

Noninvasive Breast Tumor Localization Based on Ultrawideband Microwave Backscatter

Ahmad AbdulSadda

Department of Applied Science
University of Arkansas at Little Rock
Email: atabdulsadda@ualr.edu

Nidhal Bouaynaya

Department of Systems Engineering
University of Arkansas at Little Rock
Email: nxbouaynaya@ualr.edu

Kamran Iqbal

Department of Systems Engineering
University of Arkansas at Little Rock
Email: kxiqbal@ualr.edu

Abstract—The key to surgical planning for breast conservation is tumor localization. An accurate localization of the breast tumor is essential to guide the surgeon to the lesion, and ensure its correct and adequate removal with satisfactory excision margins. Current breast tumor localization techniques are invasive and often result in a cosmetic disfigurement. In this paper, we use the ultrawide band radar-based microwave breast imaging technique to non-invasively localize (impalpable) tumors in the breast. We consider four clinically important lesion features: location, size, depth and spatial orientation within the breast. A comparison of the energy of the received signal from healthy and cancerous breasts exhibits some significant differences in some frequency bands. We, therefore, use the energy spectrum of the receiving antenna signal decomposed by wavelet transform as the input to a genetic neural network (GNN) classifier. Furthermore, for improved efficiency, we optimize the structure of the GNN for optimum initial weights and number of hidden nodes. We use CST Microwave Studio to simulate benign and malignant breast conditions, and generate a data set of 1024 cancer cases with various tumor location, size, depth and spatial orientation within the breast. Our results show that the proposed algorithm gives accurate localization of the breast lesion, and possesses a high sensitivity to small tumor sizes. Additionally, it can accurately detect and classify multiple tumors.

Index Terms—Breast cancer localization, radar-based microwave breast imaging, wavelet decomposition, genetic algorithm, neural networks.

I. INTRODUCTION

A complete understanding of the distribution of tumor before therapy is a key factor in managing any case of advanced breast cancer. This is particularly true if preoperative chemotherapy is used in an attempt to allow breast preservation. The key to surgical planning for breast conservation is tumor localization. If the tumor shrinks to 50% or more after chemotherapy to less than 2 cm, then the surgeon would need an accurate guide to the lesion in order to remove it from the breast with satisfactory excision margins.

Although a number of techniques have been proposed to localize impalpable lesions prior to surgery, needle or hook-wire localization is still the most common used technique to help guide definitive resection of a cancer [1]. This procedure presents a number of disadvantages to the patient including discomfort and vasovagal syncope [2]. There are also clinical complications, which include pneumothoraces and wire migration, displacement or transection [2]. Radioguided occult lesion localization (ROLL) was proposed to reduce the

excision volume, and enable better lesion centering within the specimen. However, the procedure is still invasive, and there is the potential for radiation exposure to the patient and surgeon. Cash et al. [2] proposed a noninvasive method of breast tumor localization based on the coregistration of three-dimensional (3D) ultrasonographic (US) data with surface contour laser data. However, as the authors point out, this technique fails in general with large breast patients because the image definition diminishes with depth, thus resulting in the need for compression. Compression will distort the breast contour and result in a poor coregistration of US data with laser data.

In this paper, we propose a new noninvasive tumor localization method based on the ultrawide band microwave backscatter technique. We use the wavelet decomposition to accentuate the backscattered signal's signature in the frequency domain, and an optimized genetic neural network for accurate tumor localization. Microwave imaging is an alternative imaging method for detecting breast cancer [3]. It is based on the electrical properties of the breast tissues. Cancerous tissues exhibit electrical characteristics that show significant and consistent contrast with healthy breast tissues [4]. Ultrawideband radar-based microwave imaging uses an antenna to irradiate the breast with low-power ultra wideband pulses of microwave energy. The scattered energy is received by a receiving antenna. We propose to localize breast lesions in terms of their (i) location, (ii) size, (iii) depth, and (iv) spatial orientation within the breast. Spatial orientation of the tumor with respect to the nipple may illustrate ductal extension of tumor, and this factor may have an influence on the shape and volume of the excised specimen. We found that there are significant differences in specific frequency bands between healthy and tumor tissues. We foster this difference in the frequency domain for tumor localization by using the wavelet transform of the received signal's energy to obtain a discriminative feature vector for the GNN. An important byproduct of the wavelet decomposition is the compression of the total number of input data to few wavelet coefficients. In order to avoid the common neural network problems of local minima, convergence, and overfitting of the data, we use the genetic algorithm to select the optimum number of hidden nodes and the optimum initial weights for the network.

II. THE BREAST MODEL

We model the breast as a homogenous region, which can possibly contain one or multiple lesions. The transmitting antenna operates at a frequency of 6 GHz [5]. The distance between the breast and the antenna array is set to 2 cm with total feeding power of 300mW. The breast cancer model with the antenna position is described in Fig. 1. The received signal carries information about the tumor status inside the breast, and is related, through the Maxwell equations, to the dielectric permittivity of the breast [6].

In our breast model, we take into account the dependence of the relative dielectric permittivity of the breast, ϵ_r , on the frequency by using the following first-order Debye dispersion formula:

$$\epsilon_r = \epsilon_\infty + \frac{\epsilon_s - \epsilon_\infty}{1 + \omega^2 \tau^2}, \quad (1)$$

where ϵ_s and ϵ_∞ are the space permittivity, and the high frequency permittivity, respectively. ω denotes the angular frequency, and τ is the time constant. The values of the Debye parameters for normal and cancerous tissues have been determined experimentally in [6]. Figure 2 shows an example of the time domain measurement of the received signal for a healthy and cancerous breasts.

III. WAVELET-BASED FEATURE VECTOR

Even though the response signal of the breast to an ultra wideband microwave frequency can be easily measured using receiving antennas [5], the raw received signal cannot be used directly to fully identify the abnormal lesions within the breast due to the high amount of clutter from healthy tissues. An appropriate representative index vector has to be constructed from the received signal, and used for tumor characterization and classification. A comparison of the energy of the received signal from healthy and cancerous breasts exhibits significant differences in specific frequency bands (see Fig. 2). This is because tumors will suppress or enhance certain frequency components of the received signal, and consequently cause energy increase or decrease of the received signal. Therefore, the energy of the received signal contains information on abnormal lesions. Additionally, the energy variation of one or several frequency components of the signal can indicate a special characteristic of the lesion. In order to extract the tumor information from the breast response signal, we first decompose the signal into multiple sub-signals in various frequency bands using the wavelet decomposition as follows

$$S_{0,0}[n] = \sum_{j=1}^{2^{k-1}} S_{k,j}[n], \quad n = 0, 1, \dots, N-1, \quad (2)$$

where $S_{0,0}[n]$ denotes the sampled received signal of length N , and $S_{k,j}[n]$ is the sub-signal with orthogonal frequency band and k indicates the layer number of the tree structure of

the wavelet decomposition. The energy of the j^{th} order sub-signals can be expressed as

$$U_{k,j} = \sum_{n=1}^N |S_{k,j}[n]|^2. \quad (3)$$

Assuming that the energy of the j^{th} -order sub-signals of the healthy and cancerous tissues are $U_{k,j}^h$ and $U_{k,j}^c$ respectively, the wavelet index vector can be constructed as follows:

$$\begin{aligned} V_d &= [v_1, v_2, \dots, v_{2^{k-1}}]^T \\ &= \left[1 - \frac{U_{k,1}^c}{U_{k,1}^h}, 1 - \frac{U_{k,2}^c}{U_{k,2}^h}, \dots, 1 - \frac{U_{k,2^{k-1}}^c}{U_{k,2^{k-1}}^h} \right]^T \end{aligned} \quad (4)$$

The elements of different index vectors not only indicate the differences between healthy and cancerous tissues, but also suggest the differences between various lesion types. We will use the energy vector as the input feature vector to the genetic neural network classifier.

IV. OPTIMUM DESIGN OF THE NEURAL NETWORK USING THE GENETIC ALGORITHM

A. Genetic algorithm

The genetic algorithm (GA) is a search technique, based on an abstraction of biological evolution, whereby chromosomes are used to encode possible solutions to an optimization problem. In GA, a population of chromosomes, which are represented as binary vectors, evolves into a new population using the natural selection forces of crossover and mutation. Some chromosomes are selected and allowed to reproduce and the more ‘‘fit’’ chromosomes produce more offspring. The quality of a solution is represented by its fitness, which is the objective function to be optimized. In our case, the objective function is the mean-square error.

We use the genetic algorithm to determine the optimal initial weights and the optimum number of nodes in the hidden layer of the neural network. The optimal initial weights will guarantee a faster convergence to the global minimum of the objective function, and the optimum number of nodes in the hidden layer will avoid the bias-variance dilemma of neural networks. Given a data set, a small number of nodes in the hidden layer captures the data trend only in a small region of the pattern, whereas a large number of nodes in the hidden layer results in overfitting the data. In the tumor localization problem, it is important to obtain a high accuracy in order to design a computer-aided system, which will ultimately replace preoperative invasive techniques for tumor localization.

B. Genetic neural network

We use a three-layer artificial neural network (ANN) with: (1) an input layer that receives the index vector of the received signal, which has 32 nodes; (2) a hidden layer, which processes the data; and (3) an output layer that indicates the tumor status we are interested in classifying.

In this paper, we consider the following breast cancer statuses:

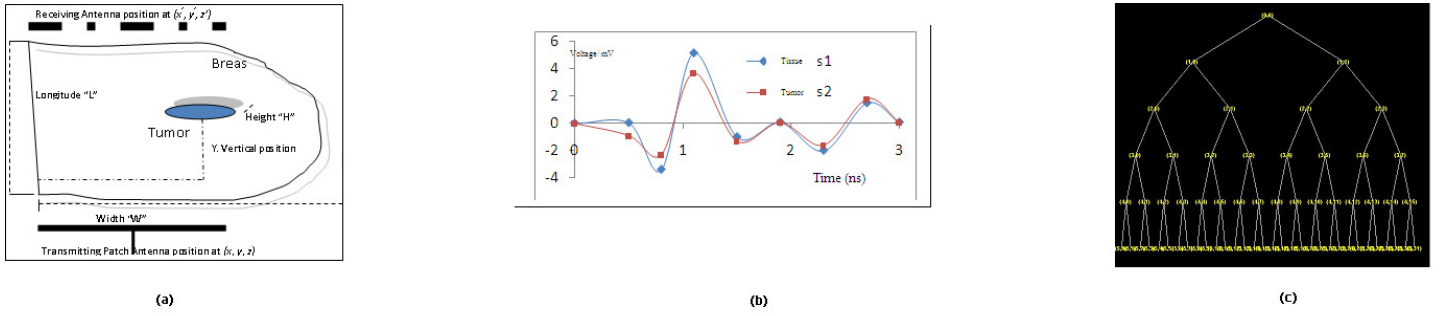


Fig. 1. (a) Breast model with transmitting and receiving antennas; (b) Time-domain measurements of the received signal for a healthy and cancerous breasts; (c) Wavelet tree decomposition of the cancerous signal in (b).

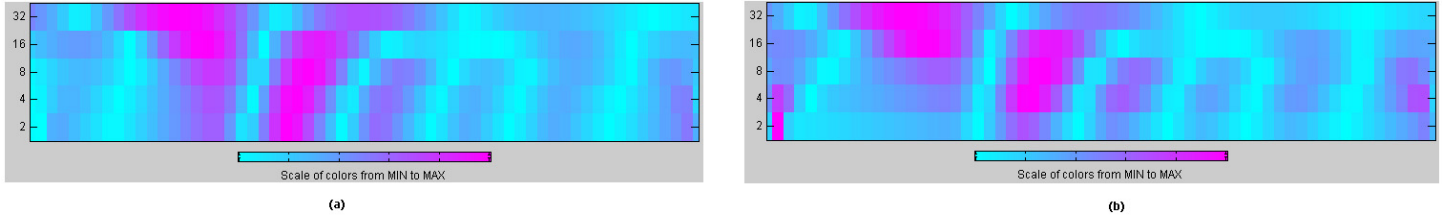


Fig. 2. Wavelet spectrogram of the (a) healthy and (b) cancerous signals shown in Fig. 1(b).

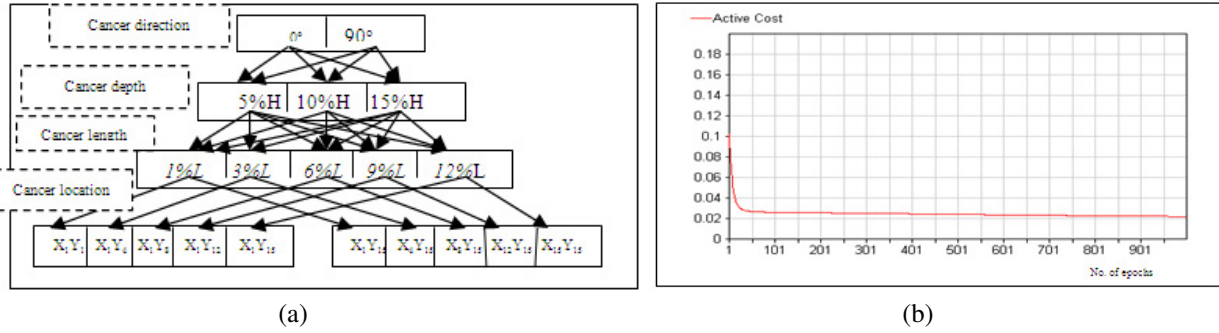


Fig. 3. (a) A representation of 60 statuses of tumor characteristics. (b) Mean square error (MSE) versus number of epochs.

- 1) Tumor spatial orientation of 0° and 90° corresponding to a longitudinal or a transverse lesion, and expressed using one bipolar nodal output of -1 or 1.
- 2) Three possible tumor depth: 5%H, 10%H, and 15%H denoted using the nodal outputs with two digits (1,-1), (-1,1), and (1,1), respectively.
- 3) 15 possible tumor cell lengths: 1%L, 2%L, ..., 15%L represented using the nodal outputs with four digits (1, -1, -1, -1), (-1, 1, -1, -1), ..., (1, 1, 1, 1), respectively.
- 4) 15 possible ranges of x -coordinates, with respect to the breast length L , and 15 possible ranges of y -coordinates, with respect to the breast width B , of the tumor location at 5%, 15%, 25%, 30%, 35%, 40%, 45%, 50%, 55%, 60%, 65%, 70%, 75%, 85%, and 95%. The x, y coordinates are described using two sets of nodal outputs with four digits (1, -1, -1, -1), (-1, 1, -1, -1), ..., (1, 1, 1, 1), respectively.

Figure 3(a) shows 60 examples of cancer statuses among the 32,767 possible cancer statuses that can be described by the NN. Observe that we need 15 nodes in the output layer of this NN. The healthy case (no tumor) will be represented by the value 0 at all 15 nodes in the output layer.

V. SIMULATION RESULTS

We use CST Microwave Studio to simulate the backscattered microwave signal from cancerous breasts. We model the breast as a homogenous cylindrical shape attached to a half-sphere, and containing a tumor, modeled as a small sphere inside the breast. In this reported set of simulations, we use a tumor of radius 8 mm. Observe that it is very difficult to observe such a small size tumor using the standard imaging methods like mammography and ultrasound imaging. The Debye parameters for healthy breast tissues are given by $\epsilon_s = 10$, $\epsilon_\infty = 7$, $\tau = 6.4$ ps, and for cancerous tissues are $\epsilon_s = 40$, $\epsilon_\infty = 5.573$, $\tau = 9.149$ ps at 6 GHz [6]. Therefore,

TABLE I
TARGET OUTPUTS AND REAL OUTPUTS OF THE GNN FOR FIVE SELECTED VERIFICATION SAMPLES.

Output bit	Sample case 1		Sample case 2		Sample case 3		Sample case 4		Sample case 5	
1	1	0.967	1	0.96	1	0.97	1	0.97	1	1.02
2	1	1.02	1	1.04	1	1.03	1	0.94	1	0.96
3	1	0.941	1	1.08	1	0.97	1	0.92	1	0.97
4	1	0.97	1	1.13	-1	-1.03	1	0.95	1	0.98
5	-1	-0.94	1	1.02	1	0.94	-1	-0.96	1	1.03
6	-1	-0.97	1	0.96	1	0.97	-1	-0.95	1	0.98
7	-1	-0.94	-1	-1.03	-1	-1.03	-1	-1.03	1	1.02
8	1	1.06	-1	-0.97	-1	-0.93	-1	-0.97	1	0.97
9	-1	-0.94	-1	-0.95	-1	-0.95	-1	-0.94	1	0.99
10	-1	-0.94	1	0.99	-1	-0.98	1	1.02	1	0.94
11	-1	-0.95	-1	-1.03	1	0.95	1	0.97	1	0.98
12	1	0.95	1	0.94	1	0.97	1	0.98	1	0.97
13	1	0.94	1	0.96	1	0.94	1	0.94	1	0.95
14	1	1.05	1	1.03	1	0.98	1	0.96	1	1.02
15	1	0.98	1	1.02	1	1.02	1	0.94	1	0.94

TABLE II
PERFORMANCE OF THE PROPOSED CLASSIFIER WITH AND WITHOUT NOISE

Noise level	Accuracy	False positive rate	False negative rate
0% noise	94%	2.8%	1.8 %
5% noise	91.6 %	4.6 %	3.7 %
10% noise	90.7 %	5.5 %	3.7 %

using Eq. (1), the relative dielectric permittivity of a healthy (resp., cancerous) breast is $\epsilon_r = 9.89$ (resp., $\epsilon_r = 38.20$).

Applying the genetic algorithm to find the optimal neural network parameters, we found that the optimal number of nodes in the hidden layer is 18. We train the NN on 1024 cases, and we test it on 216 cancer statuses, that are different from the learning statuses, and one healthy (no tumor) case. Figure 3(a) displays the training error, which decreases monotonically with the number of iteration and reaches the desired value of 0.02 after about 60 epochs. Table I shows the desired (or target) outputs and the obtained outputs of the genetic neural network for five selected samples. The proposed algorithm was able to discriminate between healthy and cancerous breasts with an accuracy of 94.9%, with a false positive rate of 2.77% and a false negative rate of 1.85%. The classification error of the cancer cases is evaluated at about 6%, thus leading to an accuracy of 94%. In real experiments, the acquired data is noisy and the noise level may reduce the accuracy of the classifier. In a second experiment, we introduce noise to the data and perform the same classification again. The results are summarized in Table II.

VI. CONCLUSION

In this paper, we showed that we can accurately localize breast tumors using the antenna spectrum of the microwave backscattered breast response, decomposed by wavelet transform and fed to an optimized genetic neural network. Our simulations show that the required number of inputs to the neural network is greatly reduced, thus reducing the time and effort needed for the genetic neural network training. The genetic algorithm gave us the optimal number of hidden nodes

with the best initial weights as well.

A computer-aided diagnosis system based on the proposed algorithm for tumor localization will save health care cost and patient discomfort by avoiding expensive (and sometimes not very safe) invasive techniques to localize the exact position of small tumors prior to surgery. Finally, it is important to mention that the proposed algorithm is not restricted to breast tumor localization, but can be applied to any tumor localization within the body as long as the tumor and healthy tissue properties exhibit significant differences in the wavelet or frequency domain.

REFERENCES

- [1] A. Haid, M. Knauer, S. Dunzinger, and et al. Z. Jasarevic, "Intra-operative sonography: a valuable aid during breast-conserving surgery for occult breast cancer," *Annals of surgical oncology*, vol. 14, no. 11, pp. 3090–30101, November 2007.
- [2] C. J. C. Cash, G. M. Treece, A. D. Purushotham, and et al. P. Britton, "Noninvasive method of preoperative localization with three-dimensional US and surface contour mapping," *Radiology*, vol. 245, no. 2, pp. 556–566, November 2007.
- [3] P M Meaney, K D Paulsen, A Hartov, and R K Crane, "An active microwave imaging system for reconstruction of 2-D electrical property distributions," *IEEE Transactions on Biomedical Engineering*, vol. 42, no. 10, pp. 1017–1026, October 1995.
- [4] E. Fear and M. Stuchly, "Microwave detection of breast cancer," *IEEE Transaction on Microwave theory and techniques*, vol. 48, no. 11, pp. 1854–1863, 2000.
- [5] S. Suh, *A comprehensive investigation of new planar wideband antennas*, Ph.D. thesis, Virginia Polytechnic Institute, Blacksburg, VA, USA, 2002.
- [6] D. Woten, J. Lusth, and M. Al-Shenawee, "Interpreting artificial network output for the microwave detection of breast cancer," *IEEE Microwave and Wireless Components Letter*, vol. 17, no. 12, pp. 825 – 827, December 2007.