Interpreting Artificial Neural Networks for Microwave Detection of Breast Cancer

Douglas A. Woten, John Lusth, and Magda El-Shenawee

***Abstract*—This letter demonstrates the applicability of artificial neural networks (ANNs) to breast cancer detection. A simplified model of the breast containing a tumor was used to determine the scattering of electromagnetic waves in the microwave band. This data was used to train an ANN which was then tested on new data to predict the presence or absence of a tumor. Synthetic variation (noise) is added to the data to realistically model the breast.**

***Index Terms*—Artificial neural networks (ANNs), breast cancer detection, wave scattering.**

I. Introduction

THE American Cancer Society estimates that over 170 000 women will be diagnosed with invasive breast cancer in 2007 in the United States alone. The survival rate of these women is inversely proportional to the stage of the cancer at the time of detection. This makes detection in the early stages when treatment is effective, vital. The current golden standard in detection methods is X-ray mammography. Unfortunately mammography screenings yield false-negative rates as high as 34% [1]. In response, microwave imaging has emerged as a promising solution to overcoming some of these problems [2].

At microwave frequencies the contrast between the electrical properties of the tumor and those of the healthy breast tissue is significant. This contrast allows for the detection of smaller tumors than traditional methods with high accuracy [2], [3]. Typical microwave systems are designed for detection or imaging, each having their own requirements. Detection requires determining whether an abnormality is present, whereas imaging requires a high degree of spatial resolution to reconstruct the object [4].

Imaging algorithms depend on an iterative process that requires considerable computer resources and time [4], [5]. This letter demonstrates the feasibility of incorporating an artificial neural network (ANN) as a preprocessor to these time intensive imaging algorithms. ANNs have been proven feasible in interpreting digital mammography scans but, to the best of our knowledge, ANNs have not been explored for microwave detection [6].

Manuscript received May 31, 2007; revised July 31, 2007. This work was supported in part by the National Science Foundation GK-12 Program, NSF Award Number ECS-0524042, the Arkansas Biosciences Institute, and the Women's Giving Circle at the University of Arkansas.

D. A. Woten is with the Microelectronics–Photonics Program, University of Arkansas, Fayetteville, AR 72701 USA (e-mail: dwoten@uark.edu).

J. Lusth and M. El-Shenawee are with the Computer Science Computer Engineering Department, Electrical Engineering Department, University of Arkansas, Fayetteville, AR 72701 USA.

Color versions of one or more of the figures in this letter are available online at http://ieeexplore.ieee.org.

Digital Object Identifier 10.1109/LMWC.2007.910466

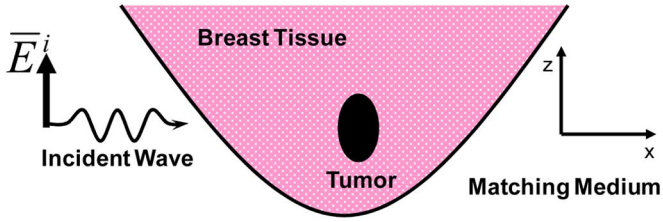

Fig. 1. Model of the breast and incident plane wave at 6 GHz with an ellipsoidal tumor (0.75 cm, 0.5 cm, 1.0 cm).

TABLE I
RANGE OF DIELECTRIC VALUES

	Breast Tissue	Tumor
Lower	9-j1	40-j10
Upper	25-j6	60-j30

II. Methodology

The breast is modeled as a homogeneous region containing a tumor. An incident plane wave at 6 GHz with vertical polarization is used (Fig. 1). The scattered fields are calculated at a point receiver. For this work the effect of the skin is ignored in our model. This simplified breast model is used to demonstrate the proof of concept of using ANNs in this application. Studies have been conducted on the effect of the skin and the error introduced by ignoring its thickness [7]. Approximately 3% error is introduced when using the scattered fields in backscatter direction upon ignoring the skin thickness. Therefore, the work presented here is restricted to the backscattered fields for both training and testing the neural network.

The dielectric properties of the tumor and surrounding breast tissue were randomly selected in the ranges shown in Table I [2], [3]. A medium, perfectly matched to the breast tissue, is assumed to surround the breast.

A computer code based on the Method of Moments (MoM) and the Ansoft package (HFSS) is used to calculate the scattered electromagnetic fields (E_x, E_y and E_z) from an ellipsoidal tumor [4]. One hundred random locations of the tumor are generated along with one hundred random sets of electrical properties of the tumor and background. The scattered fields are calculated for each of these 10 000 cases. In addition, 4 000 cases were simulated with no tumor present but with random electrical properties of the background. This leads to a total of 14 000 cases which are equally divided into training and testing sets.

Fig. 2 shows the range selected for the random tumor location, where the point receiver is located at least 4 cm from the tumor. Synthetic variations (noise) were then added to each dataset to represent white noise introduced in a laboratory environment.

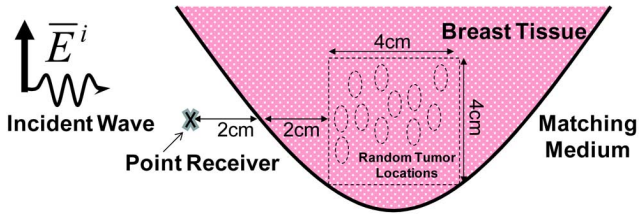

Fig. 2. Random tumor locations.

The noise added is based on a Gaussian probability density function with a mean of zero and varying standard deviations, σ_n ($\sigma_n = .001, .01, .1, .2$, and $.3$).

Multiple ANNs were prepared using each noise variation of the training data (from no noise through $\sigma_n = 0.3$) and these were tested against each variation of the testing set.

An ANN is a processing technique, loosely based on the human brain, in which interconnected layers process information simultaneously and adapt over the course of multiple runs [8]. A typical neural network consists of three layers 1) the input layer, 2) the hidden layer, and 3) the output layer. Each layer consists of a series of nodes which are connected to nodes in the other layers. For this study the input layer consists of six nodes corresponding to the real and imaginary part of E_x, E_y, and E_z. The hidden layer is sought to be as simple as possible while still allowing for an acceptable detection rate. For this study seven nodes were utilized in the hidden layer. The output layer consists of a single node.

For this project, the package NevProp 3—Nevada backPropagation Version 3 is used (open source). A fully interconnected network with seven nodes in the hidden layer is implemented using this package.

At each node in the hidden layer an activation function is called that manipulates the incoming data from the input layer. Although this function can take many forms depending on the output desired, for this study a saturation function is used which constrains the output to the range of 0–1. This function is given as:

$$f(x) = 1/(1 + e^{-\beta x}) \tag{1}$$

where β is a constant determining the slope of the function and is determined during the iterative error minimization process by the ANN.

III. Results

After running each case contained within the testing data set (7 000 cases), each case produces a neural network output between 0 and 1. The next step is determining the cutoff value using Receiver Operating Characteristic (ROC) curves [9], [10]. This cutoff value determines what constitutes the presence or absence of an abnormality. Two schemes were implemented and the results were compared. The first scheme utilizes a single cutoff yielding two possible predictions, 1) a tumor is present and 2) no tumor is present. The second scheme utilizes a double cutoff separating the output data into three regions, 1) tumor is present, 2) undecided, and 3) no tumor is present.

An ideal cutoff would be the one that corresponds to the highest probability of detection with the smallest tradeoff in false positives. To find this optimum cutoff value(s) Youden's Index, a method to minimize the false positive and false negative rates, is used [11]. The typical Youden's Index (Y) is given by

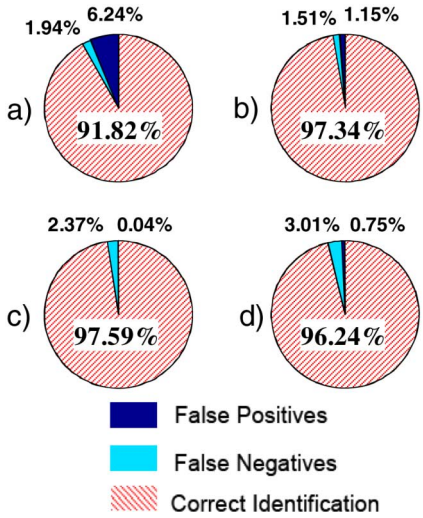

Fig. 3. Results using a single cutoff of 0.5 with a testing set of noise $\sigma_n = 0.1$ and a training set of (a) no noise (b) $\sigma_n = 0.1$ (c) $\sigma_n = 0.2$ and (d) $\sigma_n = 0.3$.

$$Y = 1 - [(\mathrm{FNR}) + (\mathrm{FPR})] \tag{2}$$

where FNR is the false negative rate and FPR represents the false positive rate.

By calculating Youden's index for each cutoff and recording the cutoff with the highest index for each test, the optimum cutoff can be found. The mode value of the highest indices, Y, is used as the optimum single cutoff, which is 0.5 in this work. Using this cutoff the number of false positives, false negatives and correct identifications is recorded for a network training of noise level $\sigma_n = 0.1$ (Fig. 3). The comparison of the ANN prediction versus the known presence/absence of a tumor in the synthetic data is used to determine the percentage accuracy of the network.

For breast cancer detection it is arguable that the advent of a false negative is more undesirable than a false positive. Although a false positive could cause undue stress on the patient, a false negative might mean a tumor goes untreated and progresses to later stages. Youden's index, by formulation, is unbiased towards the FNR or the FPR since each one is given an equivalent weight of 1. To find a double cutoff a new form of Youden's Index must be implemented. It has been shown that weights can be added to Youden's Index to allow selectivity in the results based on the severity of the FNR and FPR given by [11]

$$Y = 1 - (w_{\mathrm{sens}}\mathrm{FNR} + w_{\mathrm{spec}}\mathrm{FPR}) \tag{3}$$

where w_{sens} is the weight associated with the sensitivity and w_{spec} is the weight associated with the specificity.

The choices for these weights in (3) were chosen arbitrarily as there is no quantified value as to how much worse a false negative is to a false positive. In an attempt to minimize both the FNR and the FPR, weights were chosen as 2 and 1 and the corresponding cutoff was determined. The weights were then

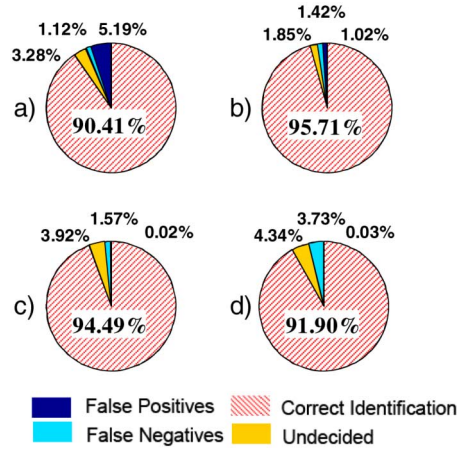

Fig. 4. Results using a double cutoff of 0.4 and 0.7 with a testing set of noise $\sigma_n = 0.1$ and a training set of (a) no noise, (b) $\sigma_n = 0.1$, (c) $\sigma_n = 0.2$, and (d) $\sigma_n = 0.3$.

changed to 1 and 2 to obtain the second cutoff. The calculated cutoffs were 0.4 and 0.7, respectively, for this work.

Using the cutoff values specified above, the number of false positives, false negatives, and undecided results were plotted in Fig. 4.

Comparing Figs. 3 and 4, it is clear that the introduction of undecided results had an impact on the overall accuracy of the network, i.e., the number of correct identifications. When using a single cutoff the number of correct identifications is higher than the double cutoff case, but with a larger number of false positives and false negatives. The introduction of undecided results lowered the overall accuracy while also lowering the number of false negative and false positives. Additionally, the number of undecided results tends to increase with the increase of noise in the training set.

It is noticed that introducing some noise into the training data improves the network performance. This observation agrees with Wang *et al.*, [8] where injection of noise into the input signal enhanced network robustness and training efficacy. If too much noise is added, however, the network performance starts to degrade as expected. In Figs. 3 and 4 this is observed at noise levels of $\sigma_n = 0.3$.

The large number of simulated cases (14 000) used here allowed the network to perform more accurately than previous work using only 200 cases and a single tumor location [10]. This implies that the benefit of using a larger training set overcomes the additional uncertainty added by varying the tumor location.

IV. Conclusion

The inherent adaptability of ANNs enables the network to be reconfigured very easily, while still providing results within milliseconds. Additionally, by setting the cutoffs arbitrarily low, the number of false negatives can be reduced to zero, while increasing the number of false positives and undecided results. This allows the network to act as a preprocessor to the more time intensive imaging algorithms without sacrificing the integrity of the results. Ideally, the false positive and undecided results will be passed to the imaging algorithm, while the cases with no tumor will not.

We observed that the worst case results occurred when training on a high level of noise and testing vs. the other highest noise (data not shown here). This yielded an overall miss rate of 32%, which is comparable to the worst case scenario of mammography 34% [1]. This is considered a reasonable performance for a preprocessor algorithm.

A smaller number of cases (200 cases) utilizing a three-region model of the breast were simulated and compared to the two-region model [9]. The network performance was not significantly affected by the inclusion of an additional region, although the required computational time drastically increased. For the two-region model, each case required approximately 2–3 min on a 64-b SUN server (2.4 GHz Opteron processor) compared to 2–3 h for a three-region model on the same platform.

References

[1] P. T. Hunynh, A. M. Jarolimek, and S. Daye, "The false-negative mammogram," *Radiograph*, vol. 18, no. 5, pp. 1137–1154, 1998.

[2] P. M. Meaney, M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE Trans. Microw. Theory Tech.*, vol. 48, no. 11, pp. 1841–1853, Nov. 2000.

[3] E. J. Bond, X. Li, S. C. Hagness, and B. D. Van Veen, "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propag.*, vol. 51, no. 8, pp. 1690–705, Aug. 2003.

[4] M. El-Shenawee and E. Miller, "Spherical harmonics microwave algorithm for shape and location reconstruction of breast cancer tumor," *IEEE Trans. Med. Imag.*, vol. 25, no. 10, pp. 1258–1271, Oct. 2006.

[5] P. Rashidi, M. El-Shenawee, and D. Macias, "Microwave imaging of malignant tumors employing an enhanced evolution strategy," in *Proc. IEEE Int. Symp. Antennas Propag.*, Washington, D.C., Jul. 2005, pp. 807–810.

[6] D. Voth, "Using AI to detect breast cancer," *IEEE Intell. Syst.*, vol. 20, no. 1, pp. 5–7, Jan./Feb. 2005.

[7] S. Pandalraju, "A Hybrid Algorithm Based on Mie Theory and Evolution Strategy for Breast Cancer Imaging," M.S. thesis, Univ. Arkansas, Fayetteville, 2006.

[8] C. Wang and J. Principe, "Training neural networks with additive noise in the desired signal," in *Proc. IEEE Int. Conf. Neural Networks*, Anchorage, AK, Jul. 1998, pp. 1084–9.

[9] D. Woten, "Artificial Neural Networks for Breast Cancer Detection using Micro Antennas," M.S. thesis, Univ. Arkansas, Fayetteville, 2006.

[10] D. Woten, M. El-Shenawee, and J. Lusth, "Interpreting artificial neural network output for the microwave detection of breast cancer," in *Proc. Appl. Comput. Electromagn. Soc.*, 2007, pp. 91–96.

[11] A. Kristopher L. and A. Pretty Ian, "Results of the 4th ABFO bitemark workshop," in *Proc. Forensic Sci. Int.*, 1999, vol. 124, pp. 104–11.