Middle-East Journal of Scientific Research 23 (6): 1030-1034, 2015
ISSN 1990-9233

DOI: 10.5829/idosi.mejsr.2015.23.06.93102

Breast Cancer Diagnosis Using Wavelet Based Threshold Method

[1]M. Kanchana and [2]P. Varalakshmi

[1]Research Scholar, Anna University, Chennai, India
[2]Madras Institute of Technology, Anna University, Chennai, India

Abstract: This paper presents an automatic support system for detection and stageclassificationofthe cancer regionin its early stagesusingthresholding based segmentation approachand probabilistic neural network. Mammogram (Breast X-ray imaging) isa reliable, cheap and effective method that is used for early detection of breast cancers.In this paper initially, the mammography images are segmentedusingwavelet based threshold method.The proposed system provides valuable information to the radiologists and is helpful in detecting abnormalities faster than the traditional methods. The proposed Computer Aided Diagnosis (CAD) system is tested using Mammography Image Analysis Society (MIAS) database and achieves an accuracy of 92.3%.

Key words: Mammogram Images · Discrete Wavelet Transform · Probabilistic Neural Network · Thresholding

INTRODUCTION

Breast cancer is a leading cause of death amongwomen [1]. Early prevention of this disease is impossible as the cause of the death still remains unidentified. Hence, early detection is significant to improve the diagnosis of breast cancer. Several breast cancers cannot be diagnosed accurately based on the tumors through visual analysis. This is because tumors are of low quality images resulting from the minor differences in X-ray dilution between breast tissues. Thus, it is imperative to develop a fully automatic technique with high sensitivity to assist the early detection of breast cancer with mammogram images. To detect the cancer from the multi resolution of scanned images, a K-means clustering algorithm is implemented [2].

Mammography is the most suitable technique for detecting breast cancer. But mammograms hide lesions or create false alarms as they are projected images. Though automatic analysis of mammograms does not supplant physicians, an accurate computer aided detection technique helps physicians to take reliable decision [3]. Several methods have been proposed for identifying lesions in mammograms such as wavelet based approach [4], neural network analysis [5], a morphological approach [6], fuzzy logic based technique and so on. It is observed that the wavelet based approach provides a better analysis of mammogram images.

This paper presents a probabilistic neural network for the classification of the breast cancerbased on discrete wavelet transform using digital mammogram images. The discriminating features extracted using Grey Level Co-occurrence Matrix (GLCM) is fed into the classifier for improving the classification accuracy.

The rest of the paper is organized as follows: section 2 presents the related work, in section 3 the proposed methodology for breast cancer detection is presented, section 4 presents results and discussion and finally section 5 concludes the paper.

Related Work: A novel methodology for improvingwavelet transform for enhancing and detecting micro calcifications in mammograms was developed [7] depending upon the supervised learning approach. In this approach, a cost function is used to indicate the difference between thepreferred output and the reconstructed image resulting from the weighted wavelet coefficients.

A Decision Support System is structured using multi objective genetic and neural network algorithm tocategorizethe tumor and recognize the levels of the cancer. The first order statistical features, spatial gray level dependent features, surrounding region dependent features, gray level run length feature and gray level difference feature havebeen considered in [8].

Corresponding Author: M. Kanchana, Research Scholar, Anna University, Chennai, India.

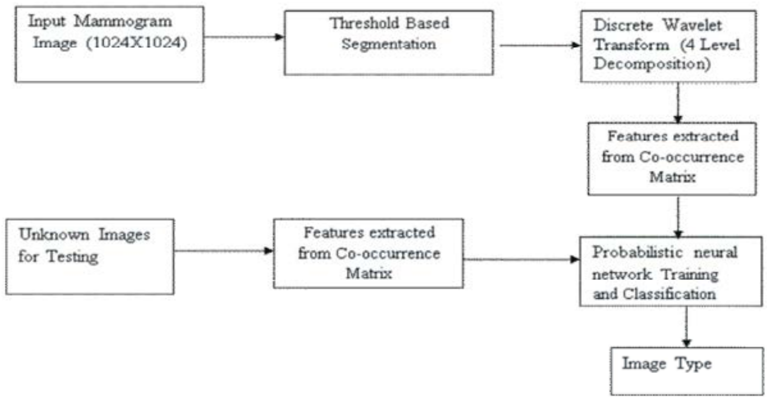

Fig. 1: Functional block diagram of the proposed scheme

Proposed methods for mass detection in the digital mammogram image of the breast cancer, in which 22 conventional features and 3 are unconventional features, are extracted from images [9] andminimum distance classifier and k-Nearest Neighbor (k-NN) classifiers are used for classification.

2-D discrete wavelet transformsbased on multi-resolution analysis was applied in [10]and the algorithm for the development of computer aided detection of microcalcification helps radiologists in the diagnosis of breast cancer.

Proposed Methodology: The objective of this paper is to diagnose breast cancer based on the mammogramimage, through a probabilistic neural network.The functional block diagram of the proposed methodology is shown in Fig 1 respectively.

Image Acquisition: The initial step is the acquisition of data in the form of digital mammogram images. The mammogram images are acquired from theMammography Image Analysis Society (MIAS) database [11]. The image acquired is in PGM (Portable GrayMap) format and it is the most widely utilized formatsas it is lossless type. Each acquired image has a resolution of about 1024 X 1024 pixels respectively.

Segmentation: After image acquisition, the next step is segmentation. Segmentation is the process of classifying the image into several regions[12]. The proposed segmentation of intensity images such as mammograms is based on thresholding technique (Ostu method). The thresholding in mammogram images involves the separation of background from the breast tissues. Based on the threshold value the pixels that are less than the threshold are considered as background and the remaining pixels are considered as breast.

Discrete Wavelet Transform: The presence of microcalcification cluster is one of the vital symptoms of breast cancer that have higher attenuation than the breast tissue and appears as rough bright spot in the mammograms. These spots appear as high frequency in the frequency domain and can be identified by using wavelet based approaches.

Basically Wavelet Transform (WT)analyzes the non-stationary signals.Wavelet Transformiscalled Multi-resolution Analysis (MRA), since it examines the signals at various frequencies providing different resolutions. The wavelet transform comprises of incorporating a signal into a broad range of translated and dilated versions of the antecedent wavelet(t).Wavelet provides a simple hierarchical framework for interpreting the image information.

The wavelet set is produced from the mother or basic wavelet which is defined as:

$$a,b(t) = \frac{1}{\sqrt{a}}\psi\left(\frac{t-b}{a}\right), a > 0 \tag{1}$$

The variable ‘a’ (inverse of frequency) reflects the width of a specific basis function such that its large value gives low frequencies and small values gives the high frequencies. The variable ‘b’ specifies its translation along x-axis in time. The term $\frac{1}{\sqrt{a}}$ is used for normalization.

Feature Extraction: Feature extraction techniques extract the most discriminating features that represent various classes of images. The extracted features provide the image properties and compare with the unknown sample image features for classification. Here, feature extraction is performed using Gray Level Co-occurrence Matrix (GLCM) approach. The Co-occurrence matrix is constructed based on the wavelet sub bands. Then the texture features, namely Energy, Contrast, Correlation and Entropy are extracted from the constructed co-occurrence matrix for diagnosis purpose. The features are given as follows.

Energy: It is a grayscale image texture measure of homogeneity changing, reflecting the distribution of image grayscale uniformity of weight and texture. The formula for calculating the Energy _ is given as follows:

$$E = \sum_x \sum_y p(x,y)^2 \tag{2}$$

Contrast: Contrast is the main diagonal near the moment of inertia, which measure the value of the matrix is distributed and images of local changes in number, reflecting the image clarity and texture of shadow depth. The contrast is computed as follows:

$$contrast(I) = \sum \sum \ (x-y)^2 \, p(x,y) \tag{3}$$

Entropy: It measures image texture randomness, when the space co-occurrence matrix for all values is equal, it achieved the minimum value. The Entropy S is given as

$$S = \sum_x \sum_y p(x,y) \log(x,y) \tag{4}$$

Correlation Coefficient: Measures the probability of combined occurrence of the specified pixel pairs.

$$correlation = \sum_{x,y=0}^{n-1} p_{x,y} \frac{(x-\mu)(x-\mu)}{\sigma^2} \tag{5}$$

Classification: Once the features are extracted they are used as input to the classifier in order to classify the images into malignant or benign. In the proposed approach probabilistic Neural Network (PNN) classifier is trained to classify the images by comparing the features of the unknown sample images with the trained features.Probabilistic Neural Networks provides a solution for pattern recognition problems.

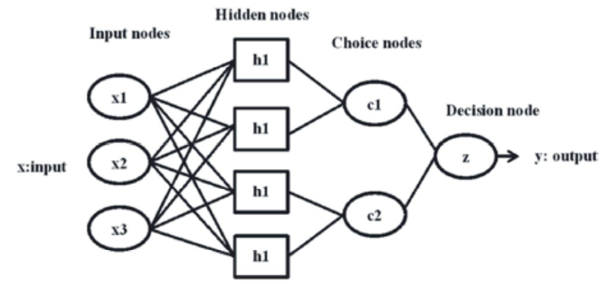

Fig. 2: PNN architecture

Figure 2 shows the architecture of the PNN classifier. The architecture of the Probabilistic Neural Networks (PNN) is same as that of the General Regression Neural Networks (GRNN) but a fundamental difference is thatnormal regression neural networks deal with regression in which the target variable is persistent but the probabilistic networks deal with classification in which the target variable is categorical.

The PNN used has multilayer structures comprising of a single RBF hidden layer of modified units that are fully connected to an output layer. The nodes of the hidden layer compute the Euclidean distance between the network input vector and the center and passes the output to the radial basis function. The output of the PNN classifier is defined as

$$c = \frac{\sum_{i=1^e}^{n} \frac{(a_i - 1)}{\propto^2}}{n} \tag{6}$$

Here, n is the total number of samples, a_i is the activation function of the node in the hidden layer and á is the smoothing factor.

RESULTS AND DISCUSSION

The acquired data set is divided into two data sets, namely the training set (25 images) and the testing set (17 images). The testing data set was used to assess the accuracy and the robustness of the trained network and the training data set was used to train the network for the classification of breast tumors. Table 1 shows the MIAS database used in this study and Table 2 shows the feature extraction results.

After feature extractions, the PNN is used to classify the mammogram images. The most important advantage of PNN is that training is easy and instantaneous. The performancemetricssuch as sensitivity,specificity and accuracy are used to evaluate the efficiency of the proposed classificationsystem.

Table 1: MIAS database images

Cases	#total images	#benign cases	#malignant cases
Normal	207	-	-
Micro calcification	25	12	13
Circumscribed masses	23	19	4
Speculated masses	19	11	8
Ill-defined masses	14	7	7
Architectural distortion	19	9	10
Asymmetry lesion	15	6	9
Total	322	64	51

Table 2: Features Extraction Results

Sl.No	Texture Features	Image Set 1 Value	Image Set 2 Value
1	Energy	0.3043	0.4198
2	Contrast	2.0344e+03	2.3203e+03
3	Correlation	0.7948	0.1131
4	Entropy	3.4002	2.8120

Table 3: Performance Metrics of the Classifier

S.No	Metrics	Value (%)
1	sensitivity	90.3 %
2	Specificity	100%
3	Accuracy(Overall)	92.3%

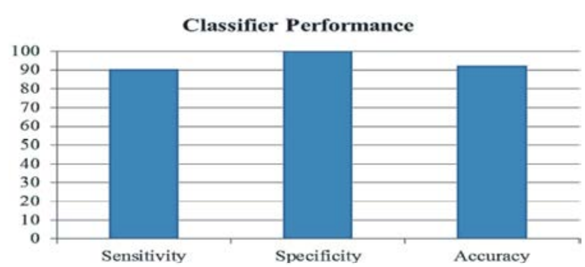

Fig. 3: Performance of the classifier

Table 3 shows the effectiveness of the proposed classifier and Fig 3 shows the graph indicating the performance of the classifier.

$$\text{Sensitivity} = \frac{\text{Number of true positive decisions}}{\text{Numbers of actual positive cases}} ?100 \quad (6)$$

$$\text{Specificity} = \frac{\text{Number of true negetive decisions}}{\text{Numbers of actual negetive cases}} ?100 \quad (7)$$

$$\text{Accuracy} = \frac{\text{Number of correct decisions}}{\text{Total numbers of cases}} ?100 \quad (8)$$

CONCLUSION

Discrete Wavelet Transform using fourlevel decomposition is used to decompose the mammogram for obtaining microcalcification clusters and the classification based on supervised learning using wavelet statistical features and target vectors. The implementation of the proposed method was carried out using MATLAB. The proposed approach is tested on MIAS database and the obtained classification accuracy is 92.30% respectively.

REFERENCES

1. Balakumaran, T., 2010. Ila. Vennila and C. Gowrishankar, “Detection Of microcalcification In Digital Mammograms Using One Dimensional Wavelet Transform”, ICTACT Journal On Image And Video Processing.
2. Sasikala, 2013. Vasanthakumar, “Breast Cancer-Classification And Analysis Using Different Scanned Images”, International Journal of Research in Engineering and Advanced Technology.
3. Giovanni Palma, Serge Muller and Isabelle Bloch, 2010. “Spiculated Lesions and Architectural Distortions Detection in Digital Breast Tomosynthesis Datasets,” IWDM 2010, LNCS 6136, pp: 712-719.
4. Kai-yang Li and Zheng Dong, 2006. “A Novel Method of Detecting Calcifications from Mammogram Images Based on Wavelet and Sobel Detector,” ICMA.
5. Ganesan, N., K. Venkatesh and M.A. Rama, 2010. “Application of Neural Networks in Diagnosing Cancer Disease Using Demographic Data”, International Journal of Computer Applications.
6. Zaheeruddin, Z., A. Jaffery and Laxman Singh, 2012. “Detection and Shape Feature Extraction of Breast Tumor in Mammograms”, Proceedings of the World Congress on Engineering.
7. Chang, R.F., K.C. Chang-Chien, E. Takada, C.S. Huang, Y.H. Chou, C.M. Kuo and J.H. Chen, 2010. “Rapid image stitching and computer-aided detection for multipass automated breast ultrasound,” Med. Phys., 37: 2063-2073.
8. Muthusamy Suganthi and Muthusamy Madheswaran, 2012. “An Improved Medical Decision Support System to Identify the Breast Cancer Using Mammogram”, J. Med Syst.
9. Salem Saleh Al-amri, N.V. Kalyankar and S.D. Khamitkar, 2010. “Image Segmentation by Using Thershod Techniques”, Journal of Computing, 2: 5.

10. Nizar Ben Hamad, Khaled Taouil and Med Salim Bouhlel, 2013. "Mammographic Microcalcifications Detection using Discrete Wavelet Transform", International Journal of Computer Applications, 64: 21.
11. http://www.mammoimage.org/databases/.
12. Ramesh Kumar, K.K. and A. Anbumani, 2014. "Medical Image Segmentation using Multifractual Analysis" International Journal of Inventions in Computer Science and Engineering (IJICSE), 1: 3.