Feature Extraction Performance Verification based on Ultra-wideband Imaging and Artificial Neural Network

V.Vijayasarveswari[1], M.Jusoh[1*],T.Sabapathy[1],M.N.Osman[1], R.A.A.Raof[1], S.Khatun[2], M.N.Yaasin[1]
{muzammil@unimap.edu.my}

[1]Bioelectromagnetics Research Grant (BioEM), School of Computer and Communication Engineering, University Malaysia Perlis, Perlis, Malaysia
[2]Faculty of Electrical & Electronic Engineering, Universiti Malaysia Pahang, Pekan, Pahang, Malaysia

Abstract: Breast cancer is a second leading case among the women. Therefore, an efficiency breast cancer system is very important for accurate early detection. The selection of number of features is very important to avoid the system complexity and large processing time. This paper proposed a modified feature selection method for early breast cancer detection. Ultra-wideband (UWB) signals are transmitted and received using a pair of antenna. 1632 features are extracted from the received UWB signal and four statistical features (mean, median, maximum and minimum numbers) are selected from the extracted 1632 features. These features are fed into feed forward backpropagation neural network for breast cancer detection. The proposed features selection method is able to use to detect the breast cancer in terms of existence, location and size with average accuracy of 86.28%. The proposed feature selection method is able to increase the system performance.

Keywords: Feature selection, feed forward backpropagation neural network, signal processing

1 Introduction

Artificial neural network (ANN) -based processing of UWB signals are found in breast cancer applications. The main aims of the ANN are to simplify complexity by representing them in a convenient form and reduce the computational time. The dimensionality of the data is linearly proportional to the amount the of the data sample. Processing a large number of features increases computational cost and time. In addition, large number of features contains irrelevant and noise features. Noise should be avoided as much as possible to reduce the error in measurements because ANN performance is affected due to noisy features. To overcome those problems, it is important to find a suitable method to reduce the number of features (Jeyachidra, 2014). Two methods are usually used i.e.: feature selection and feature extraction. Feature selection method is a process of choose/select some features from the original features and the remaining the leave out. Feature extracting is a process of extracting features to produce a new set of features (Xue, 2016). Most of researchers undergo these process to increase the performance of the classification in order to make better the system more efficient and effective.

Discrete Wavelet Transform (DWT) is applied in order to extract the features and about 15 components are extracted from the original features by performing statistical analysis

ICASI 2019, July 18-19, Banda Aceh, Indonesia

DOI 10.4108/eai.18-7-2019.2288670

[Jones, 2013]. Hybrid feature selection based on Maximum Minimum Backward Selection (MMBS) is proposed and tested using support vector machine (SVM) and only 24 features are selected from 112 features [Liu, 2016]. The simple feature extraction method is done based on Chi-square rank correlation factorization by calculating the feature weight (Li, 2017). A new feature extraction method called Distinguishability based Weighted Feature Selection using Column Wise K-neighbourhood (DWFS-CKN) is proposed where feature weight is calculated based on classifiable nature. The highest feature weight is the most important features and vice versa. The features are listed based on the feature weight [Jeyachidra, 2014]. Principal component analysis (PCA) can be used to extract the features (Wimmer, 2017; Omucheni, 2014). In order to extract some features from the original data set, PCA equation is applied, standardized data set equation. From 4096 features, using PCA, only 50 features are extracted (Santorelli, 2014). On the other hand, PCA is proposed and investigated using fuzzy k-nearest neighbour (10-fold) and SVM. It is done by seeking the largest variance among the original feature (Chen, 2013). Alshehri (2011) has used PCA feature extraction method to extract important features. The original data with a length of 4500 to 7200 features is extracted until 50-300 features (Alshehri, 2011).

Four features are selected from 1632 features. Four features which are maximum, minimum, average and standard deviation values. 98% of original value has been deducted before training. This deduction has increased the performance of the system with 99% accuracy (Reza, 2015).Rough set based feature selection is basically measured the degree of features dependency and their significance (Qamar, 2013). SVM is used to test the proposed feature selection method and proved it can get along with this classifier (Wang, 2009). Swarm optimization is used to investigate the proposed method (Fan, 2012). Processing time can be reduced by using this method and improve the accuracy. Most of the researchers are able to eliminate all unwanted and noisy data and increase the data quality and system efficiency. In this paper, a modified feature selection method is proposed and the selected features from the method are fed into the developed ANN to identify the performance of breast cancer detection in terms of existence, location and size. The main contribution of this paper is to select four statistical features from the original features and verify the performance in terms of system complexity and efficiency. This paper is organized with methodology which describes about the breast model and experimental set-up, results and discussion which demonstrates the obtained results and finally the conclusion which summarized the whole work.

2 Methodology

The developed system is consisted of hardware and software. Hardware includes UWB transceiver and a pair of directional antenna. The software includes feed-forward back propagation artificial neural network (FFBPNN) module to detect the tumor existence, size and location along with soft interface between software and hardware. Heterogeneous breast phantom is developed based on the dielectric properties (permittivity and conductivity) of the real breast as described in [Alshehri, 2011]. Low cost material, wheat flour, water and petroleum jelly are used based on breast phantom dielectric properties. Tumors are placed in breast phantom with various size and location. The antenna [Reza, 2015] is placed diagonally opposite side of breast phantom as shown in Figure 1. Forward scattering technique is used to receive the forward scattered signals. This received signal contains the signature of the tumor information. The received analogue signal is converted into digital using Discrete Cosine Transform (DCT). There are 136 data samples are collected. Each data sample consists of 1632 features.

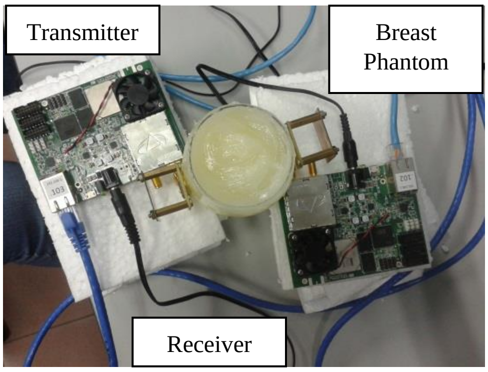


Figure 1: Experimental Set-up

2.1 Feature Selection

As mentioned, each data sample has 1632 features that encode the existence, location and size of a breast tumor/cancer which is large data set. The features are reduced by using proposed feature selection method based on (Reza, 2015). Such 1632 features are reduced into 4 features which are selected based on statistical analysis. Four features are mean, median, maximum number and minimum number. Maximum and minimum numbers normally are considered as the best feature compared to other features (Vanaja, 2014). Approximately 96% features is reduced from the original features. The median is the middle number of a group numbers. Mean, μ is expressed as in Equation 1.

$$\mu = \frac{1}{N}\sum_{n=0}^{N-1}[X_n]$$

$$= \frac{x_1 + x_2 + \cdots + x_n}{N} \qquad (1)$$

Where *N* is the total number of data points and X_n is the each feature.

2.2 Artificial Neural Network

The selected features are fed into FFBPNN module to train, validate and test. The collected 136 data samples are divided into two groups.
Group (1): 125 data samples for training, validating and testing.
Group (2): 11 data samples for real time testing.

FFBPNN module is developed by calling newff command in Matlab. Data samples (Group 1) are fed and train into developed ANN module until the result is satisfied. The number of hidden neurons and hidden layers are changed based on the data sample to optimize the training session. However, the ANN module should not be overtrained in order to avoid overfitting. Once the trained session is optimized, data sample of group (2) is fed into trained ANN module to evaluate the performance efficiency of the system.

3 Results And Discussion

The trained ANN module is tested using the untrained data sample as stated in Group (2). The detection efficiency on tumor existence, location and size are 100%, 84.48% and 85.86% and 100%, 80.81% and 88.97% for ANN module without feature selection and ANN module with feature selection as shown in Table 1 and Table 2 respectively. Table 3 shows the summarization of the Table 1 and Table 2 in terms of efficiency for existence, location and size respectively.

Table 1. Detection Efficiency ANN Without Feature Selection

Actual Target (mm)				**ANN output (mm)**				**Detection Efficiency (%)**			
x	y	z	size	x	y	Z	size	x	y	z	size
-1	-1	-1	-1	-1.00	-0.75	-0.99	-1.00	100	100	100	100
32.5	20	50	2	32.50	20.33	49.80	4.20	100	98.35	99.6	47.62
32.5	62.5	40	2	27.39	62.01	40.00	2.00	84.28	99.22	100	100
32.5	2.5	30	3	31.57	42.97	49.90	3.94	97.14	5.82	60.12	76.14
62.5	32.5	30	3	37.60	33.93	29.57	4.63	60.16	95.79	98.57	64.79
2.5	32.5	40	4	21.27	34.26	29.57	4.60	11.75	94.86	73.93	86.96
32.5	2.5	40	4	26.10	17.11	42.18	3.80	80.31	14.61	94.83	95.00
32.5	32.5	40	5	63.91	68.39	43.91	5.52	50.85	47.52	91.1	90.58
62.5	32.5	30	5	62.5	30.95	29.98	5.77	100	95.23	99.93	86.66
32.5	50	30	6	42.04	48.07	30.01	5.80	77.31	96.14	99.97	96.67
50	32.5	50	6	51.00	17.80	50.00	6.00	98.04	54.77	100	100

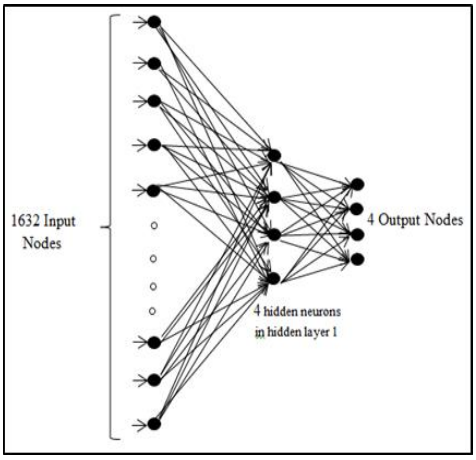
Table 2. Detection Efficiency ANN With Feature Selection

Actual Target (mm)				**ANN output (mm)**				**Detection Efficiency (%)**			
x	y	z	size	x	y	Z	size	x	y	z	size
-1	-1	-1	-1	-1.22	-0.98	-1.09	-1.12	100	100	100	100
32.5	20	50	2	25.45	16.78	42.64	1.59	72.29	80.81	82.74	74.21
32.5	62.5	40	2	25.98	56.90	36.34	1.90	74.90	90.16	89.93	94.74
32.5	2.5	30	3	25.39	10.90	25.78	3.67	71.99	22.94	83.63	77.67
62.5	32.5	30	3	69.06	35.67	35.00	3.20	89.50	90.25	83.33	93.33
2.5	32.5	40	4	3.45	40.87	45.21	4.39	72.47	74.25	86.98	90.25
32.5	2.5	40	4	45.67	15.66	35.67	3.90	59.48	15.96	87.86	97.44
32.5	32.5	40	5	37.78	36.00	41.90	4.56	83.75	89.23	95.25	90.35
62.5	32.5	30	5	53.98	45.87	30.98	5.40	84.22	58.86	96.73	92.00
32.5	50	30	6	36.89	59.98	32.70	7.68	86.49	80.04	91.00	72.00
50	32.5	50	6	40.56	32.80	48.00	6.20	76.73	99.08	95.83	96.67

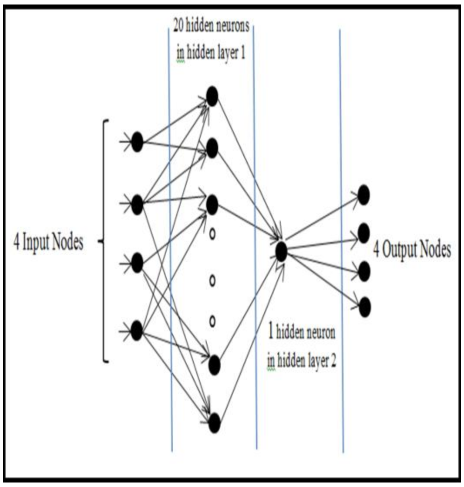
Table 3. Average Detection Efficiency ANN Without/With Feature Selection

Breast Phantom	**Average Detection Efficiency (%)**				
	Existence	**Location**			**Size**
		X	y	z	
Without Feature Selection	100	78.17	72.94	92.55	85.86
With Feature Selection	100	79.26	72.87	90.29	88.97

The overall performance efficiency increases from 85.43% to 86.28% by using the modified feature selection method. The feature selection method is to reduce the number of features of each data samples. Thus, number of hidden neurons is increased in order to get better detection efficiency. At the same time the architecture is simpler than the developed FFBPNN without using feature selection as shown in Figure 2. Figure 2 shows the proposed ANN structure. Figure 2(a) shown ANN architecture consists of 1632 inputs, 4 hidden neurons and 4 outputs while Figure 2(b) shows ANN architecture consists of 20 hidden neurons in first hidden layer and 1 hidden neuron in second hidden layer. The ANN architecture in Figure 2(a) is more complex and leads to consume more time to process as compared to ANN architecture of proposed feature selection as in Figure 2(b).



(a) Without using Feature Extraction



(b) With using Feature Extraction

Fig.2. Proposed FFBPNN Architecture

The feature selection method is proposed by Reza (2015) is validated using own data samples and compared to the modified feature selection method. The comparison is done using “size” data samples (heterogeneous breast phantom) only is as shown in Table 4. The testing efficiency was 88.97% using the modified feature selection method which better compared to the developed in Reza (2015). Reza (2015) is able to achieve 99% of training efficiency as stated in the paper which lower than proposed feature selection method (99.67%).

Table 4. Average Detection Efficiency for Reza (2015) Feature Extraction versus Modified Feature Extraction

	Average Detection Efficiency (%)			
	Mean, Standard Deviation, Maximum number, minimum number (Reza, 2015)		Mean, Median, Maximum number, minimum number	
Using Own Data samples	Training	96.76%	Training	99.97%
	Testing	63.44%	Testing	88.97%
	Overall	80.10%	Overall	94.30%
Based on Reza (2015)	Training	99.00%		

4 Conclusion

In this paper, a modified feature selection method is proposed in order to reduce the number of irrelevant and noise features. Huge features in each data sample conclude complex network architecture and increase processing time. Not all features are relevant and by using all features, the efficiency of the system can be decreased. These problems are solved by using the proposed feature selection method where it is a very helpful for breast cancer detection and enhances the system performance.

Acknowledgements

This work is supported by Ministry of Higher Education, Malaysia, Grant FRGS – 9001-00569.

References

[1] Jeyachidra, P. and Punithavalli, M., 2014. Distinguishability based weighted feature selection using column wise k neighbourhood for the classification of gene microarray dataset. *J. Applied Sciences*, *11*, pp.1-7

[2] Xue, B., Zhang, M., Browne, W.N. and Yao, X., 2016. A survey on evolutionary computation approaches to feature selection. *IEEE Transactions on Evolutionary Computation*, *20*(4), pp.606-626.

[3] Jones, M., Byrne, D., McGinley, B., Morgan, F., Glavin, M., Jones, E., O’Halloran, M. and Conceição, R.C., 2013. Classification and monitoring of early stage breast cancer using ultra wide band radar. In *The Eighth International Conference on Systems (ICONS), International Academy, Research, and Industry Association (IARIA)* (pp. 46-51).

[4] Liu, H., Jiang, H. and Zheng, R., 2016. The Hybrid Feature Selection Algorithm Based on Maximum Minimum Backward Selection Search Strategy for Liver Tissue Pathological Image Classification. *Computational and mathematical methods in medicine*, *2016*.

[5] Li, Q., Hao, C. and Xu, Z., 2017. Diagnosis of breast cancer tissues using 785 nm miniature raman spectrometer and pattern regression. *Sensors*, *17*(3), p.627.

[6] Wimmer, H. and Powell, L., 2016. Principle component analysis for feature reduction and data preprocessing in data science. In *Proceedings of the Conference on Information Systems Applied Research ISSN* (Vol. 2167, p. 1508).

[7] Omucheni, D.L., Kaduki, K.A., Bulimo, W.D. and Angeyo, H.K., 2014. Application of principal component analysis to multispectral-multimodal optical image analysis for malaria diagnostics. *Malaria journal*, *13*(1), p.485.

[8] Santorelli, A., Porter, E., Kirshin, E., Liu, Y.J. and Popovic, M., 2014. Investigation of classifiers for tumor detection with an experimental time-domain breast screening system. *Progress In Electromagnetics Research*, *144*, pp.45-57.

[9] Chen, H.L., Huang, C.C., Yu, X.G., Xu, X., Sun, X., Wang, G. and Wang, S.J., 2013. An efficient diagnosis system for detection of Parkinson's disease using fuzzy k-nearest neighbor approach. *Expert systems with applications*, *40*(1), pp.263-271.

[10] AlShehri, S.A., Khatun, S., Jantan, A.B., Raja Abdullah, R.S.A., Mahmud, R. and Awang, Z., 2011. Experimental breast tumor detection using NN-based UWB imaging. *Progress In Electromagnetics Research*, *111*, pp.447-465.

[11] Reza, K.J., Khatun, S., Jamlos, M., Fakir, M. and Morshed, M.N., 2015. Performance enhancement of ultra-wideband breast cancer imaging system: proficient feature extraction and biomedical antenna approach. *Journal of Medical Imaging and Health Informatics*, *5*(6), pp.1246-1250.

[12] Qamar, U., 2013, October. A rough-set feature selection model for classification and knowledge discovery. In *Systems, Man, and Cybernetics (SMC), 2013 IEEE International Conference on* (pp. 788-793). IEEE.

[13] Wang, Y. and Ma, L., 2009, January. Feature selection for medical dataset using rough set theory. In *Proceedings of the 3rd WSEAS international conference on Computer engineering and applications* (pp. 68-72). World Scientific and Engineering Academy and Society (WSEAS).

[14] Fan, H. and Zhong, Y., 2012. A rough set approach to feature selection based on wasp swarm optimization. *Journal of Computational Information Systems*, *8*(3), pp.1037-1045.

[15] Vanaja, S. and Kumar, K.R., 2014. Analysis of feature selection algorithms on classification: a survey. *International Journal of Computer Applications*, *96*(17).