Asian Journal of Information Technology 15 (20): 4054-4062, 2016
ISSN: 1682-3915

Combination of Mammographic Texture Feature Descriptors for Improved Breast Cancer Diagnosis

[1]S. Sasikala and [2]M. Ezhilarasi
[1]Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore, India
[2]Department of Electronics and Instrumentation Engineering, Kumaraguru College of Technology, Coimbatore, India

Abstract: Computer Aided techniques developed for diagnosing the breast cancer plays a vital role in the early diagnosis of breast cancer and treatment in reducing the mortality rate. Better accuracy will generally be achieved using a combination of features instead of single type of feature descriptor. This research aims to improve the diagnostic accuracy and to reduce the false positive detection. Six different descriptors and their combination have been used to represent mammographic texture. The individual and combined feature vectors are reduced by principal component analysis and then classified by a multilayer Perceptron neural network using back propagation algorithm. The performance of the classification is evaluated with the texturefeaturesseparately and theircombination ie. the concatenation of the feature vectors from individual feature extraction techniques on the Digital Database for Screening Mammography (DDSM) and INbreast database by computing various performance metrics. The results show that the use of feature combination improves the performance of classification when a system cannot be tuned to an individual dataset. Eighteen performance metrics including Accuracy, Sensitivity, Specificity, Mathews Correlation coefficient, F1 score, discriminant power, Youden's index etc. Al these metrics were improved for the combined features for both dataset.

Key words: Mammogram, GLCM, GLDM, steerable pyramid, k-gabor, fractal, GLRLM, texture

INTRODUCTION

Breast cancer is a most common form of cancer found among women. It is the uncontrolled growth of abnormal cells at milk producing glands in the breast or in the passages that deliver milk to the nipple and is a leading cause of increase in mortality rate of women. Early detection of breast cancer is very important in increasing in survival rates and also helps in improving treatment option. Mammography is one of the effective techniques used for early detection of breast cancer.

Computer Aided Detection (CADe) and Computer Aided Diagnosis (CADx) systems are used to assess breast images objectively to overcome the subjective analysis made by the radiologists in double screening (Tang *et al*., 2009). The CADe systems determine doubtful regions in mammograms whereas the CADx systems are used to classify the abnormal regions as benign or malignant (Jalalian *et al*., 2013). Many research studies show that the CAD systems can be used as a secondary information as a replacement for double screening and can help radiologists for improving the accuracy of breast cancer detection and diagnosis. Detection/Diagnosis using a mammographic image is based on the fact that characteristics of pixels inside a tumor area are different from the other pixels inside the breast area. These characteristics can be simply related to grey-level intensity values or local texture or morphological measures. By properly segmenting the suspicious parts and extracting some representative features, those are clearly differentiating benign and malignant abnormalities; the performance of the CAD systems can be improved. (Bhanumathi and Suresh, 2013) provided an overview of recent advances in the development of CAD systems and a detailed introduction of some basic concepts related to breast cancer detection and diagnosis. Some of the CAD techniques developed for diagnosing the breast cancer such as detection of masses, calcification, architectural distortion, bilateral asymmetry in mammograms are also detailed.

Arnau Oliver *et al*. (2010) presented and reviewed different approaches to the automatic and semi-automatic

Corresponding Author: S. Sasikala, Department of Electronics and Communication Engineering, Kumaraguru College of Technology, Coimbatore, India

methods for detection and segmentation of mammographic masses and evaluated seven frequently used strategies for mass detection by single view using ROC and FROC analysis. An overview of digital image processing techniques and pattern analysis techniques such as contrast enhancement, detection and analysis of calcifications, masses and tumors, bilateral asymmetry and architectural distortion are discussed by (Rangayyan *et al*., 2007). The authors reported that Computer-Aided Diagnosis (CAD) is effective in reducing the errors in mammographic screening to a level comparable to a level achieved with double reading. Belle *et al*. (2013) stated that computerized decision support systems for diagnosis and prognosis have proved to be instrumental in medicine and applying advances in computational methods and techniques improve the quality of information obtained from feature extraction and feature selection. Various techniques for mass detection and classification and their advantages and drawbacks are discussed and compared by Cheng *et al*. (2006). A performance analysis was made on various segmentation techniques of mammogram.

Textures are one of the important characteristics to identify objects in various images. Visual texture is typically defined (Moreira *et al*., 2012) as the statistical variation in position, size, shape and orientation of repetitive visual features over an area. In the past few decades, texture descriptors have achieved considerable interest and various techniques have been introduced for the texture classification (Liao *et al*., 2009). Texture analysis in mammography is one of the effective methods to identify the breast cancer type. Grey level co-occurrence matrix, k-Gabor method, Grey level run length matrix, Grey level difference method, Fractal and Steerable pyramids are some of the more common descriptors used to perform texture classification. Various methods of texture analysis for the detection masses and micro calcifications in mammography were discussed by Sabu and Ponraj (2012).

The efficacy of the various texture descriptors, both in seclusion and in combination were performed by Barley and Town (2014).They suggested that in the case where the optimal combination of features and machine-learning algorithms is not known, better accuracy might be achieved using a combination of features. Feature combination means the concatenation of the feature vectors from more than one feature extraction techniques. This is likely due to the fact that the weaknesses in any individual scheme may be countered by combination of schemes. They showed that the majority of machine learning techniques produce better average accuracy over all datasets using feature combinations. Liasis *et al*. evaluated the accuracy of breastdensityclassification through SVM classifiers modeled on the histograms of the threedifferenttexturefeature sets separately and theircombination on the Medical Image Analysis Society (MIAS)mammographicdatabase (Liasis *et al*., 2012).

In this serearch, statistical distributions ofsix differenttexturedescriptors computed from suspicious regions of mammograms and theircombinationare investigated with Multi layer Perceptron Neural network for objectiveclassification of breastcancer on Digital Database for Screening Mammography (DDSM) and INbreast databases.

MATERIALS AND METHODS

Mammograms from two data bases were used in this work. The first dataset is the Digital Database for Screening Mammography (DDSM) obtained from the University of South Florida and the second data base is INbreast from Portugal. Mammograms obtained from the databases are denoised, then preprocessed to remove unwanted labels and pectoral muscle. Then suspicious regions are segmented by fuzzy c means clustering technique. After segmentation, various texture feature descriptors namely Grey level co-occurrence matrix, k-Gabor wavelets, Grey level run length matrix, Grey level difference method and steerable pyramids have been extracted from the segmented suspicious regions. Finally the mammograms were classified as benign or malignant category by applying the extracted features as input to a multilayer Perceptron neural network. The performance of classification with the isolated feature descriptors as well as their combination is evaluated by using the performance metrics such as accuracy, true positive rate, false negative rate, true negative rate, false positive rate, positive predictive rate, false discovery rate, negative predictive value and false omission rate.

Databases: The DDSM database is digitized film screen mammograms with associated ground truth and other information. It contains 2620, four view, screening mammograms obtained from Massachusetts General Hospital, Wake Forest University School of Medicine, Sacred Heart Hospital and Washington University of St. louis school of Medicine. After digitization, background area of mammograms were removed by cropping. They were then manually processed to darken pixels in regions that contained patient identifiers and were stored in files using a loss-less compression algorithm. The database was obtained in lossless jpeg format from Dr. Thoma Deserno (nee Lehmann), Department of Medical Informatics, Aachen University of Technology, Germany.

INbreast dataset has 410 Full Field Digital Mammograms (FFDM) of 115 cases from which 90 cases are from patients with both breasts affected ie. two images (MLO and CC) of each breast (four images per case) and 25 cases are from mastectomy woman patients (two images per case) saved as 14-bit contrast resolution in the DICOM format. The image matrix was 3328 × 4084 or 2560 × 3328 pixels, depends on the compression plate used in the acquisition. Several types of lesions such as masses, calcifications asymmetries and distortions were included and also accurate contours made by specialists are provided in XML format. INbreast has mammograms from screening, diagnostic and follow-up cases. The database was obtained in DICOM format from Jaime S. Cardoso, breast research group, INESC porto, portugal.

Preprocessing and segmentation: Mammograms are difficult to interpret. Hence, preprocessing phase is necessary to improve the quality of the images and to make the feature extraction step more reliable. In order to limit the search for abnormalities by CADe systems to the region of the breast without undue impact from the background of the mammogram, removal of artifacts and pectoral muscle is necessary. Resolution of the mammograms is affected by quantum noise. Thus denoising of the mammograms is essential for further processing. Since the mammograms are low contrast images and the lesions are in subtle nature, they have to be enhanced prior to segmentation. Hence, the detection accuracy of early signs of breast cancer (Tang *et al*., 2009) could be improved by denoising and subsequent enhancement of mammographic images. Preprocessing is performed in three steps, namely, denoising, contrast enhancement and removal of label and pectoral muscle.

In this research, the mammograms are denoised by applying 2 Dimensional (2D) median filter with a 3×3 mask (Nagi *et al*., 2010). The denoised images are then subjected to Contrast Limited Adaptive Histogram Equalization (CLAHE) for enhancement. In CLAHE user specified maximum is imposed on the the local histogram height so as to reduce over-enhancement of noise and edge shadowing effect (Pisano *et al*., 1998) in contrast to Adaptive Histogram Equalization (AHE) where a pixel's intensity value is transformed to a value proportional to the pixel intensity's rank in the histogram of a local region. Labels are small regions in the image and provide information regarding the mammogram projections, scanning equipment etc. It does not contribute anything in the classification of masses. Removal of labels from the images gives better result for the analysis. They are removed on the basis of area constraint of the label. Initially, the individual components in the image are labeled after thresholding. Threshold values used were chosen by looking up the intensity values of the labels on the displayed images individually. Then area is calculated for each labeled component. The breast part has the largest area compared to all other individual components in the image. Based on this area constraint, the labels with smaller area are blacked out (Nagi *et al*., 2010). The pectoral muscle lies on the left or right edge of the image. It has an inverted triangle structure of brightest pixels. It must be removed before detecting the masses. Initially, the image is converted to binary image to find the pectoral boundary. To detect the position of the pectoral muscle, nonzero pixels are searched from the left and right top corner of the binary image. If the right width is smaller than the left width, the pectoral muscle is on the right side of the image else it is on the left side of the image. The start and end point of the detected pectoral muscle are joined by a line. The area above the line has the pectoral muscle. Then it is removed to get the required region for the analysis.

After preprocessing, the suspicious part is segmented from the mammogram by applying fuzzy c means clustering based on the results in (Sudharsan, 2013). In this method, each point has a degree of belonging to particular cluster as in fuzzy logic, instead of belonging completely too just one cluster. The points on the edge of a cluster may have lesser degree than points in the centre of cluster. Any point x has a set of coefficients measuring the degree of being in the k^{th} cluster and centroid of a cluster, mean of all points, weighted by their degree of belongingness to the cluster. The INbreast data set do not contain any label and also they are much poor in contrast. Thus without enhancement, nothing is visible. Hence, histogram equalization is performed first to improve the contrast, then pectoral muscle is removed and finally tumor part is segmented.

Feature extraction: Texture feature descriptors based on Grey level co-occurrence matrix, k-Gabor method, Grey level run length matrix, Grey level difference method, fractal and steerable pyramids were extracted from segmented regions.

Grey level co-occurrence method: Statistical texture features are computed from the statistical distribution of experimental combinations of intensities in the image matrix at specified positions relative to each other. According to the number of intensity levels in each combination, they are classified into first-order, second order and higher-order statistics. The Grey-Level Co-occurrence Matrix (GLCM) (Mohanaiah *et al*., 2013) is a second-order statistical texture measure which is a relation between the intensity values of neighboring pixels. It is also defined as a two dimensional histogram of gray levels for a pair of pixels, separated by a fixed spatial relationship. It is a matrix with the number of rows and columns is equal to the number of gray levels in the image. An element P (i, j | Δx, Δy) in this matrix gives the relative frequency of two pixels with intensity 'i' and intensity 'j', separated by a pixel distance (Δx, Δy) within a given neighborhood, The element P (i, j | d, θ) is the second order statistical probability values for changes between the intensity values 'i' and 'j' at a displacement distance d and at a particular angle (θ). GLCM is also called as Gray level Dependency Matrix. Fifteen statistical measures namely autocorrelation, contrast, correlation, correlation, cluster prominence, cluster shade, dissimilarity, energy, entropy, homogeneity, homogeneity, maximum probability, sum of squares, sum average, sum variance, sum entropy, difference variance, difference entropy, information measure of correlation1 and 2, inverse difference, inverse difference normalized and inverse difference moment were computed.

k-Gabor method: The Gabor filters are used in combination with k-means clustering for extracting k-Gabor features (Mamani *et al*., 2012). Gabor filters are tunable band pass filters with multi-scale, multi-resolution and have selectivity for orientation, spectral bandwidth and spatial extent. Gabor filters work with rotations and scales and they provide a means to the texture analysis of the internal structures in any level of the original image (Cope *et al*., 2010). Clustering the image regions takes the advantage of the shape characteristic of the regions. Thus combining Gabor filters with k-means clustering method integrates texture and shape features of images and provides the most salient components of an image. Hence, Gabor filters are employed to quantify the texture information from specific regions, tissues and internal structures of the images providing a concise representation for a richer image analysis. The k-Gabor feature extractor comprises two stages: each original grayscale image is clustered using the k-Means algorithm and a set of k new images are obtained for analysis;gabor features are extracted from each image generated in the first stage. Finally, all the features extracted from the set of clustered images compose the final feature vector, building the k-Gabor elements.

Grey level run length method: Galloway introduced the run-length measures to represent the texture properties (Galloway, 1975). Grey-Level Run-Length Matrix (GLRLM) is a matrix from which perform texture analysis by extracting some texture features. It is based on computing the number of gray level runs of various lengths, ie it is a way of searching the image for runs of pixels having the same gray level value, always across a given direction. Run length is the number of neighboring pixels having the same grey intensity in a particular direction. The length of the run is the number of pixel points in the run (Suganthi and Madheswaran, 2010). GLRLM is a two-dimensional matrix where each element is the number of elements j in the direction θ with the intensity I, (Radhakrishnan *et al*., 2012). Many different run-length matrices may be computed for a single image, one for each chosen direction. In addition to the 0° direction, GLRLM can also be formed in the other directions, i.e. 45°, 90° or 135°. From run-length matrix it is possible to obtain several indicators such as Short Run Emphasis (SRE), Long Run Emphasis (LRE), Gray Level Nonuniformity (GLN), Run Percentage (RP), Low Grey-Level Run Emphasis (LGRE), High Grey-Level Run Emphasis (HGRE), Run Length Nonniformity (RLN), Short Run Low Grey Level Emphasis (SRLGE) and Short Run High Grey Level Emphasis (SRHGE) as explained in (Tang ,1998; Nanni *et al*., 2013).

Grey level difference method: Grey Level Difference Method (GLDM) method uses the probability distribution function of gray level difference between two nearby pixels in an image to compute texture features (Nanni *et al*., 2013; Tang, 1998). GLDM seeks to extract texture features that describe the size and distinction of textural elements in an image. For an image I(x, y), any given row and column displacement $\delta = (\Delta X, \Delta Y)$ let the gray level difference is $I\delta(x, y) = |I(x, y) - I(X + \delta X, Y + \Delta Y)|$. $P_\delta(i)$ be the probability distribution function of $I\delta$ (x, y) is obtained from the number of times $I\delta$ (x, y) occurs for a given δ, i.e. $P_\delta(i) = P(I\delta(x, y) = i)$. From $P_\delta(i)$ several statistical texture features can be computed. If a texture is directional, the degree of spread of the values in $P_\delta(i)$ also vary with the direction of d, given that its magnitude is in the proper range. Thus, comparison of spread measures of $P_\delta(i)$ for various directions of d provides a way to analyze texture directionality. In this research, 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 0°, 45° ,90° and 135° direction respectively. And then mean, variance, entropy, skewness and kurtosis were computed from each $P_\delta(i)$.

Fractal: Fractals are a class of mathematical functions that express the geometrical properties of sets was introduced in 1982 by Mandelbrot which are self-similar and irregular in nature. Fractals are of rough geometric shapes which can be subdivided in small parts, each one of which is reduced to similar of the whole. In classical Euclidean geometry, objects are defined by parametric equations specifying their surface or volumes. The objects have integer dimensions or topological dimensions T, for example, a point is referred by zero dimension, line by one dimension, an area by two dimensions and a solid by three dimensions, whereas in fractal geometry, objects possess a non-integer or fractional dimension D which greater than topological dimension (Welstead, 1999). The topological dimension (defined as d) of an object would not change whatever be the transformation an object undergoes. The fractal dimension is an important characteristic of fractals because it has got information about their geometric structure and describes the roughness or texture. As fractal dimension increases the irregularity of object increases, ie., Lower the value of D value the smoother the object, higher the value of D the rougher the object. A fractal can be classified according to their degree of irregularity. As malignant tumors are irregular structures, the measurement of irregularity can be obtained by calculating fractal dimension which can also be used as a quantitative measure to differentiate benign tumors from malignant tumors. This similarity between fractals and tumors suggests that the calculation of fractal dimension can be used as a means of classification of mammograms (Sedivy *et al*., 1999; Crisan *et al*., 2007).

Steerable pyramids method: The steerable pyramid is a multi-scale, translation-invariant and multi-orientation with rotation-invariant representation. It is an implementation of a multi-scale, band-pass filter bank used for applications including image compression, texture synthesis and object recognition. William, Freeman and Adelson introduced the concept of steerability for oriented filters. Steerable filter refers to a class of filters in which a filter of arbitrary orientation is synthesized as linear combination of a set of basis filters. The basis/projection functions are oriented (steerable) filters, localized in space and frequency. Also, it is overcomplete to avoid aliasing and "self-inverting" (like the QMF/Wavelet transform): the projection and basis functions are identical. The mathematical phrase for a transform obeying this property is "tight frame". It is a more efficient approach to calculate the response of a filter at a different orientation rather than having a large number of filters, one for each orientation. Steerable means that the function can be written as a linear combination of rotated copies of itself. The authors discuss and derive the conditions necessary to "steer" a given filter and derive various theorems. They emphasize that all functions that are bandlimited in angular frequency are steerable, given enough basis filters.

Feature reduction: There is a maximum number of features for a given sample size above which the performance of a classifier will be degraded instead of improving. This is known as the curse of dimensionality. To overcome this difficulty, dimensionality reduction might be performed on the actual feature set. Generally, two approaches are followed to perform dimensionality reduction. First one is feature extraction: creating a subset of new features by combinations of the existing features and the second approach is feature selection: choosing a subset of all the features. Principal component analysis, a statistical procedure for feature transformation is used to reduce the number of features.

The objective of Principal component analysis is to perform dimensionality reduction while preserving as much of the randomness in the high-dimensional space as possible (Subasi and Gursoy, 2010). In PCA, the optimal approximation of an N dimensional random vector x_N is a linear combination of M ($M<N$) independent vectors which are obtained by projecting the random vector x onto the eigenvectors corresponding to the largest M eigenvalues of the covariance matrix. This reduces the sum square magnitude of the approximation error. The PCA does not consider the class label of the feature vector and hence class separability is not considered. PCA just performs a coordinate rotation that aligns the transformed axes with the directions of maximum variance.

Classification: Classification of the mammograms into benign and malignant is performed by a Multilayer Perceptron neural (MLP) network with single hidden layer using back propagation learning algorithm (Fausett, 1993). A network with a single layer can approximate any function, if the hidden layer is large enough. In this research a MLP with single hidden layer of 10 hidden neurons is used. From classifier output confusion matrix are derived. The confusion matrix or error matrix is one way to summarize the performance of a classifier for binary classification tasks. This square matrix consists of columns and rows that list the number of instances as actual class vs. predicted class ratios. In general, positive means identified and negative means rejected. The confusion matrix and the definitions used with respect to breast cancer diagnosis are tabulated in.

Definations: True positive False PositiveType I error Positive Predictive Value (PPV) False Discovery Rate (FOR) = (1-PPV)

True and false: False negativeType II error True Negative Negative Predictive Value (NPV) False Omission Rate (FOR) = (1-NPV).

Confusion matrix: True Positive Rate(TPR)/ Sensitivity False Negative Rate (FNR) = 1-TPR True Negative Rate(TNR)/ SpecificityFalse Positive Rate (FPR) =1-TNR AccuracyPrecision Error

Marina Sokolova, Nathalie Japkowicz and Stan Szpakowicz reported that higher accuracy alone does not guarantee overall better performance of an algorithm. The same conclusion applies to every performance measure if it is considered separately from others. On the other hand, a combination of measures gives a balanced evaluation of the algorithm's performance. Hence, there are 15 performance metrics for the classification are computed based on (Sokolova *et al*., 2006).

Accuracy is the percent of correct classifications. Precision error is the percent of incorrect classifications. Sensitivity and specificity approximates the probability of the positive and negative labels respectively being true and hence assess the effectiveness of the algorithm on a single class. Accuracy, sensitivity and specificity show how effectively a classifier identifies the data labels. Precision estimates the predictive value or the class agreement of the data labels, either positive or negative; depending on the class for which it is calculated i.e., it assesses the predictive power of the classifier. It is also called the Positive Predictive Value (PPV). Precision can be thought of as a measure of a classifier's exactness. A low precision indicates a large number of false positives. Recall is a measure of a classifier's completeness. A low recall indicates many False Negatives. F-scoreorF-measure is a measure of a test's accuracy. It considers both theprecisionand therecallof the test to compute the score and conveys the balance between the precision and the recall. Negative Predictive Value (NPV) gives the proportion of actual negatives in the population being tested. False Discovery Rate (FDR) shows the measure of proportion of actual positive outcomes with respect to total positive results. Similarly False Omission Rate (FOR) defines the measure of proportion of actual negative results with respect to total negative results.

In addition to the above basic performance metrics, Matthews Correlation Coefficient (MCC), Youden's index, likelihoods, AUC, Discriminant power (DP) and kappa statistics (κ) are also computed for performance analysis. Youden's index, likelihoods and Discriminant Power combine sensitivity and specificity and their complements.

Matthews's Correlation Coefficient (MCC) is first formulated by Brian W. Matthews in 1975 to assess the performance of protein secondary structure predictions. The MCC is useful in unbalanced class settings and bounded between the range 1 and -1, where 1 indicates perfect correlation between ground truth and predicted outcome, -1 indicates inverse or negative correlation and a value of 0 denotes a random prediction. Youden's index γ is a measure of avoidance of failure and hence it complements accuracy or the ability to correctly label examples.

It evaluates the algorithm's ability to avoid failure. It equally weighs the algorithm's performance on positive and negative examples. A higher value of $\tilde{a}$ indicates better ability to avoid failure. If a measure accommodates both sensitivity and specificity but treats them separately, then the classifier's performance can be evaluated to finer degree with respect to both classes. Positive and negative likelihood assess prediction ability on positive and negative classes respectively. A higher positive likelihood and a lower negative likelihood indicate better performance on positive and negative classes respectively. AUC:acc is more reliable than accuracy when used for classifier's assessment on imbalanced data. Discriminant Power evaluates how well an algorithm distinguishes between positive and negative examples. The algorithm is a poor discriminant if DP<1, limited if DP<2, fair if DP<3 and good in other cases.

Kappa Statistic is a measure of the agreement between the predicted and the actual classifications in a dataset (Wongpakaran *et al*., 2013). It is expected to be a higher value for a classifier which has more overlapping predictions and observations. Cohen's Kappa is commonly used as a measure of agreement between the two individuals. It was calculated using the Equation:

$$K=\frac{P-e(K)}{1-e(K)} \tag{1}$$

where, p is the overall percent agreement and is obtained by the following equation:

$$p=\frac{A+D}{N} \tag{2}$$

A = The number of times both raters classify a subject into category 1
D = The number of times both raters classify a subject into category 2
N = The total sample size
e(K) = The chance agreement probability

$$e(K)=\left(\frac{A1}{N}*\frac{B1}{N}\right)+\left(\frac{A2}{N}*\frac{B2}{N}\right) \tag{3}$$

RESULTS AND DISCUSSION

Features were extracted from DDSM and INbreast data sets by individually applying each one of the six

Table 1: Performance metrics computed for DDSM

Feature Extraction Technique	ACC	ERR	TPR/ SEN / REC	TNR/ SPC	FPR (1-SPC)	FNR (1-TPR)	PPV	NPV	FDR (1-PPV)	FOR (1-NPV)	LR+	LR-	F1	MCC	DP	YI	AUC: ACC	κ
Fractal	85.3	14.7	83.7	89.5	10.5	16.3	95.3	68	4.7	32	7.971	0.182	0.891	0.681	0.905	172.200	86.600	0.667
GLCM	88.2	11.2	90.7	84	16	9.3	90.7	84	9.3	16	5.669	0.111	0.907	0.747	0.943	173.70	87.350	0.747
Gabor	88.2	11.8	88.9	87	13	11.1	93	80	7.0	20	6.838	0.128	0.909	0.744	0.954	174.900	87.950	0.743
GLRLM	89.7	10.3	90.9	87.5	12.5	9.1	93	84	7.0	16	7.272	0.104	0.920	0.777	1.018	177.400	89.200	0.777
GLDM	89.7	10.3	90.9	87.5	12.5	9.1	93	84	7.0	16	7.272	0.104	0.920	0.777	1.018	177.400	89.200	0.777
Steerable Pyramid	89.7	10.3	90.9	87.5	12.5	9.1	93	84	7.0	16	7.272	0.104	0.920	0.777	1.018	177.400	89.200	0.777
Combined Features	92.6	7.4	93.2	91.7	8.3	6.8	95.3	88	4.7	12	11.23	0.074	0.943	0.841	1.203	183.900	92.450	0.841

Table 2: Performance metrics computed for Inbreast

Feature Extraction Technique	ACC	ERR	TPR/ SEN / REC	TNR/ SPC	FPR (1-SPC)	FNR (1-TPR)	PPV	NPV	FDR (1-PPV)	FOR (1-NPV)	LR+	LR-	F1	MCC	DP	YI	AUC: ACC	κ
Fractal	82.8	17.2	75.9	88.6	11.4	24.1	84.6	81.6	15.4	18.4	6.658	0.272	0.800	0.653	0.766	163.500	82.250	0.650
GLCM	86.8	13.2	87	86.4	13.6	13	93	76	7	24	6.397	0.150	0.899	0.711	0.898	172.400	86.700	0.708
Gabor	85.9	14.1	87	85.4	14.6	13	76.9	92.1	23.1	7.9	5.959	0.152	0.816	0.707	0.879	171.400	86.200	0.703
GLRLM	82.8	17.2	80	84.6	15.4	20	76.9	86.8	23.1	13.2	5.195	0.236	0.784	0.642	0.740	163.600	82.300	0.642
GLDM	87.5	12.5	87.5	87.5	12.5	12.5	80.8	92.1	19.2	7.9	7.000	0.143	0.840	0.739	0.932	174.000	87.500	0.738
Steerable Pyramid	84.4	15.6	86.4	83.3	16.7	13.6	73.1	92.1	26.9	7.9	5.174	0.163	0.792	0.674	0.828	168.700	84.850	0.668
Combined Features	89.1	10.9	88	89.7	10.3	12	84.6	92.1	15.4	7.9	8.544	0.134	0.863	0.772	0.996	176.700	88.850	0.772

texture feature techniques discussed above. From DDSM database 24 benign and 40 malignant cases were considered. In INbreast dataset 36 benign and 28 malignant cases were taken into consideration. The classification was performed by using each one of the individual features computed from both DDSM and INbreast datasets after feature reduction by Principal Component Analysis (PCA). From each case only ten PCAs are used for classification. Also, the images are classified through the combination ie. Concatenation of all seven features. There are 18 performance metrics were computed from the confusion matrix obtained from classification. The computed performance metrics are tabulated in Table 1and 2.

From Table 1 and 2, accuracy of the classification is improved to 92.6 and 89.1% for DDSM and INbreast datasets respectively. At the same time the precision error was reduced. Sensitivity and specificity are improved 93.2 and 91.7% for DDSM data set and 88% and 89.7% for INbreast with combined feature set. FPR, FNR, FDR and FOR are reduced when the features are combined. By combining the features, positive likelyhood is improved and negative likelyhood is reduced in both datasets. Also the results show that Fi measure, MCC, Discriminant Power (DP), Youden's index and kappa statistics have also been improved for the classification with combined features in both DDSM and INbreast datasets. Hence from the results it is concluded that the combination of texture features improves the performance of classification of mammograms.

CONCLUSION

In this research concatenation of various texture features is used to improve the diagnostic performance of breast cancer. This improved performance was obtained due to the fact that any details missed in one type of feature might be captured by other type of texture feature. At the same time, the PCA used here is also used to for improving the classification results. Evolutionary algorithms may be considered in future for feature reduction to further improvement in results. A simple single hidden layer MLP with back propagation learning algorithm was employed for classification. In future, an evolutionary based weight optimization such as GA, Biogeography Based Optimization (BBO) may be used to take the advantage of randomness in learning. Also, hardware implementation of the proposed detection method may be considered in future.

ACKNOWLEDGEMENTS

The researcher thank the Management and the Principal of Kumaraguru college of Technology, Coimbatore for providing the facilities. Also, they are very much thankful to the courtesy of TM Deserni, Department of Medical Informatics and RWTH Achen, Germany for providing the DDSM database and to the courtesy of Breast Research Group, INESC Porto, Portugal for providing their Inbreast data sets to this research work.

REFERENCES

Barley, A. and C. Town, 2014. Combinations of feature descriptors for texture image classification. J. Data Anal. Inf. Process., 2: 67-76.

Belle, A., M.A. Kon and K. Najarian, 2013. Biomedical informatics for computer-aided decision support systems: A survey. Sci. World J., 2013: 1-8.

Bhanumathi, R. and G.R. Suresh, 2013. Latest advances in computer-aided detection of breast cancer by mammography. Int. J. IT Eng., 1: 1-13.

Cheng, H.D., X.J. Shi, R. Min, L.M. Hu, X.P. Cai and H.N. Du, 2006. Approaches for automated detection and classification of masses in mammograms. Pattern Recognit., 39: 646-668.

Cope, J.S., P. Remagnino, S. Barman and P. Wilkin, 2010. Plant texture classification using gabor co-occurrences. Proceedings of the International Symposium on Visual Computing, November 29-December 1, 2010, Springer, Berlin, Germany, ISBN:978-3-642-17273-1, pp: 669-677.

Crisan, D.A., R. Dobrescu and P. Planinsi, 2007. Mammographic lesions discrimination based on fractal dimension as an indicator. Proceedings of the14th International Workshop on Systems, Signals and Image Processing and 6th EURASIP Conference focused on Speech and Image Processing of Multimedia Communications and Services, June 20-30, 2007, IEEE, New York, USA., ISBN: 978-961-248-036-3, pp: 74-77.

Fausett, L., 1993. Fundamentals of Neural Networks US Edn., Prentice Hall, USA., ISBN: 0133341860, pp: 169-195.

Freeman, W.T. and E.H. Adelson, 1991. The design and use of steerable filters. IEEE Trans. Pattern Anal. Mach. Intell., 13: 891-906.

Galloway, M.M., 1975. Texture analysis using gray level run lengths. Comput. Graph. Image Process., 4: 172-179.

Jalalian, A., S.B. Mashohor, H.R. Mahmud, M.I.B. Saripan and A.R.B. Ramli *et al.*, 2013. Computer-aided detection-diagnosis of breast cancer in mammography and ultrasound: A review. Clin. Imaging, 37: 420-426.

Liao, S., M.W.K. Law and A.C.S. Chung, 2009. Dominant local binary patterns for texture classification. IEEE Trans. Pattern Anal. Mach. Intell., 18: 1107-1118.

Liasis, G., C. Pattichis and S. Petroudi, 2012. Combination of different texture features for mammo graphic breast density classification. Proceedings of the IEEE 12th International Conference on Bio Informatics and Bio Engineering (BIBE), November 11-13, 2012, IEEE, New York, USA., ISBN: 978-1-4673-4357-2, pp: 732-737.

Mamani, G.H., A.J. Traina and C. Traina, 2012. K-Gabor: A new feature extraction method for medical images providing internal analysis. Proceedings of the 25th International Symposium on Computer-Based Medical Systems (CBMS), June 20-22, 2012, IEEE, New York, USA., ISBN: 978-1-4673-2049-8, pp: 1-6.

Mohanaiah, P., P. Sathyanarayana and L. GuruKumar, 2013. Image texture feature extraction using GLCM approach. Int. J. Sci. Res. Publ., 3: 1-5.

Moreira, I.C., I. Amaral, I. Domingues, A. Cardoso and M.J. Cardoso et al., 2012. INbreast: Toward a full-field digital mammographic database. Acad. Radiol., 19: 236-248.

Nagi, J., S.A. Kareem, F. Nagi and S.K. Ahmed, 2010. Automated breast profile segmentation for ROI detection using digital mammograms. Proceedings of the EMBS Conference on Biomedical Engineering and Sciences (IECBES), November 30-December 2, 2010, IEEE, New York, USA., ISBN: 978-1-4244-7599-5, pp: 87-92.

Nanni, L., S. Brahnam, S. Ghidoni, E. Menegatti and T. Barrier, 2013. Different approaches for extracting information from the co-occurrence matrix. PloS One, Vol. 8,

Oliver, A., J. Freixenet, J. Marti, E. Perez, J. Pont, E.R.E. Denton and R. Zwiggelaar, 2010. A review of automatic mass detection and segmentation in mammographic images. Med. Image Analysis, 14: 87-110.

Pisano, E.D., S. Zong, B.M. Hemminger, M. DeLuca and R.E. Johnston et al., 1998. Contrast limited adaptive histogram equalization image processing to improve the detection of simulated spiculations in dense mammograms. J Digital Imaging, 11: 193-200.

Radhakrishnan, M., T. Kuttiannan and N. Tiruchengode, 2012. Comparative analysis of feature extraction methods for the classification of prostate cancer from TRUS medical images. IJCSI. Int. J. Comput. Sci., 9: 171-179.

Rangayyan, R.M., F.J. Ayres and J.E. Leo Desautels, 2007. A review of computer-aided diagnosis of breast cancer: Toward the detection of subtle signs. J. Franklin Inst., 344: 312-348.

Sabu, A.M. and D.N. Ponraj, 2012. Textural features based breast cancer detection: A survey. J. Emerg. Trends Comput. Inf. Sci., Vol. 3,

Sedivy, R., C. Windischberger, K. Svozil, E. Moser and G. Breitenecker, 1999. Fractal analysis: An objective method for identifying atypical nuclei in dysplastic lesions of the cervix uteri. Gynecologic Oncol., 75: 78-83.

Sokolova, M., N. Japkowicz and S. Szpakowicz, 2006. Beyond accuracy F-score and ROC: A family of discriminant measures for performance evaluation. Proceedings of the Australasian Joint Conference on Artificial Intelligence, December 4-8, 2006, Springer, Berlin, Germany, ISBN: 978-3-540-49787-5, pp: 1015-1021.

Subasi, A. and M.I. Gursoy, 2010. EEG signal classification using PCA, ICA, LDA and support vector machines. Expert Syst. Appl., 37: 8659-8666.

Sudharsan, P., 2013. A comparitive analysis of segmentation techniques in mammogram images. Int. J. Eng. Res. Technol., Vol. 2,

Suganthi, M. and M. Madheswaran, 2012. An improved medical decision support system to identify the breast cancer using mammogram. J. Med. Syst., 36: 79-91.

Tang, J., R.M. Rangayyan, J. Xu, I. El Naqa and Y. Yang, 2009. Computer-aided detection and diagnosis of breast cancer with Mammography: Recent advances. IEEE Trans. Inform. Technol. Biomed., 13: 236-251.

Tang, X., 1998. Texture information in run-length matrices. IEEE. Trans. Image Process., 7: 1602-1609.

Welstead, S., 1999. Fractal and Wavelet Image Compression Techniques. 1st Edn. SPIE Optical Engineering Press, Washington.

Wongpakaran, N., T. Wongpakaran, D. Wedding and K.L. Gwet, 2013. A comparison of Cohen?s Kappa and Gwet?s AC1 when calculating inter-rater reliability coefficients: A study conducted with personality disorder samples. BMC. Med. Res. Method., 13: 13-61.