International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 11 (2016) pp 7550-7554

Performance analysis of Neural Network based classification technique for Mammogram Images

A. P. Manjunatha[1], H. S. Sheshadri[2], M. N. GiriPrasad[3]

[1]*Associatate Professor, Department of Medical Electronics, Dr. AIT, Bengaluru, Karnataka, India.*

[2]*Professor, Department of Electronics & Communication, PES College of Engineering, Mandya, Karnataka, India.*

[3]*Professor, Department of Electronics & Communication, JNTU College of Engineering, Anantpur, Andhra Pradesh, India.*

Abstract

This paper presents experimental work on mammogram image analysis. Texture analysis is carried out using segmentation technique. Here, statistical method have been used to extract features from the segmented tumour area. The obtained features are classified using different classifiers such as Radial basis function, Main Feed forward and Main Fitnet method. The method was tested on 100 clinical data. The RBF classifier achieved anaccuracy of 91% Main fit net accuracy of 99. 20%Main Feed forward accuracy about 97%.

Keywords: RBF, Main Feed forward, Fitnet statistics

INTRODCUTION

For the early detection of breast cancer Mammography technique is most widely preferred technique by majority of the radiologists. X-Ray Mammography is the most reliable diagnostic procedure for detecting the masses and calcification in the tissue of the breasts. However during the screening procedure, radiologists often miss the detection of a significant proportion of abnormalities in addition for having high rates of false positives. [1] Many computer aided detection techniques are being implemented for the segmentation of these mammogram images and obtaining the desired region of interest, which are having high probability of occurrences masses. [2]

In the digital mammogram images, various segmentation techniques are implemented for segmenting the lesions. Radial gradient index based algorithm and a probabilistic algorithm are two such techniques. These segmentation techniques begin with a point, which is known as seed point. This point will be present in the suspected lesion. [3] The segmentation techniques are generally

Consisting of gray-level thresholding, region growing or active contours methods. [4]

In this paper we have made an attempt for segmentation of these mammogram images using texture analysis [5].

This procedure is using statistical approach for finding the suspected region of interest, having high probability of the occurrence of masses or calcification. From this region of interest 6 features are obtained. Further classification of these features is performed using Artificial neural Networks (ANN), RBF, Main fitnet,, Main Feed forward, diagnosis (Malignant or Benign). Comparison of the results of accuracy obtained from these classifier algorithms is also performed. The dataset of real time mammogram images are obtained from Padmashree Diagnostic Center, Bangalore, Karnataka, India.

Texture

To quantify the texture contents of a region we will use statistical approaches for computing texture based on statistical measures.

Statistical Approaches

The texture analysis is based on the statistical properties of the intensity of histogram. [6] It is based upon Statistical Moments [paper on statistical moments]. The expression for the nth order moments about the mean is given by equation (1)

$$\mu_{n}= \sum_{i=0}^{L-1}(z_i-m)^n\, p(z_i) \quad (1)$$

Where zi is a random variable indicating intensity, p(zi) is the histogram of the intensity levels in a region, L is the number of possible intensity levels and

$$m = \sum_{i=0}^{L-1} z_i p(z_i) \quad (2)$$

is the mean(average) intensity given by equation (2). Mean is the measure of average intensity.

Standard deviation is the measure of average contrast given by the equation (3).

$$\sigma = \sqrt{\mu_2(z)} = \sqrt{\sigma^2} \quad (3)$$

Smoothness measures the relative smoothness of the intensity in a region. R is 0 for a region of constant intensity and approaches 1 for regions with large excursions in the values of its intensity levels.

The variance used in this measure is normalized to the range [0, 1] by dividing it by (L-1)2 given in the equation (4)

$$R=1-1/\,(1+\sigma^2) \quad (4)$$

Variance is the measure of contrast.

Third moment measures the skewness if a histogram. This measure is 0 for symmetric histograms, positive by histograms skewed to the right (about the mean) and negative for histogram skewed to the left. Values of this measure are brought into a range of values comparable to the other five measures by dividing μ_3 by (L-1)2 also, which is the same divisor we use to normalize the variance given by equation (5).

$$\mu_3 = \sum_{i=0}^{L-1} (z_i - m)^3 \, p(z_i) \quad (5)$$

Uniformity measure is maximum when all the gray levels are equal (maximally uniform) and decrease from there, given in equation (6)

$$U = \sum_{i=0}^{L-1} p^2(z_i) \quad (6)$$

Entropy is the measure of the randomness given by equation (7)

$$e = -\sum_{i=0}^{L-1} p(z_i) \log_2 p(z_i) \quad (7)$$

Classification of the obtained features

ANN using RBF, Feed Forward, Fitnet
Introduction to ANN

Artificial neural network is a powerful tool for pattern classification problems. Artificial neural networks are capable of capturing domain knowledge from examples. It is a computational model based on the biological neural network system of the brain and usually assume that computation is distributed over several simple units called neurons, which are interconnected and operate in parallel (hence, neural networks are also called parallel distributed-processing systems or connectionist systems).

Basically Neural Network model *Perceptron *Feed forward *Radial basis function. *Support vector Machine. Inter neuron connection strength (weights) are used to store the acquired information, that is during training. The weights of the neural network are initiated randomly.

During learning process the weights are modified in order to model the particular learning task correctly on the training example. Neural Network had been one of the most widely used approaches for pattern recognition.

The most popular neural network is multilayer perceptron, which is a feed forward network, all signals flow in a single direction from the input to the output of the network. Feed forward network can perform static mapping between an input space and an output space: the output at a given instant is a function only of the input at that instant. Implicit'Knowledge' is built into a neural network by training it. Some neural networks can be trained by being presented with typical input patterns and the corresponding expected output patterns [7]. The error between the actual and expected output is used to modify the strengths, or weights of the connections between the neuron. This method of training is known as supervised training. In a multi-layer perceptron, the back-propagation algorithm for supervised training is often adopted to propagate the error from the output neurons for computing the weight modifications.

The radial basis function (RBF) network is a special type of neural networks with several distinctive features [8, 9, 10]. Since its first proposal, the RBF network has attracted a high degree of interest in research communities. A RBF network consists of three layers, namely the input layer, the hidden layer, and the output layer. The input layer broadcasts the coordinates of the input vector to each of the units in the hidden layer. Each unit in the hidden layer then produces an activation based on the associated radial basis function. Finally, each unit in the output layer computes a linear combination of the activations of the hidden units. How a RBF network reacts to a given input stimulus is completely determined by the activation functions associated with the hidden units and the weights associated with the links between the hidden layer. Neural classifier used for their good generalization and universal approximation property have been effectively used in their system for the classification.

Materials and Methods:-

Various steps involved in the proposed algorithm:

Step 1 Load the Features

Step 2 Initializing the necessary parameters.

Step 3 Generating random numbers.

Step 4 Reading the features (data) from random Position.

Step 5 Train the data of different classifiers like RBF, Feed forward, Fitnet functions.

Step 6 New network is generated.

Step 7 Aceesing the features from new network. (Testing the data).

Step 8 Computing the necessary parameter (error, efficiency).

Step 9 Creating the vector with the mean value Of the parameter.

Step 10 Generating the vectors of TP, TN, FP, FN TP= True Positive, TN= True Negative FP= False Positive FN= False Negative

Step 11 Generating the confusion matrix.

Step 12 Plotting the graphs of *Training ratio Vs Error rate. *Training ratio Vs Efficiency *Training ratio Vs TP *Training ratio Vs TN *Training ratio Vs FP *Training ratio Vs FN.

For finding Accuracy, Sensitivity, Specificity

The following formulas are used.

$$\frac{TP+TN}{TP+TN+FP+FN} = \text{Accuracy}$$

$$\frac{TP}{TP+FN} = \text{Sensitivity}$$

$$\frac{TN}{TN+FP} = \text{Specificity}$$

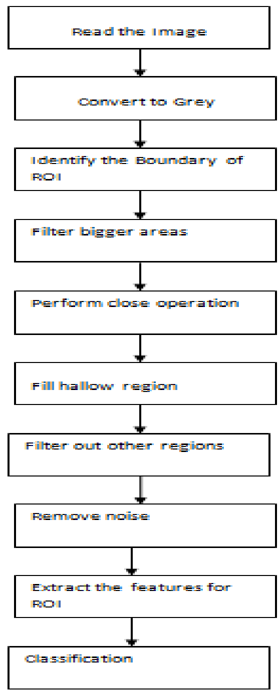


Figure 1: Flow Chart

International Journal of Applied Engineering Research ISSN 0973-4562 Volume 11, Number 11 (2016) pp 7550-7554

RESULTS AND DISCUSSION

Figure 2: Original Image

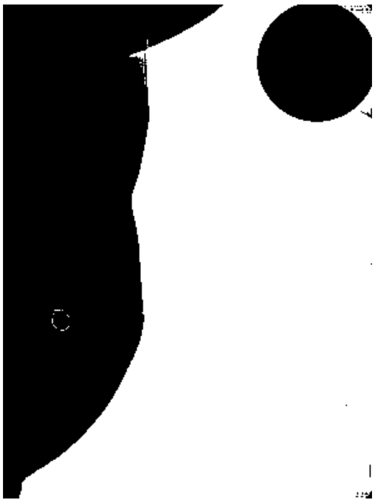


Figure 3: Highlight Marked ROI

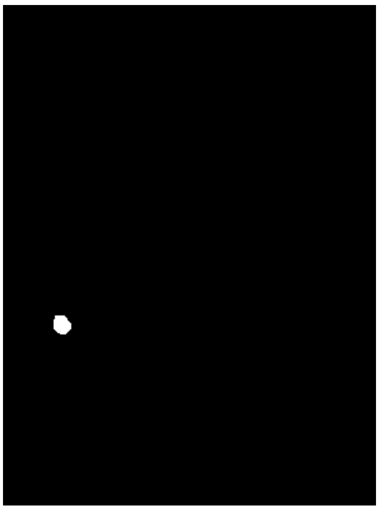


Figure 4: Segmented ROI

Features

Avg Intensity = 176. 287561,
Avg Contrast = 23. 782249,...
measure of smoothness = 0. 008623,
3rd Moment =-0. 752285,
Uniformity = 0. 020677,
Entropy = 5. 969177

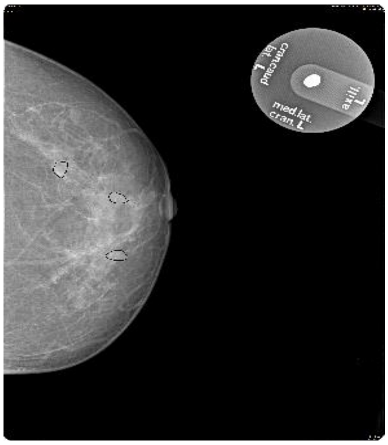


Figure 5: Original Image

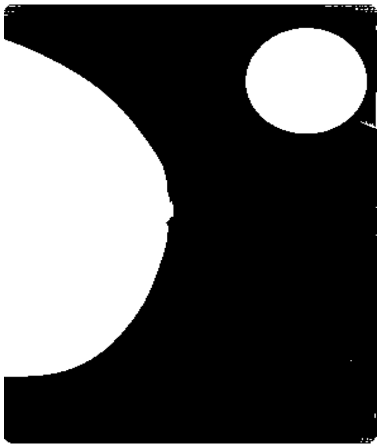


Figure 6: Binary Image

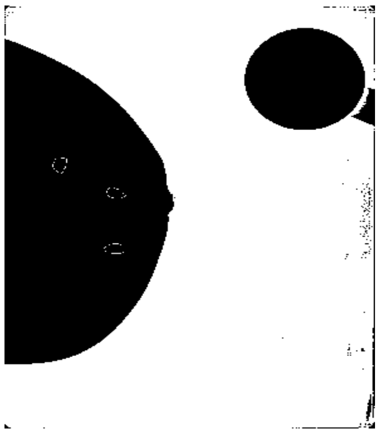


Figure 7: Highlight Marked ROI

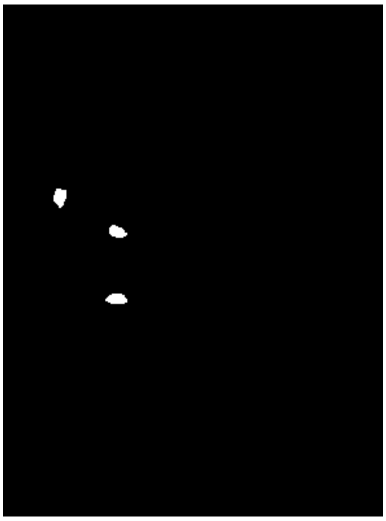


Figure 8: Segmented ROI

Features:
Region 1-----,
Avg Intensity = 179. 392412,
Avg Contrast = 25. 786781,...
measure of smoothness = 0. 010123,
3rd Moment =-0. 887256,
Uniformity = 0. 020142,
Entropy = 6. 018872

Region 2-----,
Avg Intensity = 162. 915641,
Avg Contrast = 22. 968489,...
measure of smoothness = 0. 008048,
3rd Moment =-0. 655431,
Uniformity = 0. 022192,
Entropy = 5. 864921

-----------Region 3-----,
Avg Intensity = 170. 193380,
Avg Contrast = 22. 738651,...
measure of smoothness = 0. 007889,
3rd Moment =-0. 847203,
Uniformity = 0. 031366,
Entropy = 5. 382174

The confusion matrix obtained by ANN using RBF

Table 1

	YES	NO
YES	TP	FN
NO	FP	TN

Table 2

	YES	NO
YES	85	2
NO	17	109

Training ratio 80% Efficiency 90. 89% The sensitivity is achieved 97. 70% and specitivity is about 86. 50%
In Main fitnetclassfier

Table 3

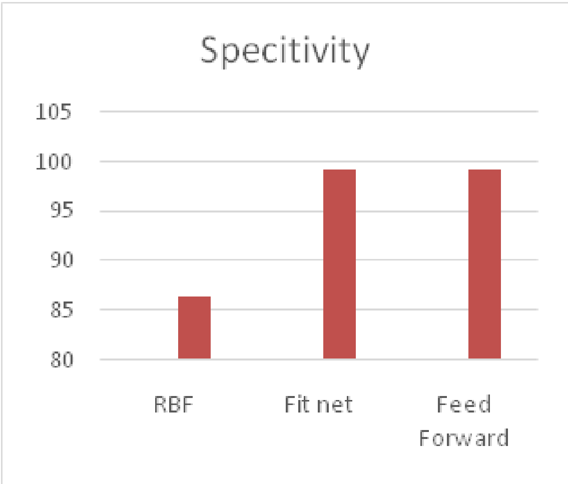
	YES	NO
YES	81	6
NO	1	125

Training ratio 80% Efficiency 95. 53%. The sensitivity is achieved 93. 1%and specitivity is 99. 20%
Feed forward

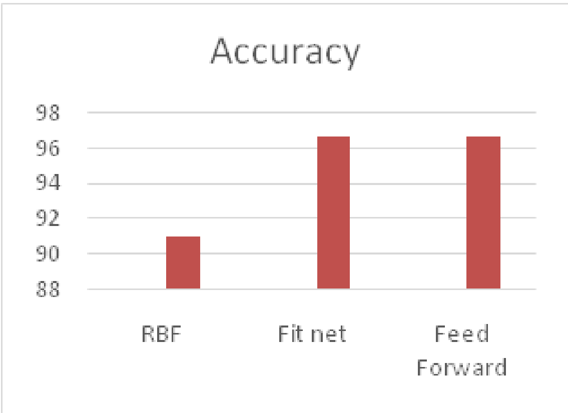
Table 4

	YES	NO
YES	86	6
NO	1	125

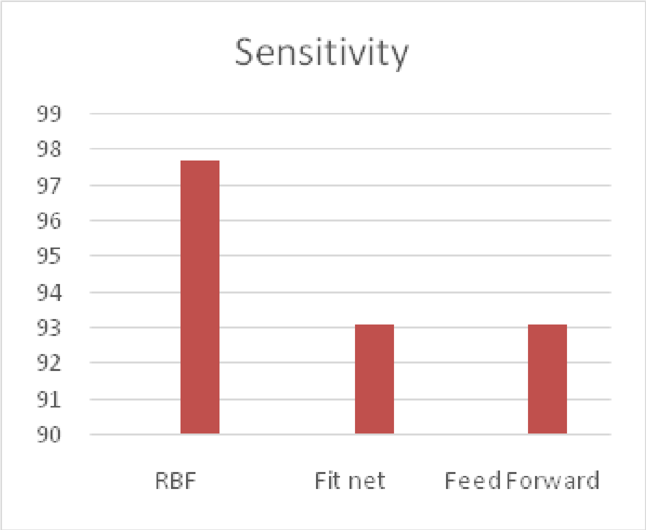
Training ratio 80% Efficiency 95. 53%
The sensitivity is achieved 93. 1%and specitivity is about 99. 20%.



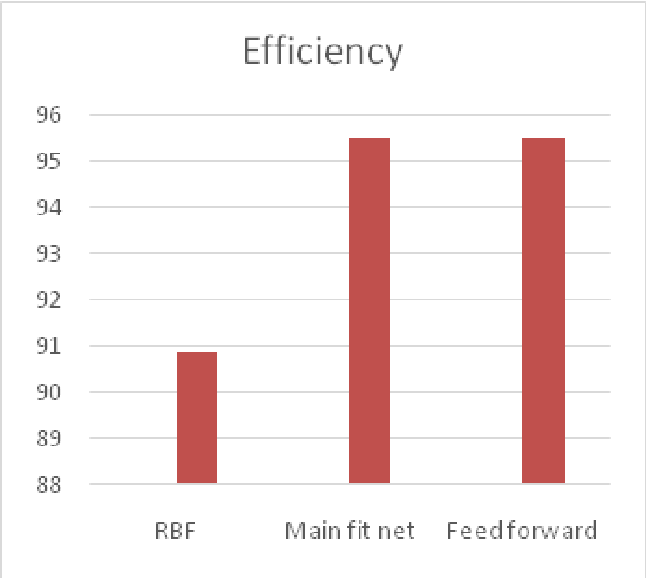
Graph (1)



Graph (2)



Graph (3)



Graph (4)

	Statistical Sequential analysis method	Proposed method
Specificity	97. 8%	99. 20%

Comparison have been made with the method employed in [11] and we have achieved better specificity of 99. 20%.

CONCLUSION

The mammogram images are segmented using texture analysis by using statistical approaches. The desired region of interest is obtained from the marked calcifications on the images. Further 6 features are obtained from the region of interest. These features are classified using RBF, Fitnet,. Main Feed Forward respective efficiency, sensitivity and specificity and accuracy are obtained from the respective classifiers and the obtained results are compared.

ACKNOWLEDGEMENT

The authors are grateful to Dr. Umashankar, Radiologist, Padmashree Diagnostic Center, Bangalore for providing Medical image database used in this study and for his valuable guidance to our work. Also we are thankful to staff of Padmashree Diagnostic Center.

REFERENCES

[1] Mohammed J. Islam, Majid Ahmadi and Maher A. sid-Ahmed "An efficient automatic Mass classification method in Digitized Mammograms using Artifical Neural Network (IJAIA) vol. 1 no. 3 pp. 1-13, 2010.

[2] R Mckenna "The Abnormal Mammogram Radiographic Findings, Diagnostic options, Pathology, and stage of cancer Diagnosis"PP 224-225, 1994 Willey online Library.

[3] Matthew a. Kupinski and Maryellen L Giger, Member, IEEE "Automated SeededLesion Segmentation on Digital Mammograms "Vol. 17, no. 4 pp 510-517, 1998IEEE Transactions on Medical Imaging.

[4] R. Campanini, E. Angelini, E. Iampieri and M. Masotti, " A fast algorithm for intra-breast segmentation of digital mammograms for CAD systems, " IWDM 2004.

[5] H. S. Sheshadri and A. Kandaswamy "Breast Tissue classification using statistical feature Extraction of Mammograms" Medical Imaging and Information Sciences Vol. 23 No. 3 September 2006

[6] J. Skodzik, V. Altmann, B. Wagner, P. Daniels and D. Timmermann, "Highly Integrable FPGA-Baesd Runtime-Configurable Multilayer Perceptron, " 2013

[7] Z. Liu, "Parameter selection for Gaussian Radial Basis Function in Support Vector Machine classification " PP-576-581, 2012.

[8] V. Kecman, Learning and Soft Computing: Support Vector Machines, Neural Networks, and Fuzzy Logic Models, The MIT Press, Cambridge, Massachusetts, London, England, 2001.

[9] T. M. Mitchell, Machine Learning, McGraw-Hill, 1997.

[10] M. J. L. Orr, "Introduction to radial basis function networks, " Technical report, Center for Cognitive Science, University of Edinburgh, 1996.

[11] Jerzy Zielinski, "Statistical Sequential Analysis for Detection of Micro calcifications in Digital Mammograms", IEEE explore PP 1-5 July 2010