K., Chen D.R., "Automatic Ultrasound Segmentation And Morphology Based Diagnosis Of Solid Breast Tumors", *Breast Cancer Res Treat.* 89: 179-185, 2005.
- [7] Cheng H.D., Juan Shan, Wen Ju, Yanhui Guo, Ling Zhang, "Automated Breast Cancer Detection And Classification Using Ultrasound Images: A Survey," *Pattern Recognition*, 43, 299-317, 2010.
- [8] Chou Y.H., Tiu C.M., Hung G.S., et al. "Stepwise Logistic Regression Analysis Of Tumor Contour Features For Breast Ultrasound Diagnosis", *Ultrasound Med Biol.*;27:1493-8, 2001.
- [9] Drukker K K., Giger M.L, Vyborny C.J., et al, "Computerized Detection And Classification Of Cancer On Breast Ultrasound", *Acad Radiol.* Nov;526-535, 2004.
- [10] Dustin Newell, Ke Nie, Jeon-Hor Chen, Chieh-Chih Hsu, Hon J. Yu, Orhan Nalcioglu, Min-Ying Su, "Selection of Diagnostic Features On Breast MRI to Differentiate Between Malignant And Benign Lesions Using Computer-Aided Diagnosis: Differences In Lesions Presenting As Mass And Non-Mass-Like Enhancement" , *Eur Radiol* 20: 771-781. 2010.
- [11] Fazal Malik and Baharum Baharudin, "The Statistical Quantized Histogram Texture Features Analysis for Image Retrieval Based on Median and Laplacian Filters in the DCT Domain", *The International Arab Journal of Information Technology*, Vol. 10, No. 6, November 2013.
- [12] Hadi Tadayyon, Ali Sadeghi-Naini and Gregory J. Czarnota, "Noninvasive Characterization of Locally Advanced Breast Cancer Using Textural Analysis of Quantitative Ultrasound Parametric Images", *Translational Oncology* 7(6), 759-767, 2014.
- [13] Hamid Behnam , Fahimeh Sadat Zakeri , Nasrin Ahmadinejad, "Breast Mass Classification On Sonographic Images On The Basis Of Shape Analysis", *J Med Ultrasonics* 37:181-186, 2010.
- [14] Haralick R.M., Shanmugam K., Dinstein I., "Texture Features For Image Classification", *IEEE Trans SMC* 3:610-621, 1973.
- [15] Hausdorff F., Dimension und auseres Mass, [*Math. Annalen* 79, 157] ,1919.
- [16] Huang Y.L., Chen D.R., "Support Vector Machines In Sonography: Application To Decision Making In The Diagnosis Of Breast Cancer", *Clin Imaging*, May-Jun;29(3):179-184, 2005.
- [17] Huang Y.-L., Chen D.-R., Jiang Y.-R., Kuo S.-J., x Wu S.-J. and Moon W. K., "Computer-Aided Diagnosis Using Morphological Features For Classifying Breast Lesions On Ultrasound", *Ultrasound in Obstetrics & Gynecology*, Vol. 32, No 4, pages 565-572, September 2008.
- [18] José Antonio Fiz, Enrique Monte-Moreno, Felipe Andreo, Santiago José Auteri, José Sanz-Santos, Pere Serra, Gloria Bonet, Eva Castellà and Juan Ruiz Manzano , "Fractal Dimension Analysis Of Malignant And Benign Endobronchial Ultrasound Nodes", *BMC Medical Imaging* 2014.
- [19] Ke Nie, Jeon-Hor Chen, Hon J. Yu, Yong Chu, Orhan Nalcioglu, and Min-Ying Su, "Quantitative Analysis of Lesion Morphology and Texture Features for Diagnostic Prediction in Breast MRI", *Acad Radiol.* Dec; 15(12): 1513-1525, 2008.
- [20] Keller J. et al., "Texture Description And Segmentation Through Fractal Geometry", *Computer Vision, Graphics And Image Processing*, 45, 150-166, 1989.
- [21] Kim KG, Cho SW, Min SJ, "Computerized Scheme for Assessing Ultrasonographic Features of Breast Masses", *Acad Radiol*, Jan;12(1):58-66, 2005.
- [22] Li, J., Sun, C., Du, Q., "A New Box-Counting Method For Estimation Of Image Fractal Dimensions", In *Proceedings of IEEE International Conference on Image Processing*, pp 3029-3022, 2006.
- [23] Li J., Du Q., "An Improved Box-Counting Method For Image Fractal Dimension Estimation", *Pattern Recognition*, 42, 2460-2469, 2009.
- [24] Li H., Wang Y., Liu KJ, et al. "Computerized radiographic mass detection-part i. Lesion site selection by morphological enhancement and contextual segmentation", *IEEE Transactions on Medical Imaging*, 20:289-301, 2001.
- [25] Liao Y.-Y., Tsui P-H, Li C-H., Chang K-J., Kuo W-H., Chang C-C., and Yeh C-K., "Classification Of Scattering Media Within Benign And Malignant Breast Tumors Based On Ultrasound Texture-Feature-Based And Nakagami-Parameter Images", *Med Phys* 38(4), 2198-2207, 2011.
- [26] Madheswaran M., and Anto Sahaya Dhas D., "Classification Of Brain MRI Images Using Support Vector Machine With Various Kernels", *Biomedical Research*; 26 (3): 505-513, 2015.
- [27] Mehdi Hassan, Asmatullah Chaudhry, Asifullah Khan, Jin Young Kim, "Carotid Artery Image Segmentation Using Modified Spatial Fuzzy C-Means And Ensemble Clustering", *Computer methods and programs in Biomedicine*, 2012.
- [28] Melania Costantini, Paolo Belli, Roberta Lombardi, Gianluca Franceschini, Antonino Mulè, Lorenzo Bonomo, "Characterization of Solid Breast Masses Use of the Sonographic Breast Imaging Reporting and Data System Lexicon", *American Institute of Ultrasound in Medicine, J Ultrasound Med*, 25:649-659, 2006.
- [29] Min Long, Fei Peng, "A Box-Counting Method with Adaptable Box Height for Measuring the Fractal Feature of Images", *Radioengineering*, Vol. 22, No. 1, 2013.
- [30] Mori Tamura S. and Yamawaki T., "Textural Features Corresponding to Visual Perception", *IEEE Transactions on Systems, Man and Cybernetics*, vol. smc-8, no. 6, June, pp. 460-473, 1978.
- [31] Neelima Bagri and Punit Kumar Johari, "A Comparative Study on Feature Extraction using Texture and Shape for Content Based Image Retrieval", *International Journal of Advanced Science and Technology* Vol.80, 2015.
- [32] Peitgen, H., and Saupe, D., "The Science Of Fractal Images", Springer-Verlag, ISBN 0-387-96608-0, New York, 1988.
- [33] Peitgen, H.; Jurgen, H. and Saupe, D., "Chaos And Fractals: New Frontiers Of Science", Springer-Verlag, ISBN 0-387-97903-4, New York, 1992.
- [34] Polakowski WE, Cournoyer DA, Rogers SK, "Computer-Aided Breast Cancer Detection And Diagnosis Of Masses Using Difference Of Gaussians And Derivative-Based Feature Saliency", *IEEE Trans. Med. Imaging*, 16:811-819, 1997.
- [35] Rangayyan RM, EL-Faramawy NM, Desautels JE, "Measures of Acutance And Shape For Classification Of Breast Tumors", *IEEE Trans Med Imaging*, 16(6):799-810, 1997.

International Journal of Innovative Research in Science, Engineering and Technology

(An ISO 3297: 2007 Certified Organization)

Vol. 5, Issue 3, March 2016

- [36] Raghesh Krishnan K., Sudhakar R., "Automatic Classification of Liver Diseases from Ultrasound Images Using GLRLM Texture Features", *Advances in Intelligent Systems and Computing*, pp 611-624, 2013.
- [37] Redouan Korchiynel, Sidi Mohamed Farssi, Abderrahmane Sbihi, Rajaa Touahni, Mustapha Tahiri Alaoui, "A Combined Method Of Fractal And GLCM Features For MRI And CT Scan Images Classification", *Signal & Image Processing : An International Journal (SIPIJ)* Vol.5, No.4, August 2014.
- [38] Russ, J.C., "Fractal Surfaces", Plenum Press, ISBN 0-306-44702-9, New York, 1994.
- [39] Shirley Selvan, M. Kavitha, S. Shenbagadevi and S. Suresh, "Feature Extraction for Characterization of Breast Lesions in Ultrasound Echography and Elastography", *Journal of Computer Science* 6 (1): 67-74, 2010.
- [40] Segyeong Joo, Yoon Seok Yang, Woo Kyung Moon, and Hee chan Kim. "Computer-Aided Diagnosis of solid Breast Nodules: Use of an Artificial Network Based on Multiple Sonographic Features". *IEEE transaction on medical imaging*, Vol. 23, No. 10, October 2004.
- [41] Tadayyon H, Sadeghi-Naini A, Wirtzfeld L, Wright FC, and Czarnota G., "Quantitative Ultrasound Characterization Of Locally Advanced Breast Cancer By Estimation Of Its Scatterer Properties", *Med Phys* 41(1), 2014.
- [42] Tingting Mu, Asoke K. Nandi, and Rangaraj M. Rangayyan, "Classification of Breast Masses Using Selected Shape, Edge-sharpness, and Texture Features with Linear and Kernel-based Classifiers", *Journal of Digital Imaging*, Vol 21, No 2 (June), 2008.
- [43] Tuan Anh Pham, "Optimization of Texture Feature Extraction Algorithm", MSc THESIS 2010.
- [44] Wei Q, Hu Y, "A Study On Using Texture Analysis Methods For Identifying Lobar Fissure Regions In Isotropic CT Images", *Conf Proc IEEE Eng Med Biol Soc.* 2009.
- [45] Wen-Jie Wu, Shih-Wei Lin, Woo Kyung Moon, "Combining Support Vector Machine With Genetic Algorithm To Classify Ultrasound Breast Tumor Images", *Computerized Medical Imaging and Graphics* 36, 627- 633, 2012.
- [46] Wolberg W. H., Street W. N., and Mangasarian O. L., "Machine Learning Techniques To Diagnose Breast Cancer From Image-Processed Nuclear Features Of Fine Needle Aspirates," *Comput. Appl. Early Detect. Staging Cancer*, vol. 77, no. 2/3, pp. 163-171, Mar. 1994.
- [47] Xiangjun Shi, H.D. Cheng, and Liming Hu, "Mass Detection and Classification in Breast Ultrasound Images Using Fuzzy SVM", *JCIS-2006 Proceedings*, 2006.
- [48] Yasmeen Mourice George, Hala Helmy Zayed, Mohamed Ismail Roushdy, and Bassant Mohamed Elbagoury, "Remote Computer-Aided Breast Cancer Detection and Diagnosis System Based on Cytological Images", *IEEE Systems Journal*, vol. 8, no. 3, september 2014.
- [49] Ying Liu, Zong Li, Zi-Ming Gao, "An Improved Texture Feature Extraction Method for Tyre Tread Patterns", *Lecture Notes in Computer Science* pp 705-713, 2015.
- [50] Yu S, Tan KK, Sng BL, Li S, Sia AT, "Lumbar Ultrasound Image Feature Extraction and Classification with Support Vector Machine", *Ultrasound Med Biol*,41(10):2677-89, 2015.