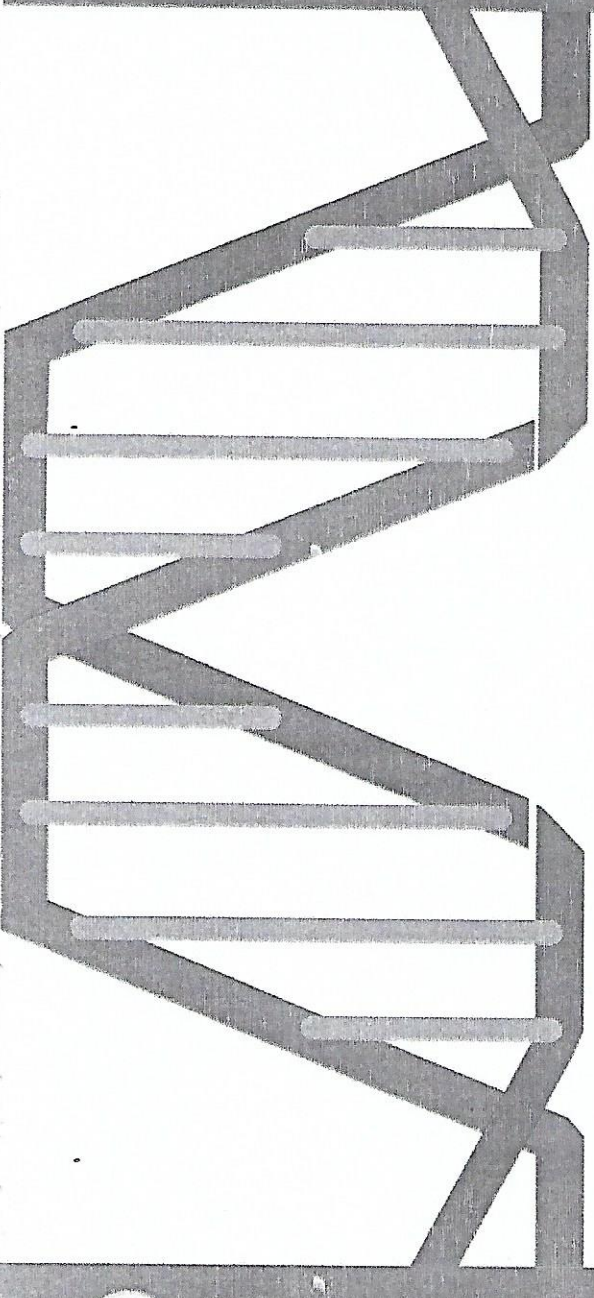
UNIVERSITY OF LAGOS

Journal of Basic Medical Sciences

Volume 2, Number 1 January - June 2014 ISSN: 2354 - 4368



University of Lagos Journal of Basic Medical Sciences is Published bi-annually by Faculty of Basic Medical Sciences of the College of Medicine of the University of Lagos, Idi-Araba

APPLICATION OF ARTIFICIAL NEURAL NETWORKS IN BREAST CANCER CLASSIFICATION: A COMPARATIVE STUDY

[1]S.C. Nwaneri, [2]E.O. Nwoye, [3]N.K. Irurhe and [4]A.M. Babatunde
[1, 2] Biomedical Engineering Department, College of Medicine, University of Lagos, Nigeria
[3, 4] Radiology Department, College of Medicine, University of Lagos, Nigeria

Correspondence Author
S.C Nwaneri
Email: snwaneri@unilag.edu.ng

ABSTRACT

Background: Breast cancer is a leading cause of death especially among women globally. The classification task of breast lump as benign or malignant is due to the experience and skill of the radiologist. However, Artificial Neural Networks (ANNs) can be developed to assist radiologists in decision making.

Objective: The purpose of this study is to develop ANN based models for breast cancer classification.

Method: The five features of retrospective breast ultrasound data obtained from Lagos University Teaching Hospital (LUTH) consisting of 83 samples were rated using Breast Imaging Reporting and Data system (BI-RADS).The data was normalized and trained in MATLAB software version (R2009a)using a feedforward multilayer ANN with 5 inputs neurons, 10 hidden neurons and one output neuron. The hidden neurons were increased in steps of 10 for different iterations to a maximum of 100 neurons in the hidden layer. The well known Wisconsin Breast Cancer Data (WBCD) comprising 699 samples of digitized data was also trained with the same algorithm and parameters.

Results: The results show that ANNs performance in both cases was quite high. It was also proved that there was no direct relationship between the performance of the network and the number of hidden neurons.

Conclusion: ANNs are efficient classifiers that can be utilized in the diagnosis of breast cancer in the country.

Keywords: Breast Cancer, Artificial Neural Networks, Breast Imaging Reporting and Data System (BI-RADS), Ultrasound, Radiologists.

1.0 INTRODUCTION

Breast cancer is a malignant tumor that has developed from cells in the breast. It is responsible for a high death rate among women suffering from cancer related deaths globally. Available statistics show that in 2007 about 202,964 women in the United States were diagnosed with breast cancer out of which 40,598 died from the disease. Symptoms vary from lumps to swelling to skin changes. Sometimes no clear symptoms are noticeable (1). In Nigeria, accurate and detailed statistics on the incidence of breast cancer is not available due to lack of statistics or under-reporting. This is evident as revealed in a study on Cancer registry literature update, in which only 1% of the literature originated from Africa compared to

34% and 42% from Europe and Asia respectively (2). The detection of breast cancer in most Nigerian women is usually at advanced stages when the disease would have spread to other parts of the body. In a study conducted in the eastern part of Nigeria, breast cancer was found to occur in 45% of patients aged 40 years and below (3). Despite the intense research which is still on-going, medical scientists are yet to discover methods of prevention and cure for the disease However, it has been proved that early detection and treatment may improve the chances of survival by preventing the spread of the disease to other cells. There are some early signs of having high possibility of cancerous cells like microcalsifications, mass, architectural distortion and breast asymmetries (4).

While mammography is considered the most effective method for the early detection of breast cancer, marnmograhic interpretation is characterized by a number of problems (5). Observational interpretation of mammographic findings has different accuracy problems mostly determined by the experience and skill of the radiologist. Screening mammography is associated with potential problems including inaccurate results for example, a benign tumor may be wrongly diagnosed as malignant (false positive result), or the wrong diagnosis of a malignant tumor as benign (false negative).

When abnormal features are detected during mammographic examination such as unusual density, odd shapes, and irregular border, the patient are then referred for biopsy. A biopsy is mostly carried out to confirm if a tumor is cancerous or not. However, only a few cases that have mammographic suspicious findings subjected to biopsy proved to be malignant. Computer aided schemes have been used as classifiers by several researchers in distinguishing between benign and malignant patterns in the interpretation of mammographic images (5)(6)(7)(8).

Despite the frequent use of mammography as a method of early detection of breast cancer, ultrasound imaging has been used as a complement. Ultrasound is commonly used for the detection and diagnosis of cysts. The functions of ultrasound have been expanded to include its use in breast cancer diagnosis. The new techniques have been used to improve the image quality of ultrasound machines. Consequently, the use of ultrasound is considered more effective for women less than 35 years of age. Several attempts have been made in the past to characterize breast tissue by complex signal and image processing schemes (9)(10).

In this work, a comparative study was done to evaluate the performance of training ultrasound scans and WBCD using ANN.

MATERIALS AND METHODS

This is a simulation study utilizing ANNs to train clinical data obtained from LUTH as well as WBCD obtained from UCI machine learning database. A Simulation study attempts to imitate the operation of a real-world process or system over a given period of time. The purpose of the design is to demonstrate the application of ANN in breast cancer classification. This was achieved using MATLAB 7.8.0(r200a) software. The stages involved in this research are as stated in the flowchart in fig 1 below. This includes the collection of data, preprocessing of data, coding in MATLAB, training and simulation and testing.

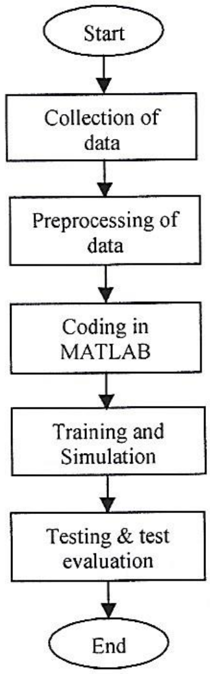


Fig. 1: Process flow diagram of Design Stages for breast cancer diagnosis

EXPERIMENTAL SETUP

Ultrasound Breast Cancer Data

The experimental setup is hereby presented. This includes the different stages, work plan and procedures for carrying out the research. In this research, an Intel(R) HP 620 T4500 dual core notebook with a RAM capacity of 3GB and Windows 7 Home Premium operating system was used for the implementation. The algorithms were coded in MATLAB 7.8.0(R2009a).

In this study, retrospective data was obtained from the radiology department of Lagos University Teaching Hospital for breast ultrasound examination carried out between January 2012 and January 2013 for 83 female patients out of which 19 were diagnosed as malignant and 64 were diagnosed as benign which were confirmed by histopathology. The ages of the patients range from 18-60years. The data consists of results of breast ultrasound imaging in which BI-RADS values had been assigned to the data by radiologists. The features used were the patients' ages and BI-RADS descriptors. The BI-RADS values are derived from features of the breast ultrasound images. The features include spiculation, shape, acoustic shadow, echogenicity, angular margins, and microlobulations. The approach used by(11) was adopted in this study to assign BI-RADS values as presented in tables 1, 2 and 3 respectively.

Table 1: BI-RADS Values of Breast Ultrasound for Shape Features

S/N	Features	BI-RADS Values
1.	Round shape	1
2.	Oval shape	2
3	Lobular	3
4	Irregular	4

Table 2: BI-RADS Values of Breast Ultrasound for Margin Features

S/N	Features	BI-RADS Values
1	Circumscribed	1
2	Microlobulated	2
3	Obscured	3
4	Ill-defined	4
5	Spiculated	5

Table 3: BI-RADS Values of Breast Ultrasound for Chogenic Features

S/N	FEATURES	BI-RADS
1.	Hyperchoic	1
2	Isoechoic	2
3	Hypoechoic	3
4	Fat- containing	4

Wisconsin Breast Cancer Database (WBCD)

The Wisconsin breast cancer database (WBCD) which is a well known source of data developed by Dr. William Wolberg of

the University of Wisconsin and obtained from University of California Irvine was used as control. The WBCD has been used by several researchers in pattern recognition and machine learning research (12)(13)(14)(15).
The database consists of 699 samples. The dataset contains 9 attributes as shown in table 4 below were assigned values on a scale of 1 to 10. Lower values were assigned to features with benign tendencies while the degree of malignancy increased as the values increased.

Table 4: Attribute Information of WBCD

S/N	Attribute	Domain
1	Clump Thickness	1 - 10
2	Uniformity of Cell	1 - 10
3	Marginal Adhesion	1 - 10
4	Single Epithelial	1 - 10
5	Single Epithelial	1 - 10
6	Bland Chromatin	1 - 10
7	Normal Nucleoli	1 - 10
8	Mitoses	1 - 10

The dataset consists of samples obtained by Fine Needle Aspiration from 699 patients in which 458 were confirmed to be benign and 241 malignant.
The WBCD comprises features computed from a digitized image of the FNA of breast masses. The FNAs were prepared by extracting viscous fluid from breast masses using syringes. The aspirates were imaged by a video camera and converted to computer files which were displayed on the computer monitor. The nuclei to be analyzed were roughly outlined by an operator using a mouse. The computer software generated a "snake" that enclosed each designated nucleus. The computer calculated 10 features for each nucleus. These features or attributes describe the characteristics of the cell nuclei present in the image. Dr. H.Wolberg and his team analyzed each cell by using computer software to compute the nuclear features of each cell as shown in table 1 above.

Parameter Settings
The data was trained with a multilayer feed-forward ANN consisting of 3 layers namely input, hidden and output layers was used for training the data using the Neural Network Pattern Recognition tool in the Graphical User Interface (GUI) of MATLAB. The input layer consisting of 5 neurons which are breast attributes in each case. The number of neurons in the hidden layer was initially set to 10 neurons and gradually increased in steps of 10 until the number of neurons became 100.A single output layer was employed to determine whether the lump was benign or malignant. A tansigmoid transfer function was used as the activation function for the hidden and output layers while a linear transfer function was use for the input layer. The maximum number of epochs was set to 1000.

MATHEMATICAL EQUATIONS

(i) Computation of the Training Algorithm

The ANN was trained with the Levenberg-Marquardt algorithm computed using the following formula:

$$J^T J \delta_p = J^T \epsilon \qquad (1)$$

Where J is the Jacobian matrix, J^T the transpose value of the jacobian matrix, and ϵ the error.

(ii) Computation of the Mean Squared Error:

The Mean squared error measures the network's performance according to the mean of squared errors. The error is calculated as the difference between the target output and the network output. The goal is to minimize the average of the sum of these errors.

$$\mathbf{MSE} = \frac{1}{Q}\sum_{k=1}^{Q} e(k)^2 = \frac{1}{Q}\sum_{k=1}^{Q}(t(k) - a(k))^2 \qquad (2)$$

Where e= error or difference between the target(desired output) and actual output

t(k) = desired output
a(k) = actual output
The LMS algorithm adjusts the weights and biases of the network in order to minimize this mean square error.

Training, Validation and Testing
The breast ultrasound data and WBCD were trained separately in Neural Network Pattern Recognition toolbox in MATLAB (R2009a) version.The data set was divided into training (70%), validation (15%) and testing (15%) in each case. This was accomplished by writing m-file codes in each case and running the program. The progress of the training was displayed after running the m-file codes. In addition, fig. 2 also displayed important information as the epoch, the duration of the training, the performance levels, the gradient and the validation checks.

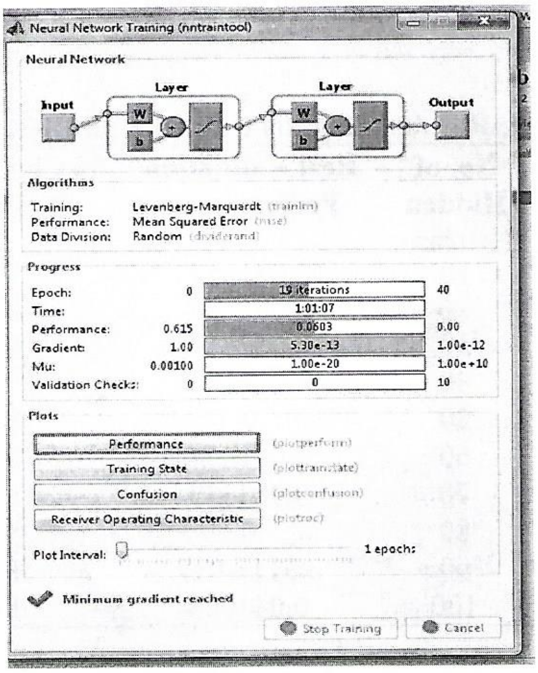


Fig. 2 Display of training in progress

RESULTS AND DISCUSSION

ANN Training Results For Breast Ultrasound Data
The results obtained for training input features of breast ultrasound imaging obtained from LUTH by varying the number of hidden neurons in steps of 10 are presented in a tabular form in table 5 below:

Table 5: Results for ANN trained with BI-RADs values of Ultrasound Breast Images

S/N	No. of Hidden Neurons	Best Validation Performance	% Classification Rate	Sensitivity	Specificity
1	10	0.041657	97.6	94.1	**98.5**
2	20	0.083333	97.6	94.1	**98.5**
3	30	0.047417	95.1	76.5	**100**
4	40	0.083333	93.9	88.2	**95.4**
5	50	0	97.6	94.1	**98.5**
6	60	0.041677	92.7	70.6	**98.5**
7	70	0.083333	97.6	94.1	**98.5**
8	80	0	94.1	96.9	**96.3**
9	90	3.311e-017	97.6	94.1	**98.5**
10	100	0.041608	96.3	94.1	**96.9**

ANN Training Results for WBCD

Table 6: Results for ANN trained with WBCD

S/N	Number Of Hidden Neurons	Best validation Performance	% Classification Rate	Sensitivity	Specificity
1	10	0.039594	98.6	98.7	98.3
2	20	0.035378	98.6	99.1	97.5
3	30	0.082304	97.6	99.1	94.6
4	40	0.021511	97.7	97.6	97.9
5	50	0.024808	98.6	99.3	97.1
6	60	0.013496	99.4	99.6	99.2
7	70	0.019373	96.6	98.0	93.8
8	80	0.046460	98.0	99.1	95.1
9	90	0.004618	98.4	98.5	98.3
10	100	0.038102	96.6	97.6	94.6

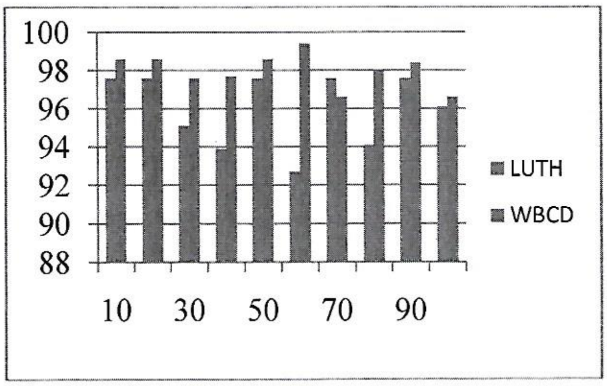


Fig. 3 Bar chart showing the performance of ANN training for Breast ultrasound Data and WBCD

The bar chart in fig. 3 shows the results of training for the ANN models for the breast ultrasound data from LUTH and the WBCD. The blue and red colored bars represent the percentage accuracy of classification for the ultrasound data from LUTH and WBCD respectively. It was observed that the accuracy of classification was high in all cases with slight variation in accuracies for the different number of hidden neurons. Also, the WBCD performed slightly better in all cases.

STATISTICAL ANALYSIS

The results of the ANN training for both sets of data were therefore analyzed in order to determine the level of significance of the difference in results for the two sources.

Table 7: Group Statistics

Grouping	N	Mean	Std. Dev.	Std. Mean Error
Ultrasound	10	96.0100	1.90348	.60193
WBCD	10	98.0100	.90486	.28614

The above table 7 displays the mean, standard deviation, and standard error for both groups. An independent sample t-test was done to determine whether the difference between sample averages may likely represent an actual difference between populations and also whether the difference is significant.

The mean of the classification rate were 96.01% and 98.01% respectively for the ultrasound results and WBCD respectively. The standard deviations for each network training for ultrasound results and WBCD data revealed that the classification rate were more variable with respect to the ultrasound scans than to WBCD data.

The 95% Confidence Interval of the Difference provides an estimate of the boundaries between which the true mean difference lies in 95% of all possible random samples of network training for 10 different times similar to the ones involved in this study. The t statistic is achieved by dividing the mean difference by its standard error. The Sig. (2-tailed) column displays the probability of obtaining a t statistic whose absolute value is equal to or greater than the obtained t statistic. From the table above, the Levene's Test for equality of variances determine if the ultrasound and WBCD results above have about the same or different amounts of variability between the classification rates. The second row was selected giving a p-value of 0.010. This means that the difference in means is statistically insignificant at the 0.005 level.

DISCUSSION

ANNs are highly efficient pattern recognition tools used in medical diagnosis. They consist of distributed parallel processing systems that can acquire knowledge and make it available for use (16). The application of ANNs in breast cancer diagnosis has received considerable research attention. In this study, feed-forward multilayer ANN was used in breast cancer classification based on input features fed into the network. Two different sets of data were trained using similar parameters as earlier stated.

Table 8 Independent Samples Test

	Levene's Test for Equality of Variances		t-test for Equality of Means						
	F	Sig.	t	df	Sig. (2-tailed)	Mean Diff.	Std.Error Difference	95% Confidence Interval of the Difference	
Equal variances assumed	10.181	.005	-3.001	18.0	.008	-2.00	.66648	-3.40023	-.59977
Equal variances not assumed			-3.001	12.87	.01	-2.00	.66648	-3.44133	-.55867

The two sets of data include:
1. Retrospective data obtained from breast ultrasound imaging.
2. Wisconsin Breast Cancer Datasets (WBCD)

The results indicated in tables 5 and 6 shows that in all cases the ANNs gave good performance. This means that based on the input features fed into the networks, the ANNs were able to predict whether the tumors were malignant or benign, except in some few instances of misclassification. Therefore it can be a valuable tool for radiologists and other medical practitioners in breast cancer diagnosis especially in few screening cases where the patients may not need to undergo the procedure of biopsy to confirm that their tumor is benign.

Also, the results also show that there was slight variation in the accuracy of classification as the number of hidden neurons was increased. However, an increase in the number of hidden neurons did not correspond to an increase in the accuracy of network classification. Hence it was established that the accuracy of classification is not proportional to an increase in the number of hidden neurons.

A comparison between the results of both sets of data revealed a slight variation in the accuracy of classification. For the ultrasound data, the least accuracy observed was 92.7% when the network was trained with 60 neurons in the hidden layer, while the best accuracy observed was 97.6% in 4 different instances of 10, 20, 70 and 90 neurons in the hidden layer. On the contrary, training the WBCD gave different results for the different number of hidden neurons. The lowest accuracies were observed when the network was trained with 70 and 90 hidden neurons, while the best accuracy of 99.4% was observed for 60 hidden neurons. The best and worst performance was not recorded for the same number of hidden neurons. It follows therefore, that in both cases there was no proportional increase in the accuracy of performance with increase in the number of hidden neurons. Hence, there is no guarantee that increasing the number of neurons in the hidden layer will improve the accuracy of classification. Similarly, it was revealed in both cases that the sensitivity values were closely related to the accuracy of classification of the networks. Therefore it was observed that the worst values for accuracy and sensitivity were recorded for networks with 60 hidden neurons for the ultrasound data. Table 6 indicates similar results for the WBCD as the best values were obtained for the network with 60 hidden neurons. The training results have demonstrated the effectiveness of the LM backpropagation training algorithm used in this study, which was proved in previous studies to be faster for medium sized problems than all other training algorithms (14). The best value of 99.4% was achieved which is quite high when compared to the value of 99.28% is obtained by (17).

CONCLUSIONS

A detailed study was carried out on the application of ANNs in breast cancer classification. The need to apply machine learning algorithms to classify breast tumors either as malignant or benign is of utmost importance to avoid the subjectivity associated with human judgment. ANNs learn by experience and have the ability to generalize. Therefore, they can easily predict whether a tumor is benign or malignant. If properly applied in the diagnosis of breast cancer in Nigeria, it will assist radiologists in decision making by improving their accuracy of diagnosis. This will help reduce mortality and patients being subjected to the traumatic experience of having to undergo biopsies due to false positives.

Furthermore, the study revealed that BI-RADs values can be assigned to observed features of breast ultrasound imaging and applied as input values to ANNs and used to classify a tumor as malignant or cancerous. However, a high level of accuracy is required in the assignment of the BI-RADs values to avoid inaccurate results. The skill and experience of the radiologist will to a large extent determine the accuracy of these inputs. it was further proved that there is no significant difference in the accuracy of classification using either the well known WBCD or local ultrasound data in which accurate BI-RADS values have been assigned to the features.

Finally, the study was able to show that training a feed-forward multilayer neural network using LM backpropagation algorithm, there is no direct relationship between the accuracy of classification and the number of hidden neurons. There may be variations in performance for the different number of hidden neurons; no relationship was established between the network performance and the number of hidden neurons.

REFERENCES

1. Centre for Disease Control and Prevention. *Breast Cancer Statistics* viewed 25 January 2012. 2010. http://www.cdc.gov/cancer/breast/statistics/.

2. Bare-Bones. Literature Review: Cancer Registration Literature Update 2006-2008. *Asian Pacific J of Cancer Prevention*. 2008; **9**.

3. Ekanem VJ, Aligbe JU. Histopathological Types of Breast Cancer in Nigerian Women: A Twelve Year Review 1993 – 2004. African *J of Reproductive Health.* 2006; **10(1).**

4 Adam A, Omar K. Computerized breast cancer diagnosis with Genetic Algorithm and Neural Networks. 2007.

5. Wu Y, Giger ML, Doi K, Vyborny CJ, Schmidt RA, Metz CE. Artificial Neural Networks In Mammography: Application to Decision Making In the Diagnosis Of Breast Cancer. 1993.

6 Mohammed WA, Kadah YM. Computer Aided diagnosis of

digital mammograms. *IEEE Trans. Biomedical Engineering,* 2007; 1-1366.

7. Chang Y, BinZheng Wang X, Good W.F. Computer-Aided Diagnosis of Breast Cancer Using Artificial Neural Networks: Comparison of Backpropagation and Genetic Algorithms. *in Proc. IJCN*:1999; 3674 – 3679.

8 Irurhe NK etal. Accuracy of Ultrasonic in detecting breast Cancer. NQSHP 1(5) 30012 or (1) 2013

9. Song JH, Venkatesh SS, Conant EF, Cary TF, Arger PH, Sehgal CM. Artificial Neural Network to aid differentiation of malignant and benign breast masses by ultrasound imaging. 2005.

10. Minavathi Murali. S. Dinesh M.S. Classification of Mass in Breast Ultrasound Images using Image Processing Techniques. *Int J of Computer Applications.* 2012; **42(10)**: 0975 – 888.

11. Duffy SW, Tabar L, Chen HH. The impact of organized mammography service screening on breast carcinoma mortality in seven Swedish counties. Cancer 95. 2002; 458-469.

12. Elgader HAA, Hamza M.H. Breast Cancer Diagnosis Using Artificial Intelligence Neural Networks. *J of Sci. and Tech* 2011; **12(1)**.

13. Raad A, Kalakech A, Ayache M. Breast Cancer Classification Using Neural Network Approach: MLP and RBF. *The Int Arab Conf on Info Tech ACIT* 2012.

14 Paulin F, Santhakumaran A. Classification of Breast cancer by comparing Backpropagation algorithms. *Int J on Comp Sci and Eng (IJCSE).* 2011.**3:** 327 - 332.

15. Wolberg WH, Street WN, Heisey DM, Mangasarian OL. Computer Derived Nuclear Features Distinguish Malignant From Benign Breast Cytology. Gordon and Breach Science Publishers. 1992; 21-32.

16. Theodoridis S, Koutroumbas K. *Pattern Recognition.* Fourth Edition, Elsevier Inc., California, USA. 2009.

17. Paulin F, Santhakumaran A, Classification of Breast cancer by comparing Backpropagation algorithms. *Int J on Comp Sci. and Eng (IJCSE) 2010;* **3:** 327 - 332.