Experimental UWB based Efficient Breast Cancer Early Detection

V. Vijayasarveswari¹, M. Jusoh¹ and S. Khatun²

¹Embedded Network and Advance Computing Research Cluster (ENAC), School of Computer and Communication Engineering, University Malaysia Perlis, Perlis, Malaysia; vijayajo@yahoo.com, muzammil@unimap.edu.my
²Faculty of Electrical and Electronic Engineering, University Malaysia Pahang, Pekan, Pahang, Malaysia; sabira@unimap.edu.my

Abstract

Objectives: The effective detection system can help to detect breast cancer in the early stage. Focus of this paper is to improve Ultra Wideband (UWB) based early breast cancer detection efficiency. Methods/Statistical Analysis: Hence, a set of 125 data samples was created through forward scattering by placing two home-made UWB antennas between breast phantoms with and without tumor. Samples were divided into 5 groups with 25 samples each. K-fold cross validation method was used to train, test and validate the data set based on Feed Forward Back Propagation Neural Network (FFBPNN). Findings: The developed system detect breast tumor with average accuracy of (1) 100%, 78.17%, 70.66%, 92.46% and 86.86% and (2) 100%, 85.17%, 87.89%, 94.11% and 83.86 for existence, x, y, z and size respectively for (1) FFBPNN and (2) k-fold cross validation based FFBPNN. K-fold cross validation based FFBPNN showed in average 90.23% detection efficiency in terms of tumor existence, location (x, y and z) and size) which is nearly 5% improved compared to conventional FFBPNN (85.43%). The detected tumor is visualized in 2D and 3D environment. Application/Improvements: Hence, K-fold cross validation FFBPNN can be a future candidate as better and faster breast cancer detection system and saves precious lives.

Keywords: Breast Cancer, Feed Forward Back Propagation, K-fold Cross Validation, Neural Network, UWB

1. Introduction

Breast cancer is a common disease among women. This is a type of disease needs early detection and diagnoses to prevent death. Many traditional detecting techniques like mammogram, MRI scan and ultrasound are available. But these types of techniques are expensive and have a high false negative (high chances of wrong decision) detection probability. Moreover, those are not safe to use by patients regularly due to ionization effect and painfulness.

Ultra-Wideband (UWB) technology is widely used throughout the world to replace (and or complement) the traditional techniques as a breast screening tool. It has the ability to detect breast cancer in early stage effectively. It is low-cost, secure, non-ionizing and non-invasive. Although UWB technology provided as a safe breast screening tool, but accurate and precise diagnosis are in concern. Artificial Neural Network (ANN) is already being widely used in medical sector such as in diagnosis, biochemical analysis, drug development and image analysis. ANN can be categorized as a precise model based on its performance, accuracy, classification or detection ability, and cost.

Zahra Nematzadeh et al. has analyzed and compared various types of machine language (Naive Bayes, Neural Network (NN) and Support Vector Machine (SVM)) based k-fold cross validation method for breast cancer classification. It was proved that more folds contribute to better accuracy but high cost. S. A. Mojarad et al. has investigated the effectiveness of Multilayer Perceptron NN (MLP) based k-fold cross validation in classifying breast cancer. They proved that an artificial neural network is the best candidate and gives better detection.
However, this accuracy is so far approximately 60% and can lead more false negative detection. Hamid Haji et al. showed Least Square Support Vector Machine (LS-SVM) technique was the best machine language in classifying breast cancer with 97.81% accuracy. However, all these the above research was just focused on detecting the existence of breast cancer with available Wisconsin Breast Cancer Dataset (WBCD). Different machine languages or techniques based k-fold cross validation was investigated but feed forward NN was left.

NN can achieve better accuracy and precise model if the input data samples are large enough. It is because a limited number of data samples will lead to insufficient training and decrease the performance accuracy of the model. This leads to wrong decision making in determining the existence of cancer. This paper focuses to overcome this shortcoming by training the NN using available data samples through k-fold cross validation method based feed forward back propagation NN. It can help detection system to make better decision in order to find cancerous cells in breast tissue. This is to estimate how k-fold cross validation method can be used in small data samples to improve model/system quality. At the same time it can detect the existence, location and size of tumor in breast tissue.

2. Material and Method

Designed breast phantom is based on the real dielectric properties of breast as shown in Table 1. This research is done using UWB transceivers, a pair of home-made antenna, and PC interfacing. The detail methods taken into account is shown in Figure 1 and presented in the following sections.

2.1 Data Samples Collection

Based on, breast phantom and tumor were developed using low-cost and available non-chemical ingredients. The heterogeneous breast is developed using the combination of petroleum jelly, flour and water. This is to make real-like of breast. The experimental set-up is shown in Figure 2, whereby, breast phantom was placed in between two home-made antennas (one as transmitter while another as receiver). Antennas were connected to UWB transceiver. Transmitted signals were travelled through the breast phantom and received by the receiver antenna. The total of 140 forward scattered signals was captured by the receiver from other end of phantom. The received signals were saved in PC and processed to obtain related discrete 1632 data points per sample. The data collection was done repeatedly for various phantoms with different tumor location and sizes.

2.2 Artificial Neural Network

Considered ANN consists of input, hidden neuron and output as shown in Figure 3. A feed forward back propagation NN was developed using Matlab software, where input moves forward in one direction. So, data inputs go through some hidden nodes with different hidden layer (if exists) and finally to output node. The performance of the NN module can be increased by changing the number of hidden neurons. However the number of hidden neurons should not be

![Figure 1. Overall System.](image-url)
very high because it may consume more time to compute. The numbers of optimized input nodes, hidden layer and output nodes were 4, 20, and respectively as shown in Figure 3. The performance of the NN module was examined with collected data samples. Mean Square Error (MSE) of NN output was calculated by subtracting output from target data\(^1\).

**Table 1. Dielectric Properties of Considered Breast Phantom\(^2\)**

| Breast Phantom Part | Material                        | Permittivity | Conductivity |
|---------------------|--------------------------------|--------------|--------------|
| Fatty tissue        | Pure petroleum jelly           | 2.36         | 0.012        |
| Tumor               | Mixture of water and wheat flour (55%) | 23           | 2.57         |
| Skin                | Glass                          | 3.5–10       | Negligible   |

**Figure 2. Experiment Set-up.**

**2.3 K-fold Cross Validation and Training**

In this paper, 5-fold cross validation was considered. Then it was interfaced with feed forward back propagation NN. This new module was trained using same data samples to investigate its performance. The training data set consists of 125 data samples. Each data sample has 1632 elements that encode the existence, location and size of a breast tumor/cancer. This data samples are divided into 5 small sets. Each set consist about 25 data samples. Sets are divided for training and testing as shown in Table 2 below.

**Table 2. Training and Testing Set**

| K  | Training | Testing |
|----|----------|---------|
| 1  | 1        | 2       |
| 2  | 1        | 2       |
| 3  | 1        | 2       |
| 4  | 1        | 3       |
| 5  | 2        | 3       |

**Figure 3. Architecture of ANN.**

**2.4 Signal Processing and Feature Extraction**

Before training, through feature extraction 1632 elements were reduced to 4 elements for a data sample. Feature extraction is done to decrease computing time and simpler structure. Different type of feature extraction is tested but the best result is only can be obtained by using the method as in\(^4\). Since the elements are in small group after feature extraction, the training can optimize and give the better result. The extracted features were mean, median, maximum value and minimum value of each data sample. Initial train was done using conventional feed forward back propagation NN, followed by using k-fold cross validation method until the performances of the NN modules were optimized. The performance of the NN module can be increased by changing the number of hidden neuron. However the number of hidden neurons should not be very high because it may consume more time to complete. The amount of used input, hidden and output layers were 4 (after extracting the feature from 1632 data points), 20 (one hidden layer), and 4 (trained for each output).
Experimental UWB based Efficient Breast Cancer Early Detection

separately) respectively as shown in Figure 3. The NN parameter is used as shown in Table 3.

**Table 3. NN Parameter for Matlab Training**

| NN Parameters                  | Value          |
|-------------------------------|----------------|
| Number of nodes in Input layer | 4              |
| Number of nodes in Hidden layer 1 | 20            |
| Number of nodes in Output layer | 4              |
| Transfer function             | tansig         |
| Training function             | traingdm       |
| Learning rate                 | 0.009          |
| Momentum constant             | 0.6            |
| Maximum number of Epochs      | 100000         |
| Minimum performance gradient  | 1e-25          |

Performance for both modules is examined. The detection performance accuracy is calculated as in Equation 1. Example of accuracy calculation for output 20.33 and for actual target is 20 in Table 4.

\[
\text{Accuracy (\%)} = \left(1 - \frac{|O - A|}{100}\right) 
\]

where O is the NN’s output and A is actual target.

**Table 4. Accuracy Calculation**

| Output | Actual Target | Calculation | Accuracy (%) |
|--------|---------------|-------------|--------------|
| 20.33  | 20            |             | 98.35        |

### 3. Results and Discussion

After completing the train of the NN modules, real time testing was done to observe the performance efficiency
of various tumors/cancer locations ($x, y, z$), size and existence detection.

Table 5 shows the testing results in terms of targeted output and NN output for both NN modules. Here to mention that, any output (location ($x,y,z$) and/or size) with −ve value indicates absence of tumor in breast tissue (i.e., healthy breast as shown in shaded rows in Table 5). The presence of the tumor in breast tissue can be determined with the location and size with +ve NN outputs only. It also indicates the 100% tumor existence detection accuracy as each individual −ve and +ve input (actual target) accurately resulted −ve and +ve output respectively.

The conventional feed forward back propagation NN module detection performance accuracy for tumor existence, location ($x,y,z$), and size is shown in Table 5.

Based on Table 6, the tumor can be detected with 100% in terms of existence. It is hard to find the exact location and size of the tumor because of its complex structure. Since, real human breast consists of glandular, heterogeneous breast need to be considered here.

### Table 6. Overall Performance Accuracy

| Method              | FeedForward Back Propagation NN | K-Fold Cross Validation |
|---------------------|---------------------------------|-------------------------|
| Existence           | 100                             | 100                     |
| Location $x$        | 78.17                           | 85.17                   |
| Location $y$        | 70.66                           | 87.89                   |
| Location $z$        | 92.46                           | 94.11                   |
| Size                | 85.86                           | 83.96                   |

Table 7 shows the overall average performance accuracy comparison between conventional BPFFNN module and k-fold cross validation based BPFFNN module. It can be seen that the performance accuracy has been increased from 85.43% to 90.23% by using k-fold cross validation, which is around 5% improvement by showing its superiority.

### Table 7. Overall Performance Accuracy

| Method                                  | Overall Training Performance | Performance Accuracy (%) |
|-----------------------------------------|------------------------------|--------------------------|
| Feed forward back propagation           | 3.1209                       | 85.43                    |
| K-fold cross validation based feed forward back propagation | 9.1495                       | 90.23                    |

One of the detected tumor with size if 2mm at location (32.5mm, 62.5mm, 40mm for $x, y, z$ respectively) is visualized in 2D and 3D environments using the developed detection system in Figure 4 and Figure 5 respectively. The system is independent software to replace existing costly UWB based system. The detection results are shown at right below side of the image.

![Figure 4. 2D Imaging.](image-url1)

![Figure 5. 3D Imaging.](image-url2)
4. Conclusion

K-fold cross validation method is combined with feed forward back propagation NN module for improved UWB based breast tumor/cancer detection. The proposed method is suitable for a small number of data samples with improved efficiency as expected.

Our present focus is to examine the various K values for optimization and to obtain the best accuracy in terms of early tumor/cancer detection.

5. Acknowledgement

This work is supported by Ministry of Higher Education, Malaysia, Grant FRGS – 9003-00418

6. References

1. World Health Organization. Date accessed: 05/07/2014: Available from: http://www.who.int/topics/cancer/en/
2. Huynh PT, Jaroilme AM, Daye S. The false-negative mammogram. Radio Graphics. 1998 Sep; 18(5):1137-54. Available from: Crossref. PMid:9747612
3. Dimitrios Siganos. Neural networks in medicine. Date accessed: 15/01/2016: Available from: http://www.doc.ic.ac.uk/~nd/surprise_96/journal/vol2/ds12/article2.html.
4. Nematzadeh Z, Ibrahim R, Selamat A. Comparative studies on breast cancer classifications with k-fold cross validations using machine learning techniques. 2015:10th Asian Control Conference (ASCC). 2015 May 31; p.1-6. Available from: Crossref
5. Mojarad SA, Day SS, Woo WL, Sherbet GV. Breast cancer prediction and cross validation using multilayer perceptron neural networks. 2010. 7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP). 2010 Jul 21; p. 760-64.
6. Fiuji H, Almasi BN, Mehdkhan Z, Bibak B, Pilevar M, Almasi ON. Automated Diagnostic System for Breast Cancer Using Least Square Support Vector Machine. American Journal of Biomedical Engineering. 2013; 3(6):175-81.
7. AlShehri SA, Khatun S and et al. Experimental breast tumor detection using NN-based UWB imaging. Progress in Electromagnetics Research. 2009; 111:79-93. Available from: Crossref
8. AlShehri SA, Khatun S, Jantan AB, Raja Abdullah RS, Mahmud R, Awang Z. Experimental breast tumor detection using NN-based UWB imaging. Progress in Electromagnetics Research. 2011; 111:447-65. Available from: Crossref
9. AlShehri SA, Khatun S, Jantan AB, Raja Abdullah RS, Mahmud R, Awang Z. 3D experimental detection and discrimination of malignant and benign breast tumor using NN-based UWB imaging system. Progress in Electromagnetics Research. 2011; 116:221-37. Available from: Crossref
10. Khan ZH, Mohapatra SK, Khodiar PK, Ragu KS. Artificial neural network and medicine. Indian journal of physiology and pharmacology. 1998 July; 42(3):321-42. PMid:9741647
11. Albu A, Ungureanu L. Artificial neural network in medicine. Symposium on Applied Computational Intelligence. 2005; p. 2-5.
12. A basic introduction to feedforward back propagation neural networks. Date accessed: 16/5/2016: Available from: http://www.webpages.ttu.edu/dleverin/neural_network/neural_networks.html.
13. Difference between feed forward and feed forward back propagation. Math Works: (N.D.). Date accessed: 16/1/2016: Available from: http://www.mathworks.com/matlabcentral/answers/179038-difference-between-feed-forward-feed-forward-back-propagation.