Optimization of Artificial Neural Network Breast Cancer Detection System based on Image Registration Techniques

Satish Saini
Research Scholar, Department of Electronics
Banasthali University
Jaipur, Rajasthan, India

Ritu Vijay
Associate Professor, Department of Electronics
Banasthali University
Jaipur, Rajasthan, India

ABSTRACT
The paper presents a Feed-forward back-propagation Artificial Neural Network (ANN) model for detection of breast cancer using Image Registration Techniques. Gray Level Co-occurrence Matrix (GLCM) features extracted from the known mammogram images are used to train the ANN based detection system. The ANN based detection system will be investigated for different number of neurons and layers on the basis of Mean Square Error (MSE) and optimum number of neurons and layers will be chosen.

General Terms
Artificial Neural Network, Digital Image processing, Breast Cancer Detection

Keywords
Artificial Neural Network, Image Registration Techniques, Mammogram.

1. INTRODUCTION
Breast cancer represents the second leading cause of deaths in women today and it is the most common type of cancers in women [1]. Mammography is currently the most effective method for detection of breast cancer before it becomes clinically palpable [2]. One of the difficulties with mammography is that mammograms generally have low contrast. This makes it difficult for the radiologist to interpret the results. Double reading of mammograms can enhance the probability of proper diagnosis but it costs high. Using CAD system can improve the detection rate of breast cancer in early stages. 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. 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 [5-10].

2. MATERIALS AND METHODS
MATLAB has been used as the simulation tool. The objective of the work is to develop and optimize an ANN based breast cancer detection system. Features extracted from the known mammogram images will be used to train the ANN breast cancer detection system. 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

2.1 Materials
The 42 mammogram Images were selected from Post Graduate Institute, Chandigarh for ANN based detection system. The mammogram images consist of 30 benign and 12 malignant images.

2.2 Preprocessing
As mammograms are difficult to interpret preprocessing is required to improve the quality of the image and make the feature extraction phase easy and reliable. Preprocessing basically represents the data cleaning phase or noise-reducing step which includes removal of unwanted or irrelevant areas and to make prominent the area of interest by increasing the contrast and is performed before feature extraction. The mammogram images may undergo certain Image Pre-processing operations on the image such as defining ROI using Image Cropping process, Image Enhancement for contrast adjustment, Segmentation before feature extraction process.

2.3 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. Gray Level Co-occurrence Matrix (GLCM) features are used in this paper. The GLCM (Gray Level Co-occurrence Matrix) is statistical method considers the spatial relationship between pixels of different gray levels. 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 (0^0, 45^0, 90^0 and 135^0). Once the GLCM is calculated several second-order texture statistics or 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 [11, 12].

Table 1. GLCM features

Feature No.	Feature Name
1.	Autocorrelation
2.	Correlation
3.	Maximum probability
4.	Dissimilarity Energy
5.	Entropy
6.	Inverse Difference Normalized

2.4 Classification

The GLCM features, extracted from the known Mammogram Images, are used as inputs to train an ANN based breast cancer detection system in training phase. In testing/recognition the trained ANN compares the extracted features with the features of the unknown sample Image of Mammogram and classifies the new mammogram images into benign and malignant.

3. RESULT AND DISCUSSION

In this paper a model for breast cancer detection based on Artificial Neural Network has been designed using Image registration Techniques. Feed-forward back propagation architecture was used and Levenberg–Marquardt back-propagation algorithm was used to train the network for classification of mammogram images into benign and malignant. The performance is measured on the basis of Mean Square Error (MSE). Mammogram images are divided into training set, testing set and validation set randomly. GLCM features are extracted from the mammogram images will be used to train the ANN based Breast cancer detection system. In training phase the features extracted from training set of images are used to train the Artificial Neural Network. While in testing phase, the extracted features from the unknown or testing set of mammogram images are compared in the trained ANN for classification of the mammogram images.

Because there are no well established theoretical methods for designing an ideal ANN, the best designs are typically determined through trial and error. The ANN model for breast cancer detection was tested to determine the predictive accuracy of detection for different number of neurons [13] and layers and optimum number of neurons and layers is chosen. Percentage Accuracy for different number of neurons is shown in Table 2 and corresponding percentage accuracy for different number of layers is shown in Table 3. Figure 1 shows graph between the percentage accuracy of detection of ANN model achieved for different number of neurons and layers respectively. It can be seen that the system gives optimum performance for 7 neurons and it is also evident that 8 layers are required to obtain the best result. The achieved accuracy of the system is 92.8% and 87.5 % for neurons and layers respectively.

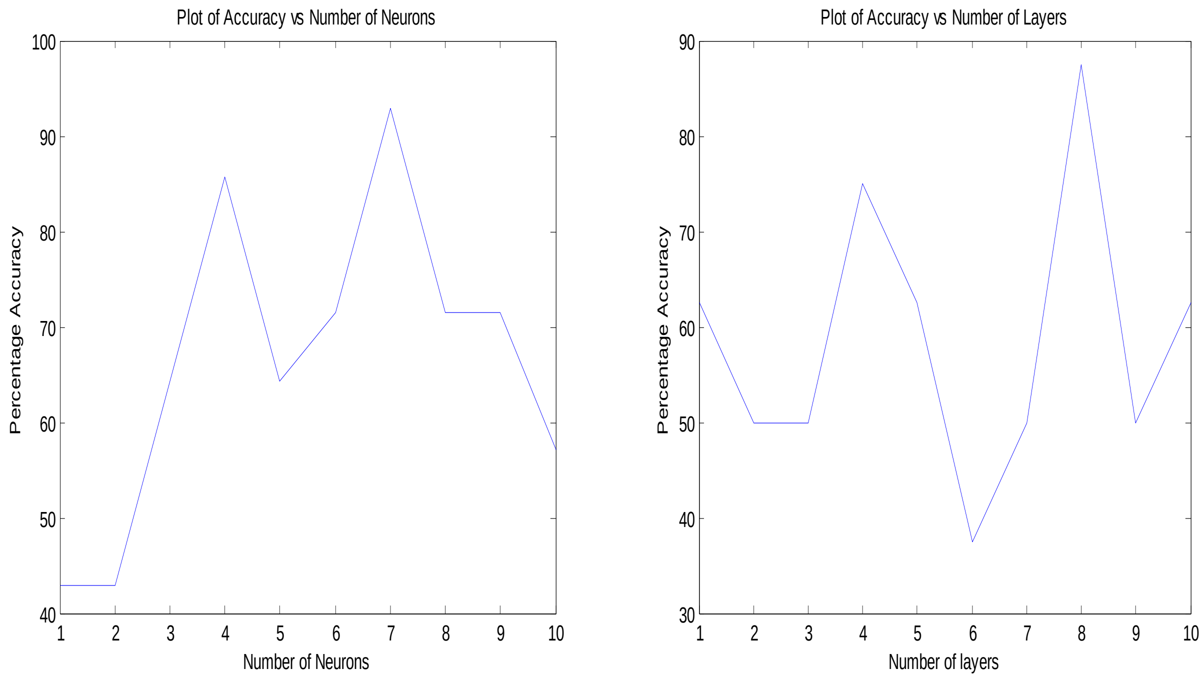
Table 2. Accuracy for different No. of Neurons

S. No.	No. of Neurons	Percentage Accuracy
1.	1	42.8
2.	2	42.8
3.	3	64.2
4.	4	85.7
5.	5	64.2
6.	6	71.4
7.	7	92.8
8.	8	71.4
9.	9	71.4
10.	10	57.1

Table 3. Accuracy for different No. of Layers

S. No.	No. of Layers	Percentage Accuracy
1.	1	62.5
2.	2	50
3.	3	50
4.	4	75
5.	5	62.5
6.	6	37.5
7.	7	50
8.	8	87.5
9.	9	50
10.	10	62.5

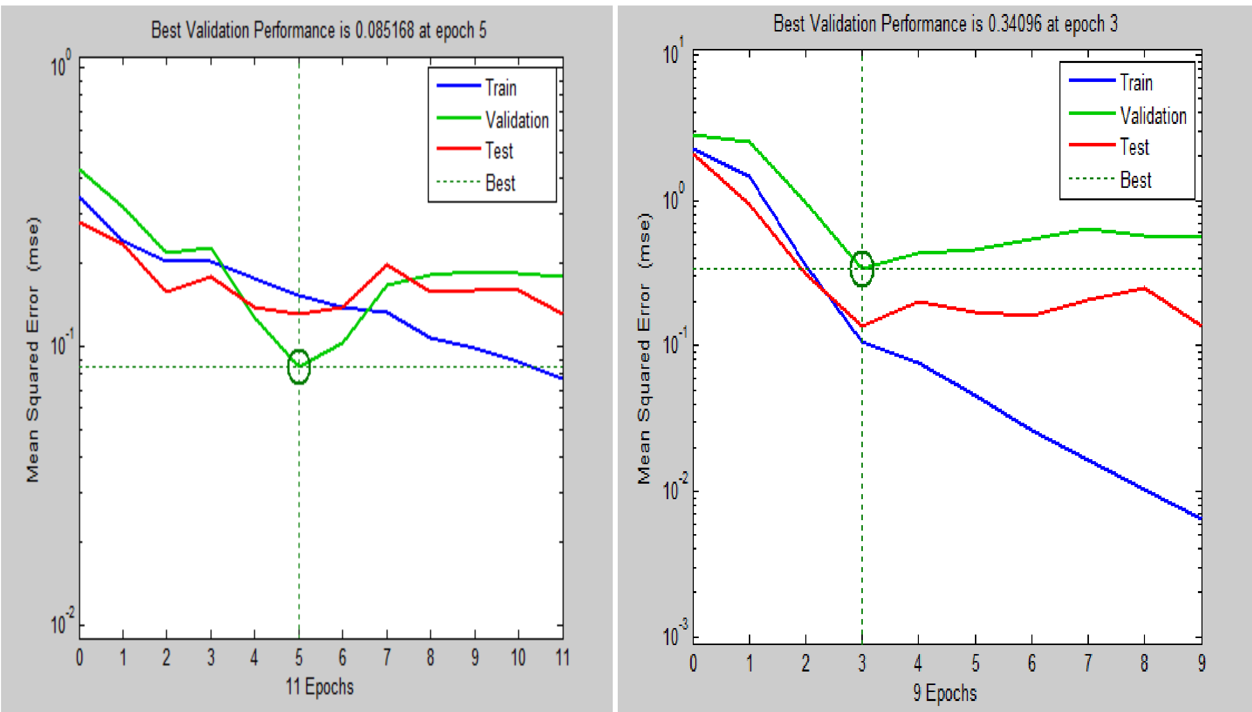
The Performance graph of training process for neurons and layers is shown in Figure 2. The graph shows MSE for training, testing and validation process for optimum number of neurons and layers. The best validation performance is 0.085168 at 5 epochs for neurons and 0.34096 at 3 epochs for layers. Mean Square Error (MSE) for training, testing and validation is shown in corresponding curves. The training state of the system for optimum number of neurons and layers is shown in Fig 3. The training state shows that the gradient is 0.025982, mutation is 0.1 and best validation check is 6 at 11 epochs for optimum number of neurons. For optimum number of layers the gradient is 0.013976, mutation is 0.01 and best validation check is 6 at 9 epochs.



a) **b)**

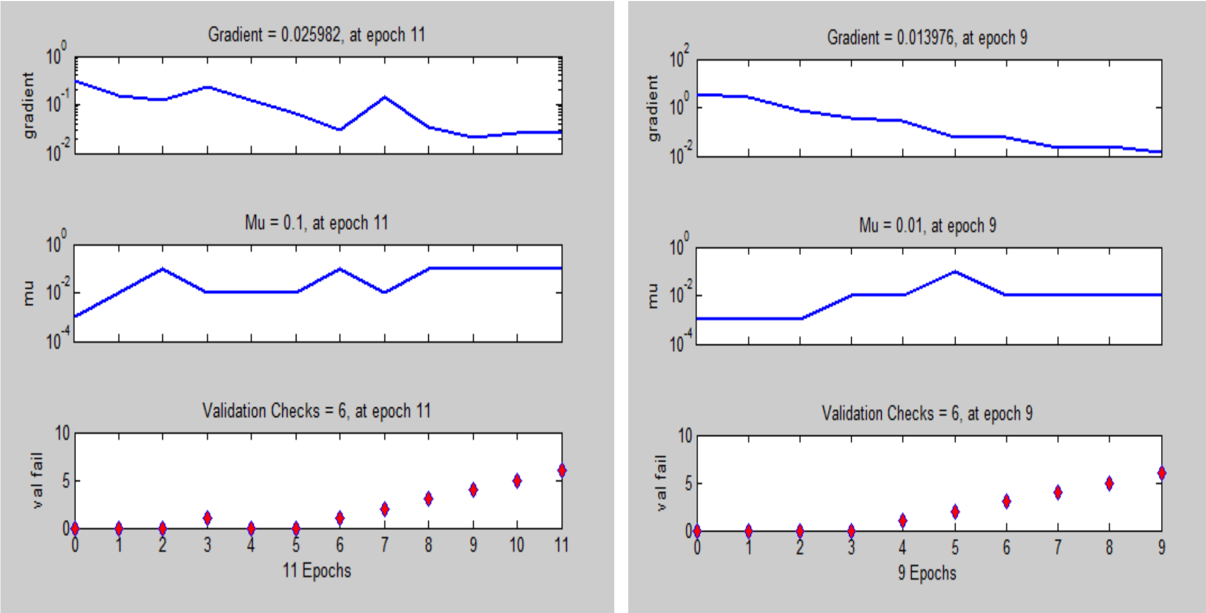
Fig 1: Accuracy for different number of a) neurons and b) layers

.



a) **b)**

Fig 2: Performance of the Training Process a) Neurons and b) layers



a) **b)**

Fig 3: Training State of the System for a) Neurons b) Layers

4. CONCLUSION

In this paper Image registration techniques were used for designing ANN based breast cancer detection using feed forward backpropagation network. Cooccurence matrices are calculated and GLCM features extracted from the mammogram images were used to train the ANN based detection system. Then the system was optimized by computing optimum number of neurons and layers on hit and trial method. The accuracy of the system was enhanced by optimization of number of neurons and layers. The performance of the system was evaluated on the basis of Mean Square Error (MSE). Thus it can be concluded from this work that ANN with 7 neurons and 8 layers is ideally suitable for predicting breast cancer and achieved accuracy is 92.8% and 87.5 % respectively. The system has been tested on mammogram images and can be designed to classify other types of cancers after few modifications. The accuracy of the system can be improved by including more sets of mammogram images and texture features. The system can also be designed and compared for different architectures of Artificial Neural networks.

5. REFERENCES

[1] http://www.cancer.org. American Cancer Society, 2012.

[2] Fred S. Azar, "Imaging Techniques for Detecting Breast Cancer: Survey and Perspectives", Technical Reports (CIS), University of Pennsylvania, 2000, pp.1-6.

[3] Wang T., Karayiannis N., "Detection of microcalcifications in digital mammograms using wavelets", IEEE Transactions on Medical Imaging, 1998, vol. 17.4, pp. 498–509.

[4] Manjusha P. Deshmukh & Udhav Bhosle, "A survey of image registration", International Journal of Image Processing (IJIP), 2011, 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. 2010. 1-4.

[6] 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, 2011, 349–357.

[7] 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.

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

[9] 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.

[10] 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, 2014, vol.24,745–753.

[11] Gonzalez R. C., Woods R. E., Richard E, Steven L. E.s, " Digital Image Processing" ,Second Edition, Pearson Education, New Delhi,India.

[12] Devendram V., Thiagarajan H., " Texture based scene categorization using Artificial Neural Networks and Support Vector Machines: a comparative study", ICGST-GVIP, 2008, vol.8, 45-52.

[13] Saini S., Vijay R., "Performance Analysis of Artificial Neural Network Based Breast Cancer Detection System", International Journal of Soft Computing and Engineering, 2014, vol. 4,Issue 470-72