Breast Cancer Detection Using BPN Classifier And Grey Level Co-Occurrence Matrix

Gayathri J [1], Arunpriya C [2]
[1]Dept of Computer Science
[2]Assistant Professor, Dept of Computer Science
[1, 2] PSGR Krishnammal College for Women, Coimbatore, Tamil Nadu, India

Abstract- *This paper describes a computer-aided detection and diagnosis system for breast cancer, the most common form of cancer among women, using mammography. The system relies on the Multiple-Instance Learning (MIL) paradigm, which has been proven useful for medical decision support in previous works. In the proposed framework, the initial step is Partitioning; breasts are first partitioned adaptively into regions. The Grey level cooccurrence Matrix (GLCM) Features are extracted from wavelet sub bands. Then, features derived from the appearance of textural features as well as detection of lesions (masses and micro calcifications) are extracted from each region and combined in order to classify it into examinations of mammography as "normal" or "abnormal". Whenever an abnormal examination record is detected, the regions that induced the automated diagnosis can be highlighted. There arise two strategies to define this anomaly detector. In a first scenario, manual segmentations of lesions are used to train an NN that assigns an anomaly index to each region; local anomaly indices are then combined into a global anomaly index.*

Keywords- Breast Cancer, Classifier, Neural Network and GLCM

I. INTRODUCTION

Recent studies have shown that breast cancer is the most common type of cancer among women, accounting for about one third of newly diagnosed cancers in the US. The mortality rate of breast cancer is also high, accounting for 17% of deaths relating to cancer in general. Accurate detection and assessment of breast cancer in its early stages is crucial when it comes to reducing the mortality rate. Mammography is the most useful tool for general population screening. However, the accurate detection and diagnosis of a breast lesion solely based on mammography findings is difficult and highly depends on the expertise of the radiologist, which leads to a high number of false positives and additional examinations. Breast cancer has emerged as the main reason behind most cancers' deaths among women. However it is extremely complicate to discover associated diagnoses tumours at a premature stage, cancer should be handled at the very early stage. The out of control development of cells in an organism called tumours can be cancerous. There are two kinds of tumors, benign and malignant. Benign or non-cancerous tumors are not spreading and are not life intimidating. In the other hand, malignant or cancerous tumor is expanding and life threatening [1]. Malignant breast cancer is defined when the growing cells are in the breast tissue. Breast cancer is the second overall cause of mortality among women and the first cause of dead among them between 40 and 55 ages. Regular breast cancer diagnosis followed by appropriate cancer treatment can reduce the unwilling risk. It is suggested to do tumor evaluation test every 4-6 weeks. Based on that reason, benign and malignant detection based on classification features become very important. Careful diagnosis in early detection has been proven to lessen the dead rate because of breast cancer. Depending on the expertise, mistakes can be made by medical professionals while identifying a disease. Manual analysis of cancer is found to be incompetent in several scenarios and extremely time-consuming process. As a result, there exists a choice for sensible schemes that identifies the cancerous cell, simultaneously deprived of any participation of people and with excessive accuracy [2]. Moreover, Digital mammography is the frequently used for breast most cancers detection. Labour-intensive analysis completed through using oncologist will no longer provide results with supreme accurateness and it is extremely time consuming. Here, provided the automated discovery of breast most cancers the usage of image processing schemes and artificial neural network (ANN). Mammogram image graphs have considered from MIAS (Mammography Image Analysis Society) [3, 4].Picture processing consists of diverse strategies to enable the virtual mammogram image ultimate for ANN community. Initially, the input image is processed with pre-processing, image enhancement, noise removal, image restoration, object recognition, segmentation and feature extraction. It must be noted that the statistical constraint is an important step in classification of mammography. The dominant extraction function is texture constraint, by means of which the irregularities can be identified with any concern. Texture is a technique for the purpose of extracting pattern. Statistical constraints comprise texture, entropy, mean, fashionable deviation, electricity, correlation [5]. In this paper, a new scheme for the computer-aided diagnosis (CAD) of micro calcification clusters (MCCs) detection in a Multi Instance Learning (MIL) framework is proposed. In the MultiInstance Learning framework, each training example is regarded as a bag of instances. A bag is positive if it contains at least one positive instance, otherwise negative. The method proposed in this paper to detect malignant tumors consists of a two-step process. The first step is intended to detect tumor candidates. The second step is an evaluation of their malignancy in order to reduce the number of false positives. This constraint could be provided as input to the particular classifier. There are specific classifiers employed for the purpose of investigation of different types of cancer. ANN is the main classifier employed in recent times. ANN is employed for the purpose of classification among cancerous and non cancerous images.

II. LITERATURE REVIEW

Jai-Andaloussi S, et al [6] offered detection of breast cancer in digital mammograms with the use of more than one concentric layer. Here, the application of more than one concentric layer (mcl) is effectively used for the purpose of mammogram segmentation. Here, this approach possessed the automated detection of cancer in digital mammograms. It effectively executes the image to array converter, at that time the image can be productively segmented. Following this, binary mask image are characteristically acquired through thresholding process from the gray scale image. The image granulation phase regulates that group of pixels which are sturdily linked based on spatial region and depth variety.

Chandra Prasetyo Utomo [7] Breast cancer is the second cause of dead among women. At early cancer detection followed by appropriate treatment can reduce the deadly risk. Medical professionals can also make mistakes while identifying a disease. The help of technology such as machine language and data mining can significantly improve the diagnosis accuracy. Artificial Neural Networks (ANN) has been widely used in breast cancer diagnosis. Although the standard Gradient - Based Back Propagation Artificial Neural Networks (BPANN) has some limitations, there are parameters to be set in the beginning, possibility to be trapped in local minima and long time for training process. In this research, we implemented ANN with extreme learning techniques for diagnosing breast cancer based on Breast Cancer Wisconsin Dataset. Results showed that Extreme Learning Machine Neural Networks (ELMANN) has better generalization classifier model than BP ANN. The development of this techniqueis promising as intelligent component in medical decision support systems.

C. R. Maurer Jr., R. Qi, and V. Raghavan[8] have proposed Computer Aided Detection and Diagnosis of Breast Cancer in Mammogram images. They described about an overview of recent advances in the development of CAD (Computer aided diagnosis or detection) systems and related techniques for breast cancer detection and diagnosis. This chapter focuses on mammogram image classification based on Rough Fuzzy Neural Networks. To the authors' knowledge, no researcher has applied Rough Fuzzy Neural Networks for mammogram image classification. Hence, the results of the proposed Rough Fuzzy Neural Networks method are compared with the Back Propagation Neural Networks, Fuzzy Neural Networks and Rough Neural Networks counterpart. The accuracy of the proposed method is validated through a tenfold validation method.

B. Monica Jenefer, v. Cyrilraj[9] Nowadays it is immediate need for best pre-screening tool to identify the abnormality of the mammogram images in the earlier stage itself. In this paper it is discussed about a tumor segmentation and classification algorithm from mammogram. The proposed approach concentrates on the result of two issues. One is the way to recognize tumors as suspicious regions may be very weak contrast to the background and the next is the way to concentrate properties which classify tumors. The proposed technique follows step by step procedures such as (a) Image Enhancement (b) Tumor Segmentation. (c) The extraction of properties from the segmented tumor region. (d) The utilization of SVM classifier. The improvement could be characterized as change of the image originality to a superior and more reasonable level. The mammogram enhancement can be obtained by removing the noise and improve the quality of the image using speckle noise removal and EM algorithm respectively. The most well-known division technique utilized is Modified Watershed Segmentation method. The features are extracted from the segmented tumor region and classify the regions utilizing the SVM classifier.

III. PROPOSED METHODOLOGY

The incidence of Breast Cancer has risen by about 1 percent annually over the past 50 years, with that rate of increase slowing over the past decade. Computer-aided detection and diagnosis algorithms have been introduced to reduce the false positive findings and hence unwanted biopsies can be eliminated by reducing false positive findings. In this paper, Back Propagation Neural Network (BPNN) has been proposed for the classification of cancer stage. The input image is enhanced by Wavelet decomposition. The classification is based on the comparison between the values of the GLCM (Grey Level Co-occurrence Matrix) features extracted from the input image and the database images.Mammogram uses X rays to create images of the breast. Earlier there were film mammography's exists in which the images were stored on films, but now a day digital mammograms are widely used because they are captured and stored directly on digital computers as well as all the corners and nooks are visible for easy detection. In breast ultrasound the images are created using sound waves. But this is not used nowadays as it is done with a handheld device, it will generate false positives and false negatives when the person who operates it is not well experienced or skilled and thus the quality of the image will vary [10]. The detailed detection of the images is done using texture extraction and feature extraction techniques. This leads to another wide variety of technique research area which have been used for segmentation, feature extraction, enhancement. Done mainly by clustering, wavelet techniques using GLCM matrix etc.., which are clearly described in the related works. So the utmost aim of this survey is to provide detection and classification techniques and different enhancement for early breast cancer detection.

A. Image Pre-Processing

Pre-processing refers to the process of operation on images which are available as raw data. Its main aim is to reduce the unwanted distortions and to enhance some features of the image which will be helpful for future processing and analysis. In pre- processing three level processing is done. The first processing is conversion of RGB image to Grey scale image because color increases the complexity of the image, so inherent complexity of the grey level image is smaller than the color image [11].The second processing involves resizing of the image to 256 rows and 256 columns in order to convert the image into standardized format. The third level processing is removing of Gaussian white noise from the image by using Gaussian smoothening filter. Adaptive Histogram Equalization (AHE) method has been employed for the enhancement of the image.

B. Wavelet Decomposition

Image processing includes a recent wavelet transform known as discrete wavelet transforms (DWT). An image is decomposed into different sub bands using this DWT, which is known as reconstruction of an image either by decomposition or by analysing noise image [12]. When a list of 2 n numbers is provided as input, the input values are paired up using the Haar wavelet transform, in which the difference is stored and sum is passed. The sum is paired by repeating the process recursively in order to prove the next scale, which gives a final sum and 2n -1 difference.

C. Feature Extraction Using GLCM

An image feature is a distinguishing characteristic of an image. Some features are natural, the visual appearance of an image, while others are artificial features which result from specific manipulations of an image. Natural features include the luminance of a region. The GLCM is also known as co-occurrence distribution. It is a second-order statistical method for textural analysis. The GLCM is a tabulation of how often different combinations of grey levels co-occur in an image or image section Energy: The homogeneousness of an image which is calculated from the normalized Co-occurrence matrix is measured as Energy. The disorder in the texture image can be detected [13].The GLCM Features are extracted from rippling sub bands. The, features derived from the detection of lesions (masses and micro calcifications) as well as textural features, are extracted from each region and combined in order to classify mammography examinations as "normal" or "abnormal".

D. Back Propagation Neural Network

An algorithm for supervised learning of artificial neural networks is given by the Backpropagation neural network also given as "backward propagation of errors" by using gradient descent. When an artificial neural network and an error function are given, the gradient of the error function with correspondence to the neural network's weights is calculated.

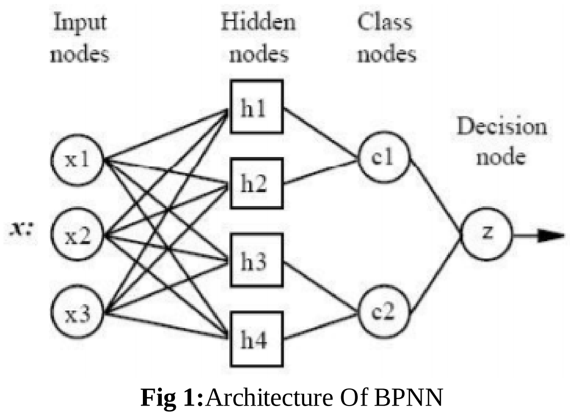

Fig 1: Architecture Of BPNN

The BPN with FF is trained with reference features set and desired output using 'newff' and 'train' command. Here, target 1 for dataset1, 2 for dataset2 and dataset3 are taken as desired output. After the training, updated weighting factor and biases with other network parameters are stored to simulate with input features [14, 15]. At the classification stage, test image features are utilized to simulate with trained network model using 'sim' command. Finally, it returns the classified value as 1 ,2 or 3 based on that the decision will be taken as Normal, Benign or Malignant.

IV. EXPERIMENTAL RESULT

The mammography image which was given as the input was subject to pre-processing where adaptive histogram equalization was used to obtain better clarity of the image. It was followed by edge detection using the combination of Sobel and Laplacian of Gaussian gradients. The edge detected image was divided into four sub-bands by the wavelet decomposition. The GLCM features were extracted and compared with that of the database images to classify them into “normal”, “benign” or “malignant” using the back propagation neural network classifier.

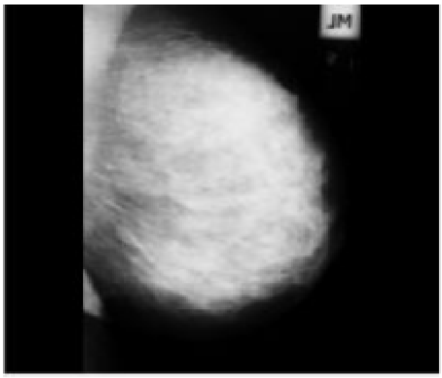

Fig 2:Input of the mammography image

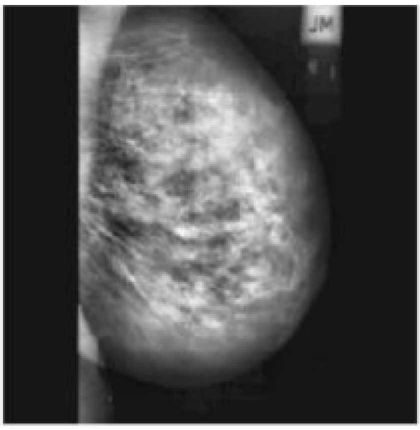

Fig 3: Pre-processing image

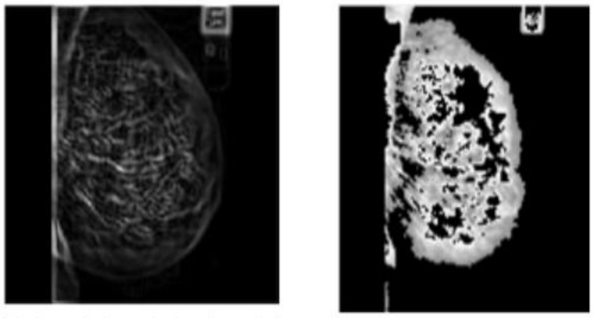

Fig 4: Normal and Abnormal Classification

The network generates the performance metrics, 1) Sensitivity: 91.6667% 2) Specificity: 100% 3) Accuracy: 93.3333%

V. CONCLUSION

Back propagation neural network was used here to classify test image into normal or abnormal based on supervised training and its training time is fast and has low complexity. Finally, the classifier performance was evaluated and BPNN is a supervised algorithm in which error variations between the actual output and obtained output is back propagated. The procedure is repeated during learning to decrease the error rate by updating the weights through the back propagation of error and it supports high speed classification. Finally, the classifier performance was evaluated and proved that the proposed system gives better prediction accuracy then prior methodologies.

REFERENCES

[1] P. Král and L. Lenc, "LBP features for breast cancer detection," published in IEEE International Conference on Image Processing (ICIP), Phoenix, AZ, 2016, pp. 2643-2647.

[2] K. L. Kashyap, M. K. Bajpai and P. Khanna, "Breast cancer detection in digital mammograms," published in IEEE International Conference on Imaging Systems and Techniques (IST), Macau, 2015, pp. 1-6.

[3] R. J. Nandi,A. K. Nandi,R. M. Rangayyan,D. Scutt,“ Classifi- - cation of breast masses in mammograms using genetic programming and feature selection”, Medical and Biological Engineering and Computing August 2006, Volume 44, Issue 8, pp 683–694.

[4] S. Baeg, N. Kehtarnavaz,”Texture based classification of mass abnormalities in mammograms”, 13th IEEE Symposium on Computer-Based Medical Systems, 2000.

[5] GORGEL Pelin& SERTBA Ahmet&KilicNiyazi&UcanOsman , Mammographic mass classification using wavelet based support vector

machine, JOURNAL OF ELECTRICAL & ELECTRONICS ENGINEERING ,Vol. 9. No. 1,2009.

[6] Jai-Andaloussi S, et al, “Mass segmentation in mammograms by using BidimensionalEmperical Mode Decomposition BEMD”, Engineering in medicine and biology society (EMBC), (2013).

[7] Chandra PrasetyoUtomo, “Breast Cancer Diagnosis using Artificial Neural Networks with Extreme Learning Techniques” IJARAI, Vol. 3, No. 7, 2014

[8] C. R. Maurer Jr., R. Qi, and V. Raghavan, “A linear time algorithm for computing exact Euclidean distance transforms of binary images in arbitrary dimensions,” IEEE Trans Pattern Anal Mach Intell, vol. 25, no. 2, pp. 265–70, Feb 2003.

[9] B.MonicaJenefer, v.Cyrilraj “An Efficient Image Processing Methods For Mammogram Breast Cancer Detection”, Journal of Theoretical and Applied Information Technology , November 2014. Vol. 69 No.1

[10] Saejoon Kim, Sejong Yoon, and Donghyuk Shin "computer aided design of Cross-Institutional mammograms utilizing SVMs with Feature Elimination", Proc. IEEE Frontiers in the Convergence of Bioscience and Information Technologies, pp. 396-402, Oct. 2007.

[11] R. Venkatesan, P. Chandakkar, B. Li, and H. K. Li, “Classification of diabetic retinopathy images using multiclass multiple-instance learning based on color correlogram features,” in Proc IEEE EMBS, vol. 2012, 2012, pp. 1462–5

[12] Khan AK &Noufal P, "Wavelet based automatic lesion detection using improved active contour method", IJERT, Vol.3, No.6, (2014).

[13] Mohamed Fathima M, Manimegalai D, Thaiyalnayaki S. “Automatic detection of tumor subtype in mammogram based on GLCM and DWT features using SVM.” Information Communication and Embedded Systems (ICICES) DOI: 10.1109/ICICES.2013.6508213. Page(s):809-813

[14] Nithya R &Santhi B, “Comparative study on feature extraction method for breast cancer classification”, JATIT & LLS,Vol.33, No.2, (2011).

[15] Hinton, G.E.; Srivastava, N.; Krizhevsky, A.; Sutskever, I.; Salakhutdinov, R.R. Improving neural networks by preventing co-adaptation of feature detectors. arXiv 2012, arXiv:1207.0580.