Breast Cancer Diagnosis & Analysis Using Artificial Neural Network System

Shivani Jain[1], Neetu Sikarwar[2]
[1]*Student, Institute of Engineering Jiwaji University, Gwalior, M.P.*
[2]*Professor, Institute of Engineering Jiwaji University, Gwalior, M.P.*

Abstract- **Breast cancer is one of the leading cancers among women in developed countries including India. Early diagnosis of the cancer allows treatment which could lead to high survival rate or avoids further clinical evaluation or breast biopsy reducing the unnecessary expenditure. This paper aims to build Artificial Neural Network (ANN) model for detection of breast cancer based on Image Registration techniques. Gray Level Co-occurrence Matrix (GLCM) features are extracted and are used to train the ANN. The performance is analysed on the basis of Mean Square Error (MSE) for different number of neurons of ANN.**

Index Terms- Artificial Neural Network, Breast Cancer Detection, Mammogram.

I INTRODUCTION

Breast cancer is the second leading cause of cancer deaths in women [1]. Several imaging techniques are available for detecting breast cancer such as ultrasound imaging, MRI imaging and digital Mammography. Mammography is the most commonly used imaging technique today throughout the world. When a patient undergoes a mammography a beam of X-rays transverses the breast and creates a projected image on a film [2]. Breast tissue composition varies with the age and hormone levels in a woman. Reading mammogram is a challenge for Radiologists. Mammography is susceptible to a high rate of false positives as well as false negatives, causing a high proportion of women without cancer to undergo further clinical evaluation or breast biopsy, or miss the best time interval for the treatment of the breast cancer. CAD systems can improve the detection rate of cancer in its early stages.

II. RELATED WORK

Different methods have been used to detect and classify mammogram images such as wavelets [3], statistical methods and most of them used feature extracted using image processing techniques. Artificial neural networks are frequently used as classifier due to simplicity of the implementation. Image Registration i.e. aligning two mammogram images acquired by same/different sensors, at different times or from different viewpoints, is very helpful for accurate and early detection of breast cancer [4]. Artificial Neural Networks can be used to enhance effectiveness of Image Registration Techniques for cancer detection in Digital mammography.

Several works have been done using feature selection extracted using image processing to develop ANN based breast cancer detection system [5]. Ritthipravat et al. have proposed Artificial Neural Networks in Cancer Recurrence Prediction [6]. J. Dheeba proposed Modified Genetic Algorithm (MGA) tuned Artificial Neural Network for detection of tumors in mammograms [7].

Salim et al. have developed an approach in breast cancer diagnosis by using Hybrid Magnetoacoustic method (HMM) and Artificial Neural Network [8]. Ahmad et al. have discussed the Probabilistic Neural Network for breast cancer detection [9]. Guzman-Cabrera et al. proposed an approach to effectively analyze digital mammograms based on texture segmentation for the detection of early stage tumors [10]. Dheeba et al. investigated a new classification approach by using Particle Swarm Optimized Wavelet Neural Network (PSOWNN)[11]. Senapati et al. proposed a new hybrid for breast cancer detection by extending the application of a variation of particle swarm optimization called K-particle swarm optimization (KPSO) [12].

III. MATERIALS AND METHODS

The process of building the breast cancer detection system using ANN may be divided into two phases: Learning/Training Phase and Recognition/Testing Phase. In Learning/Training Phase the ANN is trained for recognition of breast cancer from known benign and malignant Mammogram Images. Features, extracted from the known mammogram images, are stored in the knowledge base and given as input to the ANN based Diagnosis System. In Testing/Recognition Phase the extracted features of known and unknown Mammogram images are compared for classification of images into malignant and benign using ANN.

A. Materials

The 42 mammogram Images selected which belong to two categories: benign and malignant. The mammogram images consist of 30 benign and 12 malignant images.

B. Feature Extraction

Feature extraction methodologies extract the most prominent features that are representation of various classes of images. The extracted features give the property of the image and are compared with the features of unknown sample Image for classification.

Intensity Histogram

Intensity Histogram analysis like mean, variance, entropy, energy, skewness and kurtosis has been extensively researched in the initial stages of development of for feature selection.

Gray Level Co-occurrence Matrix (GLCM)

The GLCM (Gray Level Co-occurrence Matrix) is statistical method considers the spatial relationship between pixels of different gray levels [28]. GLCM is a matrix S that contains the relative frequencies with two pixels: one with gray level values i and other with gray level value j which are separated by a distance d at certain angle θ in the image. Each element (i, j) in the GLCM is the sum of the number of times that the pixel with value i occurred in the specified relationship to a pixel with value j in the raw image. Co-occurrence matrices may be calculated for four directions i.e. horizontal, right diagonal, vertical and left diagonal (00, 450, 900 and 1350). Once the GLCM is calculated several second-order texture statistics or features may be computed. The features like Inertia Autocorrelation, Contrast, Correlation, Dissimilarity Energy, Entropy, Homogeneity, Maximum Probability, Sum of Squares, Sum Average, Sum Variance, Sum Entropy, Difference Variance, and Inverse Difference Normalized may be computed from each co-occurrence matrices which are computed in each of four angles [13, 14]. GLCM features are used in this work to train the ANN.

C. Classification

Once the GLCM features related to known Mammogram Images are extracted and selected, the features are used as inputs to train an ANN classifier to classify the images into benign or malignant. In testing/recognition the trained ANN is used to classify new images by comparing the extracted features with the features of the unknown sample Image of Mammogram.

IV. RESULT AND DISCUSSION

In this paper feed-forward back propagation Artificial Neural Network had been trained for the classification of Mammograms images into benign and malignant. Levenberg–Marquardt back-propagation algorithm was used to train the network and performance is measured on the basis of Mean Square Error (MSE).

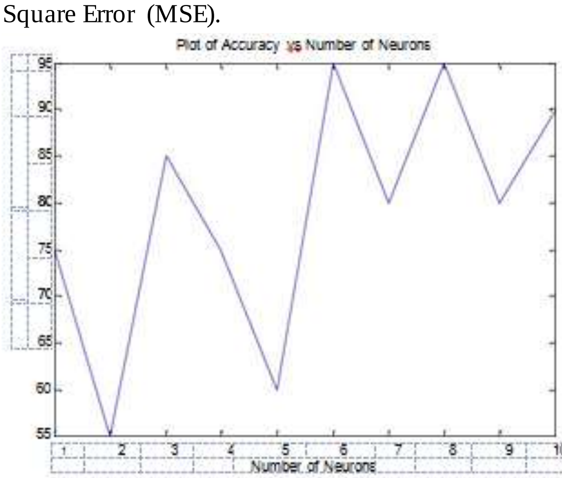

Fig. 1 Accuracy vs Number of Neurons

Mammogram images are divided into training set, testing set and validation set randomly. GLCM features are extracted from the mammogram images. The extracted features data will be used to train the

ANN based Diagnosis system. In training phase the features extracted from training set of images are saved in a dedicated database and used to train the Artificial Neural Network. While in Recognition phase, the extracted features from the unknown or testing set of mammogram images are passed through the trained ANN for classification.

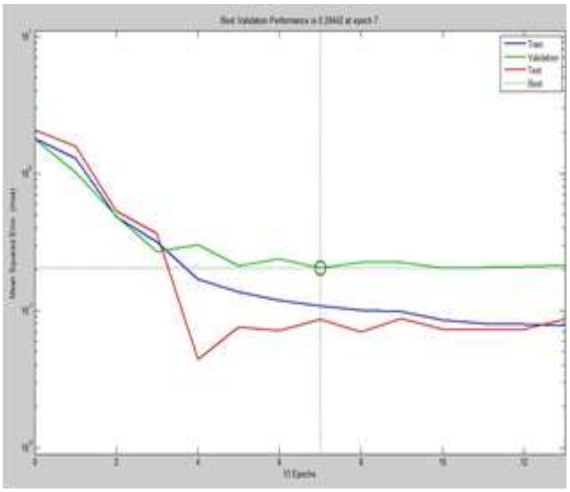

Fig. 2 Performance of the Training Process

The ANN model for breast cancer detection was tested to determine the predictive accuracy of detection for different number of neurons. The graph for the accuracy of ANN model for detection for different number of neurons is shown in Fig1.The graph shows that the detection system gives optimum performance for 6 and 8 neurons. The optimum number of neurons is chosen to be 6 as the number of resources requires for implementation are less as compared to that of 8. Because there are no well established theoretical methods for designing an ideal ANN, the best designs are typically determined through trial and error.

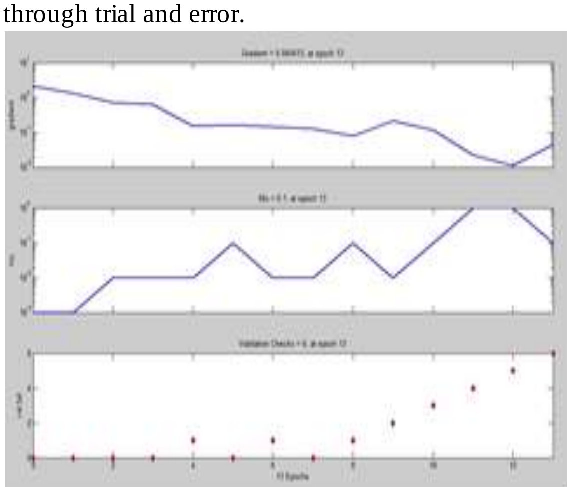

Fig. 3 Training State of the System

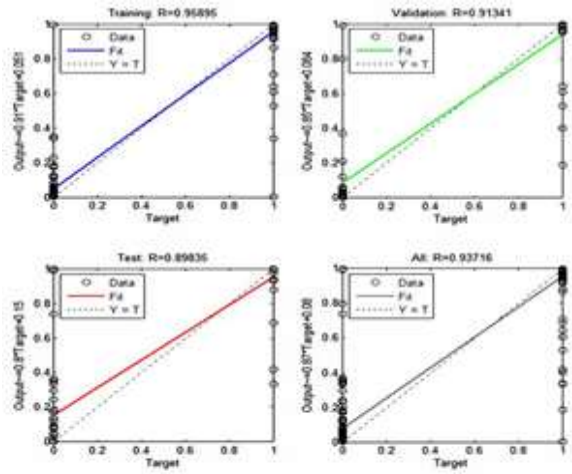

Fig. 4 Regression Plots of the System

The Performance graph of training process in Fig 2 shows Mean Square Error (MSE) for training, testing and validation process for different number of epochs. The training state of the system showing the gradient, mutation and validation check graphs for ANN based diagnosis system is shown in Fig 3.The regression plots for Training, testing and validation along with overall system are shown in Fig 4. The plot depicts that actual output is in well agreement with the desired output.

V. CONCLUSION

A Breast cancer diagnosis system based on Image registration techniques and ANN was successfully designed. Texture features are used for training of feed forward backpropagation network. Cooccurence matrices for different directions are calculated and GLCM features are extracted from the mammogram image. The GLCM features are used to train the ANN based detection system. Total optimum numbers of neurons are computed on hit and trial method and performance of the network is evaluated on the basis of Mean Square Error (MSE).The system has been tested on the mammogram images. The system can be designed to classify other types of cancers after few modifications. The accuracy of the detection can be improved by including more sets of mammogram images.

REFERENCES

[1] [Online] Available: http://www.cancer.org. American Cancer Society, 2017.

[2] Fred S. Azar, "Imaging Techniques for Detecting Breast Cancer: Survey and Perspectives", Technical Reports (CIS), University of Pennsylvania, 2009, pp.1-6.

[3] Wang T., Karayiannis N., "Detection of microcalcifications in digital mammograms using wavelets", IEEE Transactions on Medical Imaging, 2014, vol. 17.4, pp. 498–509.

[4] Manjusha P. Deshmukh & Udhav Bhosle, "A survey of image registration", International Journal of Image Processing (IJIP), 2015, vol.5.3, 1-6.

[5] Keivanfard, F., Teshnehlab, M., Shoorehdeli, M. A., Ke Nie, Min-Ying Su. "Feature selection and classification of breast cancer on dynamic Magnetic Resonance Imaging by using artificial neural networks", 17th Iranian Conference of Biomedical Engineering, 3-4 Nov. 2016. 1-4.

[6] Ritthipravat P., "Artificial Neural Networks in Cancer Recurrence Prediction", International Conference on Computer Engineering and Technology, 22-24 Jan 2009, vol. 2, 103-107.

[7] Dheeba J., Tamil Selv S., "A CAD System for Breast Cancer Diagnosis Using Modified Genetic Algorithm Optimized Artificial Neural Network", SEMCCO (Springer). Part I. LNCS 7076, 2013, 349–357.

[8] Salim MI, Ahmad AH, Arffin I, "Developmemnt of Breast cancer detection tool using Hybrid Magnetoacoustic method and Artificial Neural Network", International Journal of Biomedical Engineering, 2012, vol.6, 61-68.

[9] Ahmad Taher Azhar, Ahmed Shaimaa, El-Said, "Probabilistic Neural Networks for Breast cancer Classification", Neural Computing and Applications, Springer, 2013, vol. 23, 1737-1751.

[10] Guzman-Cabrera R., Guzman J. R., "Digital Image Processing Technique for breast cancer detection", International Journal of Thermo physics, Springer, 2013, vol.34, 1519.1531.

[11] Dheeba J., Albert Singh N., Tamil Selvi S., "Computer-aided detection of breast cancer on mammograms: A swarm intelligence optimized wavelet neural network approach", Journal of Biomedical Informatics, Elsevier, 2014, vol.49, 45–52.

[12] Senapati M. R., Panda G., Dash Hybrid P. K., "Hybrid approach using KPSO and RLS for RBFNN design for breast cancer detection", Neural Computing & Applications, Springer, 2018, vol.24,745–753.

[13] Gonzalez R. C., Woods R. E., Richard E, Steven L. E.s, “Digital Image Processing”, Second Edition, Pearson Educat ion, New Delhi,India.

[14] Devendram V., Thiagarajan H., " Texture based scene categorization using artificial neural networks and support ve ctor machines: a comparative study", ICGST-GVIP, 2018, vol.8,45-52.