Breast Cancer Classification using Global Discriminate Features in Mammographic Images

Nadeem Tariq[1], Beenish Abid[2], Khawaja Ali Qadeer[3], Imran Hashim[4], Zulfiqar Ali[5], Ikramullah Khosa[6]

The Department of CS and IT, The University of Lahore 1km off Defence Road, Lahore, Pakistan[1, 2, 3]
The Department of Mathematics & Statistics, The University of Lahore 1km off Defence Road, Lahore, Pakistan[4]
The Department of CS and IT, The University of Lahore 1km off Defence Road, Lahore, Pakistan[5]
Department of Electrical Engineering, COMSATS University Islamabad Lahore Campus, Lahore, Pakistan[6]

Abstract—Breast cancer has become a rapidly prevailing disease among women all over the world. In term of mortality, it is considered to be the second leading cause of death. Death risk can be reduced by early stage detection, followed by a suitable treatment procedure. Contemporary literature shows that mammographic imaging is widely used for premature discovery of breast cancer. In this paper, we propose an efficient Computer Aided Diagnostic (CAD) system for the detection of breast cancer using mammography images. The CAD system extracts largely discriminating features on the global level for representation of target categories in two sets: all 20 extracted features and top 7 ranked features among them. Texture characteristics using co-occurrence matrices are calculated via the single offset vector. Multilayer perceptron neural network with optimized architecture is fed with individual feature sets and results are produced. Data division corresponds as 60%, 20%, and 20% is used for training, cross-validation, and test purposes, respectively. Robust results are achieved and presented after rotating the data up to five times, which shows higher than 99% accuracy for both target categories, and hence outperform the existing solutions.

Keywords—Breast cancer; mammography; pattern recognition; classification

I. INTRODUCTION

The death rate in women due to breast cancer is high. According to the American cancer society, about 178,000 cases of breast cancer diagnosed, and 41,000 women expire due to this disease each year in the United States. In developing countries, breast cancer patient's ratio is increasing since 2000. According to an estimate, 12 million people will pass away due to breast cancer in 2030 [1]. In Asia, Pakistan has the highest rate of breast cancer cases which causes the death of nearly 40,000 women in Pakistan every year [2]. According to WHO (World Health Organization), 450,000 patients die each year worldwide due to breast cancer. Mortality rate due to breast cancer can be cut by the help of an efficient screening method at the earlier stage of cancer, before the appearance of major physical symptoms. The leading measure for screening involves taking X-Ray of breast region called a mammogram. The mammogram is very effective for initial diagnosis since it is capable of detecting a small change in the tissues which are sometimes too small to be felt by a doctor or the patient herself. Such a small change may indicate the presence of cancer [3-4]. Most commonly used techniques to diagnose breast cancer are mammography,

biopsy, thermography, and ultrasound imaging [5-7]. A biopsy is a standard clinical approach used to diagnose cancer at initial stage under a microscope, however, this approach is complex, costly as well as time-consuming. The medical expert after examining the mammogram may suggest a biopsy in case any abnormality is found. Due to the subjective nature of human interpretation, the radiologists may have different opinions on similar mammograms. A false negative diagnosis at this stage may lead to serious consequences for the patient. In case of no treatment after a malignant tumor is detected, infected cells spread to another part of the body and ultimately cause death [8]. On the contrary, a false positive interpretation may suggest an unnecessary biopsy, and so leading to a redundant painful procedure.

Development of an efficient CAD (Computer Aided Diagnosis) system would help the pathologist in determining the need for biopsy as it'll provide aid to enhance confidence to manual diagnosis. The proposed system will categorize the test sample as Benign (no-cancer) or Malignant (cancer) by estimating the probability of cancer in the patient via examining the mammographic image of the breast region. The proposed system characterizes a modest selection of features for class representation and careful selection of classification strategy. Such a scheme is a potential candidate for an automatic support system along with manual diagnosis for early detection of the presence of cancer.

II. RELATED WORK

Sharanya Padmanabhan in [9] proposed a CAD system for enhancement of Breast Cancer detection in digital mammogram by employing wavelet transform for feature extraction. The system was developed using the MATLAB tool and Mini MIAS database was used for testing. This work claimed accuracy of 75.3% for normal and 92.3% for malignant. Rehman, Chouhan, & Khan [10] used Digital Database for Screening Mammography (DDSM) dataset with six statistical features: standard deviation, third momentum, mean, randomness, smoothness, and uniformity. In this research, texture features were extracted using Local Binary Pattern (LBP). Feature vectors of Region of Interests (ROI) were obtained from taxonomic indices that were based on phylogenetic trees. Local binary patterns and statistical measures were used to train the SVM (Support Vector Machine) classifier for binary classification. Maximum accuracy achieved by using this system on DDSM dataset was

80%. In [11] Nithya, & Santhi calculated Gray Level Co-occurrence Matrices (GLCM) were calculated in four directions (0o, 45o, 90o and 135o) at four distances (1, 2, 3 and 4). Five statistical measures; entropy, energy, the sum of square variance, correlation and homogeneity were computed from GLCMs. A three-layer Artificial Neural network (ANN) was used as a classifier. In this CAD system, an experiment was conducted on DDSM dataset: network trained using 200 mammograms and tested with 50 mammograms. The maximum classification accuracy achieved by using this system was 96% whereas sensitivity and specificity rate was 100% and 93% respectively.

Mohanty, Swain, Champati, & Lenka in [12] proposed a system using Mini MIAS dataset consisting of 322 mammograms. A total of 26 features including histogram features and GLCM features were calculated. Oscillating search for features selection was a new approached that was proposed in this work to select optimal features from the given feature's subspace. Selected features were used for classification. An accuracy of 97.7% was achieved by using this model. Xie, Li, & Ma, in [13] presented a CAD system for the diagnosis of breast cancer that was based on the Extreme Learning Machine (ELM). A level set function was proposed in this work for image segmentation. Significant features were extracted by combining ELM and support vector machine. The system achieved an average accuracy of 96.02% by using mini MIAS and DDSM databases. The proposed system in [14] by Makandar, & Halalli was based on extracting the suspicious region from the breast. The pre-processing was done to remove the background and pectoral muscle. For image segmentation, region growing method has used that work in two ways: based on pixel values; and selection of seed point that is of two kinds; automatic and manual. In the automatic method, seed point was selected based on histogram on the highest intensity that represents the peak value of the histogram, while in the manual method user selects the seed point. Images were enhanced by using a Wiener filter. Suspicious mass from the mammograms was extracted by using combine techniques of a watershed and active contour-based segmentation. The efficiency of the system was measured using Mini-MIAS database. The reported accuracy was 95.86%.

Using ROI extraction, Jaleel, Salim, & Archana in [15] extracted texture features from mammograms by using Discrete Wavelet Transform (DWT) and GLCM. ANN was used for classifying mammograms into target classes: begin or malignant. System performance was checked with a mini-MIAS database. The accuracy achieved by using this model was 93.7% with GLCM and 94.6% by using DWT features respectively. In [16] DWT was used for features extraction. Normalized features were used with classifiers to categorize the mammograms. The performance was checked with the mini-MIAS database by using K- NN, SVM and Radial basis function neural network (RBFNN) with different texture features for mammograms. RBFNN with DWT features showed better results as compared to K-NN classifier and SVM classifier. The achieved accuracy by RBFNN, K-NN and SVM was 94.6%, 87.8%, and 90.54%, respectively.

III. MATERIALS AND METHODS

The key tasks in developing a CAD system for binary classification of mammograms include image processing, discriminate feature extraction and selection of an appropriate state of the art classifier. Fig. 1 shows the key computational blocks of a CAD system.

A. The Database used for the Experiment

The mini-MIAS database of Mammogram is used in the proposed system that is freely available (Suckling et al., 1994). This data set contains 322 mammograms: 270 sample images are normal (non-cancerous) and the rest 52 samples are malignant (cancerous). Each sample is a 24-bit RGB image with a standard resolution of 1024x1024 pixels. A sample of database images belonging to the target categories is shown in Fig. 2.

As discussed in the previous section, many image processing techniques have been employed by the researchers to analyze the samples and enhance their visual resolution for detection and interpretation of regions of interest. We converted the 24-bit image samples to 8-bit grayscale image and used them for extracting discriminate features. From here, for the purpose of notation, we'll use a positive sample for a malignant category, and negative sample for the benign category.

B. Feature Selection

Feature extraction plays a critical role in pattern detection and classification. Several types of features from images have been investigated and exercised for applications of pattern matching and categorization. Texture characteristics among them are frequently used for representation [10-12, 15-18]. Gray Level Co-occurrence Matrix (GLCM) is the classical and efficient feature matrix, which provides texture analysis of an image [19]. We calculated one GLCM from each sample image at an angle of 0o with an offset distance equal to 1. The size of GLCM is estimated based on existing pixel intensities in the image. From each GLCM (representing the sample image), we calculated 20 texture features as listed in Table 1.
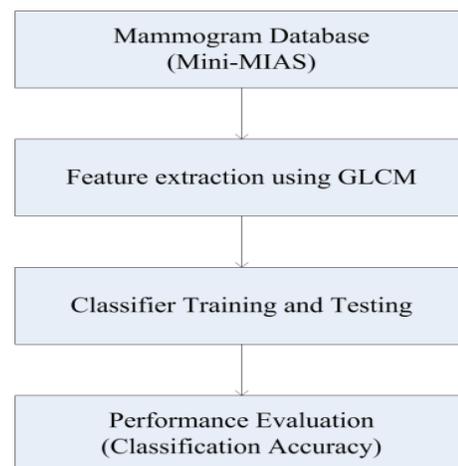

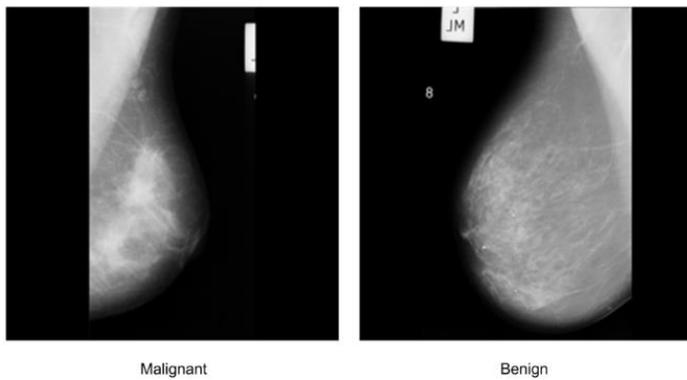Fig. 1. Key Computational Steps Involved in a CAD System (Top to Bottom).

Fig. 2. Sample Mammographic Images with Target Categories (Malignant: Cancer, Benign: No Cancer).

C. Feature Subset Selection

In addition, to considering all extracted features for data classification, we selected fewer among them as a subset of these features also, to take advantage of computational efficiency with significant, lower feature space [20].

Feature reduction is carried out by the feature ranking method where an independent evaluation for all features is carried out for binary classification, and features are sorted based on their significance towards satisfying evaluation criteria. Hence features are sorted from top to bottom according to their contribution for classification. For a lower feature space, features from rank 1 to 7 are selected including information measure of correlation, sum variance, correlation, the sum of square variance, autocorrelation, dissimilarity and sum average respectively. Fig. 3 shows the data plots using the top three ranked features.

TABLE I.   TEXTURE FEATURES EXTRACTED FROM GLCMs

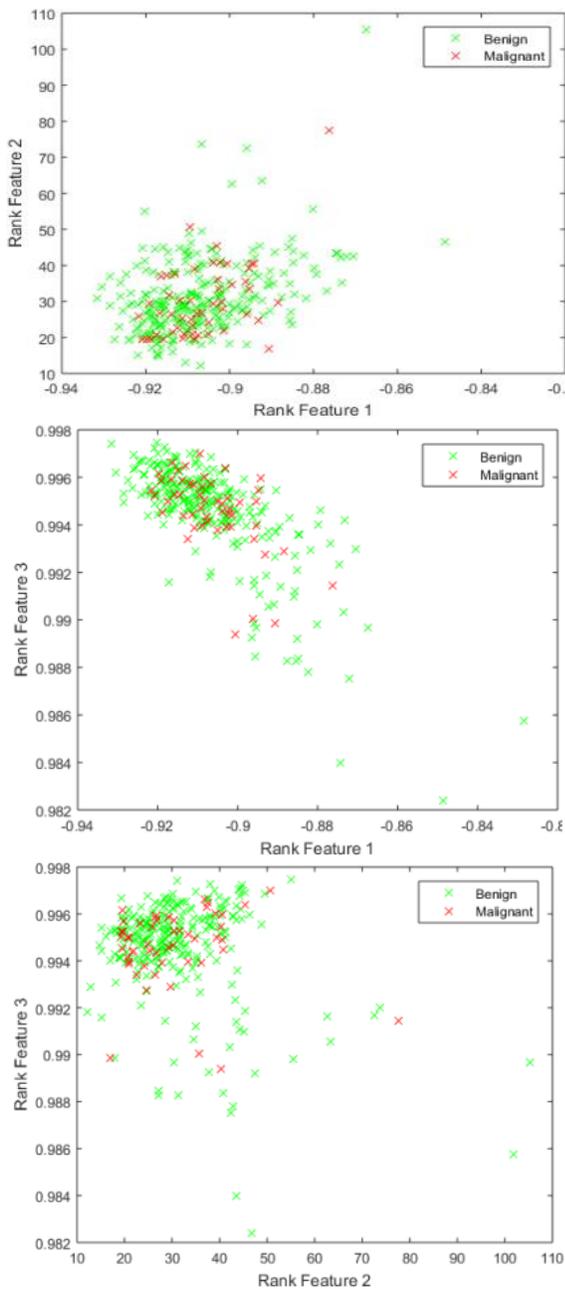| Notation	Name	Description
f1	Autocorrelation	In any time series containing non-random patterns of behavior, it is likely that any the particular item in the series is related in some way to other items in the same series
f2	Contrast	The difference in luminance or color that makes an object distinguishable
f3	Correlation	A single number that describes the degree of relationship between two variables
f4	Cluster Prominence	measure of Asymmetry
f5	Cluster shade	a measure of skewness of the matrix and is believed to gauge the perceptual the concept of uniformity.
f6	Dissimilarity	Variation of grey level pairs in an image.
f7	Energy	Energy returns the sum of squared elements in the GLCM. Energy is also known as uniformity. The range of energy is [0 1].
f8	Entropy	Entropy is a statistical measure of randomness that can be used to characterize the texture of the input image. Entropy can also be used to describe the distribution variation in a region.
f9	Homogeneity	Homogeneity returns a value that measures the closeness of the distribution of elements in the GLCM to the GLCM diagonal.
f10	Maximum Probability	It calculates grey-level having maximum probability in the GLCM.
f11	Sum of Square: Variance	Variance puts relatively high weights on the elements that differ from the average value of p (i,j).
f12	Sum Average	The relation between clear and dense areas in an image.
f13	Sum Variance	It reveals spatial heterogeneity of an image.
f14	Sum Entropy	It is a measure of the sum of micro (local) differences in an image.
f15	Difference Variance	A measure of the local variability.
f16	Difference Entropy	Is a measure of the variability of micro differences.
f17	Information Measure of Correlation1	In this feature two derived arrays are used, the first array represents the summation of rows, while the second one represents the summation of columns in the GLCM.
f18	Information Measure of Correlation2	A feature that is used to calculate mean correlation.
f19	Inverse Difference Normalized	Another measure of the local homogeneity of an image.
f20	Inverse Difference Moment Normalized	Is expected to be large if the grey levels of each pixel pair are similar.

Fig. 3.   Data Plots using the Top Three Ranked Features.

Fig. 4.   Neural Network Functional Diagram.

IV.   CLASSIFICATION

Artificial neural network (ANN) classifier is used for classification in the proposed system as it is state of the art tool for pattern classification and widely used in similar applications [21-26]. A Neural network is composed of simple parallel elements that are inspired by nodes of the biological nervous system. ANN is trained to perform a specific task by adjusting weights between the elements. Such adjustment is based on a comparison with the output and the corresponding target until the network output matches the target. ANN classifier involves two operations: training and testing. A well-trained network is likely to produce better classification accuracy on unseen data. The functional diagram of a neural network is shown in Fig. 4.
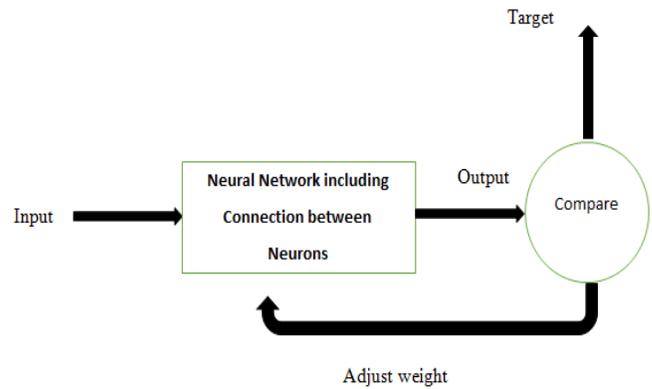
A Multilayer neural network contains an input layer, one or more hidden layers, and an output layer. A network with one hidden layer is sufficient to map any input-output relationship; however, a neural network with multiple hidden layers is often useful for complex mapping. In the proposed system, we used a multilayer neural network with two hidden layers, based on recording meager performance by employing a single hidden layer at first. To estimate the optimized network architecture, we performed a regression analysis between network response and the target value while observing Mean Square Error (MSE). The LM (Levenberg-Marquardt) algorithm is used for learning. Table 2 shows the stats of regression analysis.

The data distribution for the estimation of optimized network architecture, as well as the classifier's performance, is made as 60%, 20% and 20% for training, cross-validation, and test data respectively. The data, however, is rotated up to five times to approximate the robust estimation.

The parameter in the third column in Table 2 'm' represents the slop and 'b' corresponds to y-intercept of the best linear regression that relates target to the network outputs. If output exactly matches the target i.e. perfect fit then the slope would be 1 and the y-intercept would be 0. The third variable 'r' is the correlation coefficient between network output and the target.

Fig. 5 shows the regression analysis for the choice of 22 and 6 as a number of neurons for hidden layer 1 and 2, respectively. Network outcome is plotted versus the target output; the solid line shows the perfect fit and dashed line shows the best linear fit.

Table 3 shows the selection of different combinations of hidden layers' neurons and the respective network performance in terms of average error rate. It shows that the larger the size of the hidden unit of the network, the better the network performs. This is an obvious motivation for adopting a larger number of hidden neurons for better performance. The size, on the contrary, directly relates to the computational efficiency of the network. It is preferred to select the appropriate size based on the estimation of the optimized tradeoff between computational efficiency and classification accuracy. Considering the fact, we estimated 20 and 6 as a number of neurons in the first and second hidden layer respectively.

TABLE II. REGRESSION PARAMETERS' ANALYSIS FOR DIFFERENT COMBINATIONS OF HIDDEN LAYER'S SIZES

Hidden neurons (Layer 1)	Hidden neurons (Layer 2)	m	b	r
5	1	0.3825	0.0997	0.6191
7	1	0.1261	0.1411	0.3551
9	1	0.2318	0.1241	0.4814
12	1	0.5782	0.0690	0.7569
5	2	0.2920	0.1142	0.5414
7	2	0.4637	0.0866	0.6810
9	2	0.4159	0.0943	0.6449
12	2	0.5927	0.0658	0.7699
14	3	0.8867	0.0182	0.9420
16	5	0.9410	0.0073	0.9222
18	5	0.9639	0.0070	0.9521
20	6	0.9542	0.0074	0.9768
22	6	0.9653	0.0069	0.9885

TABLE III. NETWORK MEAN SQUARE ERROR FOR DIFFERENT COMBINATIONS OF HIDDEN LAYERS' SIZES

Hidden neurons (Layer 1)	Hidden neurons (Layer 2)	MSE
5	1	0.10352
7	2	0.0789
9	3	0.04532
11	3	0.0437
15	3	0.1225
16	4	0.02506
18	5	0.01046
20	6	0.00736
22	6	0.00617

These parameters are defined as;

$$Sensitivity = \frac{TP}{TP+FN} X100 \qquad (1)$$

$$Specificity = \frac{TN}{TN+FP} X100 \qquad (2)$$

$$Accuracy = \frac{TP+TN}{TP+TN+FP+FN} X100 \qquad (3)$$

The output from the classifier is compared with the target class to categorize it among True Positive (TP), True Negative (TN), False Positive (FP) and False Negative (FN).

TP = positive sample correctly identified

TN = negative sample correctly identified

FP = negative sample incorrectly identified as positive

FN = positive sample incorrectly identified as negative

Performance is two-way evaluated: 1) using the total extracted features and 2) by using the top seven ranked features (selected as a subset from total features). To ensure the robustness, data is rotated five times, and an average of all outcomes is used as the final classification outcome.

Fig. 5. Network Regression Response for 22:6 Size as Hidden Layer1: Hidden Layer 2.

For the smaller feature space (with seven features), the same procedure is followed i.e. by analyzing the network performance against different combinations of hidden layer sizes described in Table 3. Concretely, the same size of the hidden unit of ANN was estimated as an optimized choice after carrying out the error analysis.

A. Performance Measures

The problem under consideration is binary classification, so a few well-known parameters for evaluating such a classification task are selected including accuracy, specificity, and sensitivity. Specificity measures the proportion of negatives which are correctly identified and Sensitivity measure the proportion of positives which are correctly identified.

V. RESULTS AND DISCUSSIONS

For classification of test (unseen) data, the classifier is employed with estimated architecture and performance is evaluated by using both the total extracted features and the fewer - rank features. As described in the earlier section, to achieve robustness of classifier, the data is rotated each time and results are recorded. Finally, an average of five results was calculated as the final outcome.

Table 4 shows the classification results of the network for different data categories using the total features and the rank features. Using the total extracted features, the results are promising with an overall test accuracy of 99.4%. It showed good performance in identifying both negative and the positive samples by achieving 99.58% and 99.37% sensitivity and specificity rate respectively. Hence the texture characteristics of sample images, calculated from GLCM (which is computed using a single axis only) proved excellent choice as features for this classification task.
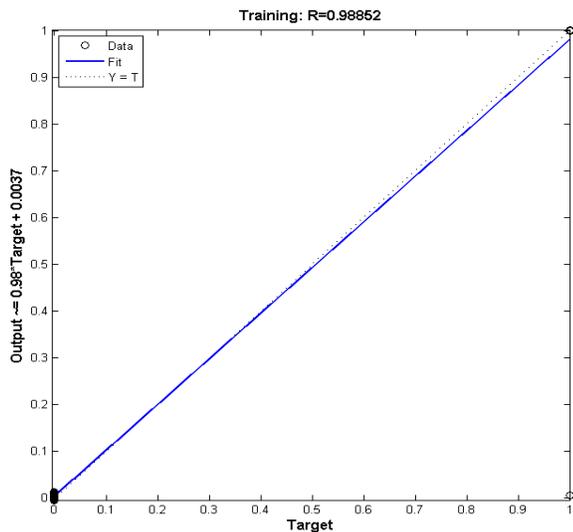
Later, the rank features (fewer significant features) are employed; the network still showed comparable results to those obtained using all features. There is hardly a lack of 0.3% in performance between the two feature set, however, the computational efficiency due to lower feature space is obvious. Considering the unit computational time for the calculation of each feature, 65% of computation time can be saved by compromising merely 0.2% of accuracy. Since the network is trained offline, hence after the robust training accuracy is achieved (as presented in Tab. 4), it will be more than 65% computationally efficient using rank features than the total extracted features, for binary classification of mammographic image data.

On the contrary, the sensitivity rate (rate of correct identification of positive samples) obtained is slightly higher than specificity rate (rate of correct identification of negative samples) for either feature set, as well as both are higher (>99%) which is desired in such classification tasks.

Concretely, the proposed Computer Aided Diagnostic (CAD) system achieved significant accuracy in classifying the mini-MAIS mammographic image database. It achieved incredibly good results with the proposed features and estimated ANN architecture, by showing more than 99% rate in correctly identifying both the target category samples. The obtained results outperform the existing studies by comparing classification accuracy. Table 5 shows a performance comparison of the proposed system with existing studies with different similar databases.

VI. CONCLUSION

In this research work, breast cancer detection is presented using mammographic images. The freely available mini-MIAS mammographic image database is used which contain 322 mammograms in total: 270 are normal and 52 are malignant. A co-occurrence matrix is calculated using each sample and statistical texture features are extracted. Features were then sorted using their individual contribution and a smaller feature set was prepared in addition to the all-feature set. Sixty percent of data was used for training, 20 percent for cross-validation, and the rest 20% for test purposes. An estimated architecture of multilayer neural network, optimized for feature sets, is employed to classify the test data. An average result is produced by rotating the data up to five times. The classifier achieved more than 99% accuracy for identification of benign as well as malignant image samples, using both feature sets and so outperformed previous studies for this database.

TABLE IV.   CLASSIFICATION RESULTS USING INDIVIDUAL FEATURE SETS

Data Category	Sensitivity (%)		Specificity (%)		Accuracy (%)	
	Rank features	*Total Features*	*Rank features*	*Total Features*	*Rank features*	*Total Features*
Training Data	99.62	99.85	99.46	99.78	99.48	99.79
Validation Data	99.23	99.6	98.64	98.81	98.73	98.93
Test Data	99.4	99.58	99.15	99.37	99.2	99.4
Total Data	99.5	99.74	99.23	99.5	99.27	99.54

TABLE V.   PERFORMANCE COMPARISON OF PROPOSED SYSTEM WITH EXISTING SOLUTION

Method	Database	Sensitivity	Specificity	Accuracy
SharanyaPadmanabhan [9]	Mini MIAS	-	-	75.3% (normal) 92.3% (malignant)
Awais [10]	DDSM	-	-	80%
R. Nithya [11]	Mini MIAS	100%	93%	96%
Aswinikumar [12]	Mini MIAS	-	-	97.7%
Weiyingxie [13]	Mini MIAS+DDSM	-	-	96.02%
Makandar[14]	Mini MIAS	-	-	95.86%
Jaleel, J. Abdul [15]	Mini MIAS	-	-	93.7% with GLCM and 94.6% with DWT
Jaleel, J. Abdul, and SibiSalim [16]	Mini MIAS	-	-	Using RBFNN 94.6% K-NN 87.8% and SVM 90.54%
Proposed System (Total features)	Mini MIAS	99.58%	99.37%	99.4%
Proposed System (Rank features)	Mini MIAS	99.4%	99.15%	99.2%

REFERENCES

[1] J. Tang, R. M. Rangayyan, J. Xu, I. El Naqa, & Y. Yang. Computer-aided detection and diagnosis of breast cancer with mammography: recent advances. IEEE Transactions on Information Technology in Biomedicine, 13(2): 236-251, 2009.

[2] http://pinkribbon.org.pk/about.aspx/Accessed 10.04.2016.

[3] B. Verma, & P. Zhang. A novel neural-genetic algorithm to find the most significant combination of features in digital mammograms. Applied soft computing, 7(2): 612-625, 2007.

[4] C. Di Maggio. State of the art of current modalities for the diagnosis of breast lesions. Breast Cancer: 99-126, 2008.

[5] K. M. Kelly, J. Dean, W. S. Comulada, & S. J. Lee. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. European radiology, 20(3): 734-742, 2010.

[6] J. L. Jesneck, J. Y. Lo & J.A. Baker. Breast mass lesions: Computer-aided diagnosis models with mammographic and sonographic descriptors 1. Radiology, 244(2): 390-398, 2007.

[7] H. Zhi et al. Comparison of ultrasound elastography, mammography, and sonography in the diagnosis of solid breast lesions. Journal of ultrasound in medicine, 26(6): 807-815, 2007.

[8] R. Bhanumathi, & G. R. Suresh. Combining trace transform and SVD for classification of micro-calcifications in digital mammograms. In Electronics and Communication Systems (ICECS), 2nd International Conference on. pp. 1381-1386. Feb 2015.

[9] S. Padmanabhan, & R. Sundararajan. Enhanced accuracy of breast cancer detection in digital mammograms using wavelet analysis. In Machine Vision and Image Processing (MVIP), International Conference on (pp. 153-156), Dec 2012. IEEE.

[10] A. U. Rehman, N. Chouhan, & A. Khan. Diverse and Discriminative Features Based Breast Cancer Detection Using Digital Mammography. In Frontiers of Information Technology (FIT), 13th International Conference on (pp. 234-239), Dec 2015. IEEE.

[11] R. Nithya, & B. Santhi. Classification of normal and abnormal patterns in digital mammograms for diagnosis of breast cancer. International Journal of Computer Applications, 28(6): 21-25, 2011.

[12] A. K. Mohanty, S. K. Swain, P. K. Champati, & S. K. Lenka. Image mining for mammogram classification by association rule using statistical and GLCM features. IJCSI, 2011.

[13] W. Xie, Y. Li, & Y. Ma. Breast mass classification in digital mammography based on extreme learning machine. Neurocomputing, 173: 930-941, 2016.

[14] A. Makandar, & B. Halalli. Combined segmentation technique for suspicious mass detection in Mammography. In Trends in Automation, Communications and Computing Technology (I-TACT-15), International Conference on (Vol. 1, pp. 1-5), Dec 2015. IEEE.

[15] J. A. Jaleel, S. Salim, & S. Archana. Textural features based computer aided diagnostic system for mammogram mass classification. In Control, Instrumentation, Communication and Computational Technologies (ICCICCT), International Conference on (pp. 806-811), July 2014. IEEE.

[16] J. Suckling et al. The mammographic image analysis society digital mammogram database. In Exerpta Medica. International Congress Series (Vol. 1069, pp. 375-378), July 1994.

[17] J.A. Jaleel, & S. Salim. Mammogram mass classification based on discrete wavelet transform textural features. In Advances in Computing, Communications and Informatics (ICACCI, 2014 International Conference on (pp. 718-722), Sep 2014. IEEE.

[18] L. K. Soh, & C. Tsatsoulis. Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices. IEEE Transactions on geoscience and remote sensing, 37(2): 780-795, 1999.

[19] R. M. Haralick, & K. Shanmugam. Textural features for image classification. IEEE Transactions on systems, man, and cybernetics, 3(6):610-621, 1973.

[20] S. M. Lai, X. Li, & W. F. Biscof. On techniques for detecting circumscribed masses in mammograms. IEEE Transactions on Medical Imaging, 8(4): 377-386, 1989.

[21] L. M. Mina, & N. A. M. Isa. Breast abnormality detection in mammograms using Artificial Neural Network. In Computer, Communications, and Control Technology (I4CT), International Conference on (pp. 258-263), April 2015. IEEE.

[22] M. M. A. Abdelaal, H.A. Sena, M. W. Farouq, & A. B. M. Salem. Using data mining for assessing diagnosis of breast cancer. In Computer Science and Information Technology (IMCSIT), Proceedings of the International Multiconference on (pp. 11-17), 2010.

[23] H. T. T. Thein, K. M. M. Tun. An approach for breast cancer diagnosis classification using neural network. Advanced Computing, 6(1): p.1, 2015.

[24] A. Moh'd Rasoul, M. Y. Al-Gawagzeh, & B. A. Alsaaidah. Solving mammography problems of breast cancer detection using artificial neural networks and image processing techniques. Indian journal of science and technology, 5(4): 2520-2528, 2012.

[25] A. P. Dhawan, Y. Chitre, & C. Kaiser-Bonasso. Analysis of mammographic microcalcifications using gray-level image structure features. IEEE Transactions on Medical Imaging, 15(3): 246-259, 1996.

[26] A. J. Walker, S. S. Cross, & R. Harrison. Integrated data visualisation and classification using growing cell structure neural networks applied to the diagnosis of breast cancer, 1998.