

# A Breast Cancer Detection Approach Based on Radar Data Processing using Artificial Neural Network

Salvatore Caorsi<sup>1</sup>, Claudio Lenzi<sup>2</sup>

<sup>1,2</sup> Department of Electrical, Computer and Biomedical Engineering, University of Pavia, 27100, Pavia, Italy

**Abstract**—A new breast cancer detection approach is proposed as an accurate, non-invasive Yes-No diagnostic tool without the need for breast imaging. The approach is based on the processing of ultra wideband (UWB) mono-static radar signals backscattered around heterogeneous two-dimensional (2D) and three-dimensional (3D) breast models. Suitable data are extracted and input in an artificial neural network (ANN) able to detect the presence or absence of the tumor for each single radar trace. Then, a diagnostic criterion is applied, considering the collective ANN outputs. The best results were obtained for tumors positioned outside the fibro-glandular tissues. Using 2D breast models and an ideal skin artifact removal technique, tumors were detected with 80% accuracy for 2000 testing data values. When a realistic model-based skin artifact removal technique was applied, 74% accuracy was obtained. Using a realistic 3D breast model, this technique correctly detected tumors with diameters as small as 2 mm located at different distances from the chest. Moreover, for the analyzed cases, the application of the diagnostic criterion showed an accuracy of 100%. The ANN processing technique applied to radar systems realizes a simple, fast, and highly accurate breast cancer diagnostic criterion with low computational burden.

**Keywords**— Artificial neural network, breast cancer detection, inverse scattering.

## I. INTRODUCTION

Breast cancer is the most frequent type of cancer among women and it accounts for about one-third of all cancer diagnoses [1]. Its early detection is one of the most challenging and fundamental aspects in improving treatment outcomes and reducing the mortality rate [2].

At present, the standard technique for breast cancer detection is the use of an X-ray mammogram. However, this technique suffers of various problems and limitations. In addition to the pain and discomfort of compressing the breast, it is well known that exposure to X-rays subjects the women to serious health risks and can increase the chances of cancer development [3]. Moreover, the X-ray mammogram suffers from relatively high numbers of detection failures and shows sensitivity values that vary in the range of 66% and 96%; this wide variability can be attributed to several factors as reported in [4].

In order to improve performance, various combination modalities consisting of different diagnostic techniques, such as the X-ray mammogram, magnetic resonance imaging (MRI), ultrasound (US), and clinical examination, were studied in [5]. Using these combinations, some interesting results were obtained, in particular for sensitivity. Nevertheless, the values of overall accuracy, which accounts for the values of sensitivity and specificity together, vary in a range between 66.6% and 75.6% [5, 6].

In addition, in recent decades, the existing contrast between the electromagnetic properties of malignant and healthy breast tissues [7] has driven the development of microwave imaging techniques. Among these, different modalities including passive [8, 9], hybrid [10], and active approaches [7, 11, 12, 13, 14, 15, 16] have been studied and proposed. Among the active techniques, the two most promising alternatives are those based on microwave tomography and those based on ultra-wide band (UWB) radar imaging.

The microwave tomography approaches are intended to provide a quantitative estimation of the spatial profile of the breast's dielectric properties, and they are based on the solution of an inverse electromagnetic scattering problem [11, 12]. However, solving this problem usually involves long computation times and heavy computational burden, where these two parameters strongly depend on the searched resolution accuracy, the microwave frequencies used, the signal information, and the problem approximation applied (such as the Born and Rytov approximations).

The UWB radar imaging techniques are focused on achieving breast imaging to determine the presence and location of significant dielectric scatterers. These techniques are based on breast illumination using UWB pulses and on the processing of the backscattered signals using a time-domain image-formation algorithm (beamformer). Several algorithms have been proposed in the literature; they can be classified in two main categories: data dependent (DD) beamforming and data independent (DI) beamforming. Examples of DD algorithms are multistatic adaptive microwave imaging (MAMI) [17], multi-input multi-output (MIMO) [18], and time-reversal multiple signal classification (TR-MUSIC) [19, 20]. These techniques can provide high-accuracy resolution in the case where the array steering vector corresponding to the signal of interest is well known, but when dealing with realistic cases, it is difficult to determine this component. Otherwise, some promising DI beamforming techniques [21] include delay and sum (DAS), delay multiply and sum (DMAS), and improved delay and sum (IDAS). These algorithms, in order to compensate the signal attenuations and dispersions due to the propagation path inside dispersive tissues, use an assumed homogeneous dielectric breast model. As expected and according to [21], the resulting accuracies are high in the case of homogeneous breast models but they worsen with increasing breast heterogeneity [21].

Moreover, another important issue relating to microwave diagnostic techniques is that the total backscattered signals contain, in addition to the components due to the presence of

both the tumor and the internal tissues, the reflection backscattered from the air/skin interface (the so-called artifact component) [22]. This last is a predominant component because it is several orders of magnitude greater, and because it overlaps the internal tissues' reflections, it can mask the presence of tumors [22, 23, 24, 25]. On this issue, various methods have been proposed in the literature with the aim of reducing the artifact component before signal processing [22, 23, 24, 25, 26].

In this paper, we propose a UWB radar technique for breast cancer detection based on the use of artificial neural networks (ANNs). The main purpose is not to image the cancerous breast, but to propose a new diagnostic tool as an aid to the work of a medical clinician operator in order to determine the presence or absence of a tumor independent of its depth and width. The use of ANNs provides several advantages, such as short computation times, low computational burden, and the opportunity to recast the problem by considering only a few unknowns of interest.

To assess our proposed approach for a significant scenario, we used two-dimensional (2D) and three-dimensional (3D) realistic breast models that were derived from the database made available by the numerical breast phantom repository of the University of Wisconsin cross-disciplinary electromagnetic laboratory (UWCEM) [27]. Because of the importance of removing the predominant reflections due to the presence of the skin, our ANN-based approach was assessed and tested using both an ideal skin artifact removal technique and a model-based one [26].

This paper first presents a detailed description of the main parts that constitute the proposed method. Second, by using realistic 2D breast models and applying an ideal cleaning technique, we present the results obtained in the case of tumors positioned both outside and inside the fibro-glandular tissues. Moreover, during these first analyses, the use of two different UWB pulses is assessed. In the third part, because of the best results obtained especially for tumors located outside the fibro-glandular tissues, we focused on this problem by considering a more generic and realistic scenario of testing data. In particular, we assessed the performance in the case where a realistic model-based skin artifact removal technique [26] was used. Moreover, some test on a realistic 3D breast model with ideally cleaned radar signals were also performed. Finally, the improvements introduced by the application of the proposed diagnostic criterion are presented, and a conclusions section closes the paper.

## II. MATERIALS AND METHOD

In this section, the main issues and characteristics of the proposed approach are described through the following five topics: geometries and system configuration, signal pre-processing, ANN processing, diagnostic criterion, and analysis of single backscattered radar signals.

### A. Geometries and System Configuration

We assumed the patient lying in the prone position. The acquisition system consists of a mono-static radar system that collects the backscattered signals in different space locations.

The measurement points are situated along a circumference around the breast and at different distances from the chest.

In this context, to assess the performance of our ANN-based radar data processing approach, we considered a set of 2D healthy and cancerous breast geometries. In order to work with the most significant possible scenario, these geometries were derived from the 3D realistic breast models that are freely provided by the UWCEM database [27]. Each of these geometries distinguishes eight different typologies of healthy breast tissue, namely, the skin, three types of adipose tissue, a transitional tissue, and three different typologies of fibro-glandular tissue.

In the present work, starting from such models, we built different 2D breast geometries taking into account different cross sections. In order to describe a broader set of breasts characterized by different densities, the values for the dielectric characterization of each internal tissue, according to the Debye parameters, namely the static relative permittivity ( $\epsilon_s$ ), the relative permittivity at infinite frequency ( $\epsilon_\infty$ ), the conductivity ( $\sigma$ ), and the relaxation time ( $\tau$ ), were randomly chosen within the range of the values provided by [27]. In conclusion, each healthy breast geometry was built by randomly choosing the UWCEM model, the section, and the Debye parameters ( $\epsilon_s$ ,  $\epsilon_\infty$ ,  $\sigma$ ,  $\tau$ ) for the dielectric characterization of the internal tissues. The cancerous geometries were built by starting with new healthy geometries, chosen in the same randomly way, in which we inserted a dielectric anomaly of different width, with random diameter values chosen between 0.2 cm and 1 cm, and different depths randomly chosen between 0.5 cm from the outer skin surface and the center of the probing line. The Debye parameters for the tumor were obtained by minimizing a suitable cost function [28] from the Cole-Cole representation provided in [29]. In particular, it was characterized by an  $\epsilon_s$  of 61.6,  $\epsilon_\infty$  of 14.5,  $\sigma$  of 0.7 S/m, and a  $\tau$  of 13 ps.

For each radar angular position, the breast geometry is illuminated with a UWB pulse, and the time wave of the backscattered signal is obtained through numerical simulation performed by means of the finite-difference time-domain (FDTD)-based open source software GprMax [30]. Following the latest proposals in the literature [15, 16, 23, 25, 31], the illuminating signal consists of a differentiated Gaussian pulse (DGP). Compared to the simple Gaussian pulse, the DGP provides a higher backscattered signal strength and, consequently, it is more suitable in detecting deep targets that are sparsely distributed [32].

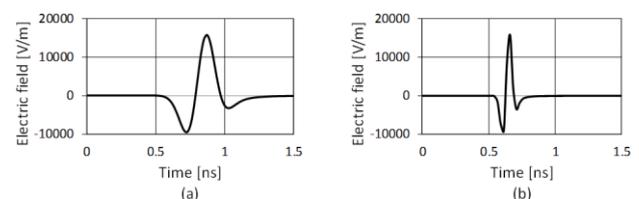


Fig. 1. UWB pulses used as incident signals: (a) DGP of 1 ns and central frequency 2 GHz; (b) DGP of 0.3 ns and central frequency 6 GHz.

In the present work, the performance of two different DGP pulses was assessed, namely, a DGP with a duration of 0.3 ns

and central frequency 6 GHz, and a DGP with a duration of 1 ns and central frequency 2 GHz. A graphic representation of these pulses is shown in Fig. 1.

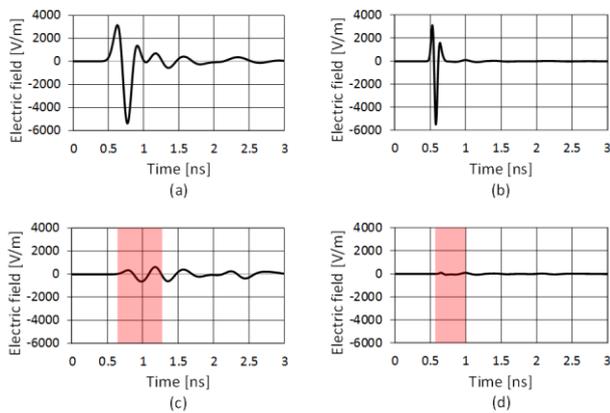


Fig. 2. Comparison between a backscattered radar trace and its cleaned version. (a), (c) DGP with central frequency 2 GHz. (b), (d) DGP at 6 GHz. (a), (b) Waveforms of the total backscattered radar signals measured on a cancerous geometry. (c), (d) The radar signals after the application of the ideal cleaning technique, where the tumor signal contribution is marked in red.

### B. Signal Pre-Processing

Once the backscattered radar signals are collected, suitable and significant data are extracted in order to train and test ANNs to detect the presence of a tumor. Unfortunately, as mentioned previously, the backscattered signals do not contain only the reflections due to the presence of the healthy and malignant internal breast tissues; they also contain those due to the presence of the skin. This is a problem because the predominant nature of the skin artifact component makes it difficult to choose suitable data that best characterize the presence of a tumor.

In order to better focus on this problem, Fig. 2 shows a comparison between the radar signals obtained before and after the application of an ideal cleaning technique that consists of the removal of the known skin response. In particular, the ideally cleaned radar signal is obtained by subtracting, from the total backscattered signal, the signal measured on a geometry with equal shape and dimensions but formed by only the skin and adipose tissue. Fig. 2(a) and 2(c) show the signals measured by using the DGP incident pulse with central frequency 2 GHz, whereas Fig. 2(b) and 2(d) show the signals obtained with the DGP at 6 GHz. Fig. 2(a) and 2(b) show the total backscattered radar signals measured on a realistic geometry, and Fig. 2(c) and 2(d) present the same radar signals after the application of the ideal cleaning technique. The total radar signals were measured on a 2D realistic cancerous breast geometry in which a dielectric anomaly, with a diameter of 6 mm, was inserted at a depth of 1.5 cm from the outer skin surface.

Comparing the cleaned signals with the corresponding total ones, it is possible to observe that the skin artifact component is predominant and of different orders of magnitude higher with respect to the other signal components. Moreover, it overlaps the reflections of the internal tissues, masking the contribution of the signal that contains the

information inside the breast geometry. These considerations highlight the necessity to apply a cleaning technique in order to extract significant information on the presence of dielectric anomalies.

Finally, because of the non-circular shapes of the 2D realistic breast geometries, the distance between the antenna and the skin is not constant, and the times of arrival and the amplitudes of the measured backscattered signals change because of the varying mono-static radar position. In order to reduce this space-temporal error, we equalized the times and amplitudes. With a cross-correlation technique, the time of arrival of the total backscattered signal is computed for each radar position, and the resulting values are used to choose a zero reference time for all the recorded radar signals. Moreover, the signal amplitude is multiplied by a coefficient that considers the temporal shift.

### C. ANN Processing

Starting from the cleaned and equalized radar signals, we propose to use the amplitudes ( $A_1, \dots, A_i, \dots, A_N$ ) and the arrival times ( $t_1, \dots, t_i, \dots, t_N$ ) of  $N$  suitable local maxima and minima as significant data to train and test the artificial neural network.

The ANN used in the present method is a multilayer feed-forward fully connected network [33]. It is formed by an input layer, an output layer, and one or more hidden layers. Moreover, each node of each layer is connected to every one of the adjacent layers. The number of nodes that we used for the input layer and for the hidden ones will be discussed in the following sections. For the purposes of the diagnostic method, it is instead important to note that we always used ANNs having as the output layer only a single node that provides a signal of type Yes/No, depending on the presence (Yes) or absence (No) of the tumor.

### D. Analysis of Single Backscattered Radar Signals

The mandatory objective of our approach is first to develop an accurate algorithm that can detect cancer by working on a single radar trace. In this context, a fundamental step is to choose the information contained in the radar signal in order to extrapolate data that best characterize the presence of the tumor. To this end we analysed many radar signals obtained for several healthy breast geometries in which a dielectric anomaly was inserted under different conditions. The reached conclusions can be described considering the situations reported in Fig. 3.

Fig. 3(a), 3(c), and 3(e) show the signals measured by illuminating the geometry with the DGP with central frequency 2 GHz, whereas Fig. 3(b), 3(d), and 3(f) show the signals obtained with the DGP at 6 GHz. In particular, Fig. 3(a) and 3(b) present the cleaned radar signals measured on the breast geometry in absence of the tumor. Fig. 3(c) and 3(d) show the radar signals measured, on the same healthy geometry, in the case where a dielectric anomaly of diameter 6 mm is inserted at 1.5-cm depth from the outer skin surface and outside the fibro-glandular tissues. Fig. 3(e) and 3(f) report the radar signals measured in the case where the same dielectric anomaly has been inserted at 1-cm depth inside the fibro-

glandular tissues, corresponding to 3-cm depth from the skin surface.

Comparing Fig. 3(a) and 3(b) with Fig. 3(c) and 3(d), shows that, if the tumor is positioned outside the fibro-glandular tissues, the only the first part of the cleaned radar signal is more influenced by its presence. In fact, the greatest differences occur in the time interval between 0.5 ns and 1.5 ns.

In contrast, if a tumor is positioned inside the fibro-glandular tissues, the first part of the cleaned radar signal does not change significantly. In Fig. 3(a) and 3(e), namely in the case of the DGP with a 2-GHz central frequency, the major differences are contained in the time interval between 1.5 ns and 2.5 ns. Similarly, in Fig. 3(b) and 3(f), with the DGP at 6 GHz, the major differences are contained in the time interval between 1 ns and 2 ns.

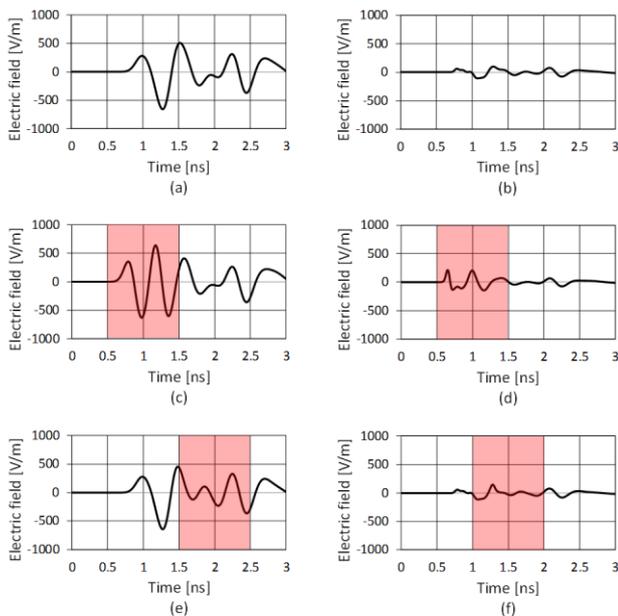


Fig. 3. Ideally cleaned and equalized radar signals measured on a healthy geometry in which a dielectric anomaly is inserted under different conditions. (a), (c), (e) DGP with central frequency 2 GHz. (b), (d), (f) DGP at 6 GHz. (a) and (b) Healthy breast geometry in absence of the tumor. (c) and (d) Tumor of 6-mm diameter positioned outside the fibro-glandular tissues at 1.5-cm depth from the skin surface. (e) and (f) Same tumor positioned 1 cm inside the fibro-glandular tissues and at 3-cm depth from the outer skin surface. The signal tumor contribution is marked in red.

These considerations suggest we research the tumors that are located outside and inside the fibro-glandular tissues separately. In this way, a process able to detect tumors positioned outside the fibro-glandular tissues is separately applied using only the first part of the cleaned radar signal, whereas another process is applied using the second part of the same signal in order to detect the presence of internal tumors.

Fig. 4 shows a schematic flow chart of the detection algorithm that we propose for single radar trace processing. Following this representation, from the cleaned and equalized radar signals, two sets of data,  $I_1$  and  $I_2$ , are extracted and provided to their properly specialized ANNs. These ANNs are described in detail in the next section.

E. Diagnostic Criterion

As previously described, the first key step of the present approach is to extract useful and suitable information from any single radar trace recorded for each different angular position. Nevertheless, in order to improve the detection accuracy, we propose a diagnostic criterion based on a collective analysis of all the obtained results.

Because our approach provides only a Yes/No answer for each single radar position, and each answer may be a false-positive or a false-negative result, we aim to assess the capability to reach a reliable diagnostic response by giving credibility to a Yes answer only when such output is obtained consecutively for a given arc, or for a given percentage of Yes inside it, and not to isolated positive detection positions. The same procedure is applied for the negative tumor diagnosis.

Based on this idea, it will be possible not only to improve the specificity and sensitivity values of the whole diagnostic process, but also to better localize the angular sector where the tumor is located.

III. RESULTS AND DISCUSSION

A. Tumors Positioned Outside the Fibro-glandular tissues

In this section, we present the ANN architecture designed for the detection of tumors located outside the fibro-glandular tissues. As mentioned in Section II.D, in these situations, it is valid to assume that the information on the tumor presence is contained in the first part of the cleaned and equalized radar signal. Following these considerations, as shown in Fig. 5, we exploited the information contained in the first two local maxima/minima of the cleaned radar signals. In particular, Fig. 5(a) highlights the first two peaks of the cleaned signal measured by using the DGP with central frequency 2 GHz, whereas Fig. 5(b) presents the case of the DGP at 6 GHz.

In this section, with the aim of preliminarily assessing the capabilities of our ANN-based detection approach, we obtained the cleaned radar signals by applying an ideal cleaning technique. As mentioned in Section II.B, it consists of subtracting, from the total backscattered radar signal, the signal that is measured on a geometry with equal shape and dimension but formed by only the skin and adipose tissues, where these two tissues have the same Debye dielectric characterization of the respective realistic geometry.

We used an ANN architecture of type 4-8-1. It receives four input data, namely the amplitudes and arrival times of the

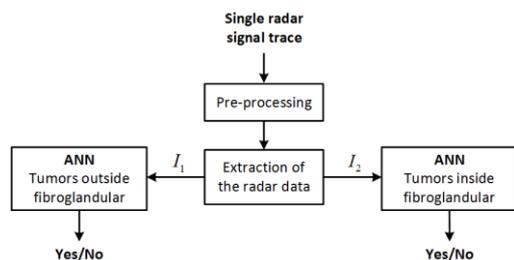


Fig. 4. Schematic flow chart of the detection algorithm applied to each single radar trace. The total backscattered radar signal is first pre-processed and suitable radar data,  $I_1$  and  $I_2$ , are extracted. Then,  $I_1$  and  $I_2$  are processed by the corresponding ANN in order to detect the presence of tumors respectively located outside and inside the fibro-glandular tissues.

first two maxima/minima measured on the ideally cleaned radar signals, and provides one output of type Yes/No, depending on the presence or absence of the tumor. This ANN has one hidden layer of eight nodes. According to the number of degrees of freedom of the ANN 4-8-1, the network is trained with 100 training data, namely 50 measured on healthy geometries (NT) and 50 measured on cancerous geometries (T). Using this ANN architecture, we trained two different networks: the first by using the radar signals measured in the case of the DGP at 2 GHz (hereafter named ANN-4-8-1-Text-2-GHz), and the second by using the DGP at 6 GHz (hereafter named ANN-4-8-1-Text-6-GHz).

For the purposes of training the ANN, we positioned the anomaly by respecting one constraint: the anomaly must be the first internal discontinuity that the radar signal encounters within the geometry of the breast. An example of such a geometry is shown in Fig. 6.

The results on the training data gave a tumor detection accuracy of 100%. This means that the ANN was able to match all the given pairs of input-output examples correctly.

In order to assess the generalization capabilities of the two ANNs, we generated two new sets of 100 test data, namely 50 T and 50 NT. The first set was obtained by using the DGP at 2 GHz, and the other by using the DGP at 6 GHz. The test data were simulated by always respecting the constraint on the tumor's position. Table I shows the results in terms of the confusion matrix for ANN-4-8-1-Text-2-GHz, and Table II shows the results obtained for ANN-4-8-1-Text-6-GHz. Table I shows that, in the case of the DGP at 2 GHz, the method can detect the presence of the anomaly with a sensitivity of 92%, a specificity of 90%, and an accuracy of 91%. By contrast, when the DGP at 6 GHz is used, the anomaly is detected with a sensitivity of 88%, a specificity of 78%, and an accuracy of 83%.

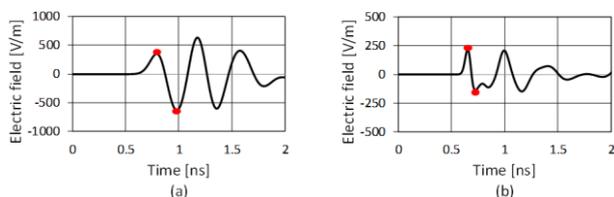


Fig. 5. Ideally cleaned and equalized radar signal measured on a 2D realistic heterogeneous geometry. (a) DGP at 2 GHz. (b) DGP at 6GHz.

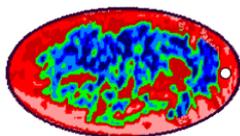


Fig. 6. Example of a cancerous geometry in which a dielectric anomaly of diameter 4 mm is located outside the fibro-glandular tissues at a depth of 0.5 cm from the skin surface. The blue colors indicate the three typologies of fibro-glandular tissue, the green parts represent the intermediate tissue, and the red colors indicate the three typologies of adipose tissue.

TABLE I. Confusion matrix, ANN-4-8-1-Text-2-GHz, 100 testing examples, constraints on the T location, and ideal cleaning technique. The overall accuracy is 91%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	50	46	4	92	/
NT	50	5	45	/	90

TABLE II. Confusion matrix, ANN-4-8-1-Text-6-GHz, 100 testing examples, constraints on the T location, and ideal cleaning technique. The overall accuracy is 83%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	50	44	6	88	/
NT	50	11	39	/	78

### B. Tumors Positioned Inside the Fibro-glandular tissues

This section focuses on the problem of detecting tumors located inside the fibro-glandular tissues. Following the concepts explained in Section II.D, in order to best characterize the presence of the tumor, we searched for suitable information by using the second part of the cleaned and equalized radar signals.

In order to train and test the new ANNs, we built new healthy and cancerous geometries with the dielectric anomaly positioned inside the fibro-glandular tissues. An example of such a geometry is shown in Fig. 7. Starting from the ideally cleaned and equalized radar signals, different network architectures—having a different number of both hidden layers and nodes, and providing only one output node of type Yes/No—were trained and tested. In particular, we trained different ANNs that receive six input data, namely the amplitudes and arrival times of the 3<sup>rd</sup>, 4<sup>th</sup>, and 5<sup>th</sup> maximum/minimum of the cleaned and equalized radar signals.

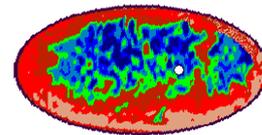


Fig. 7. A cancerous geometry in which a dielectric anomaly of diameter 4 mm is located inside the fibro-glandular tissues at a depth of 3.5 cm from the skin surface. The blue colors indicate the three typologies of fibro-glandular tissue, the green parts represent the intermediate tissue, and the red colors indicate the three typologies of adipose tissue.

TABLE III. Confusion matrix, DGP at 2 GHz, 100 testing examples, and ideal cleaning technique. The overall accuracy is 68%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	50	37	13	74	/
NT	50	19	31	/	62

TABLE IV. Confusion matrix, DGP at 6 GHz, 100 testing examples, and ideal cleaning technique. The overall accuracy is 51%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	50	28	22	56	/
NT	50	27	23	/	46

In order to test the ANNs, we generated two new sets of 100 test data, each with 50 T and 50 NT; the first obtained by using the DGP at 2 GHz and the second with the DGP at 6 GHz. Table III and Table IV show the best results in terms of the confusion matrix obtained respectively for the DGP at 2 GHz and at 6 GHz. Unfortunately, comparing the Tables III and IV shows that only the use of the DGP at 2 GHz had acceptable results by providing a tumor detection accuracy of 68%.

C. Robustness Assessment Study of ANN-4-8-1-Text-2-GHz

In the previous sections, we showed the best results, reported in Table I, obtained by using ANN-4-8-1-Text-2-GHz. This network is specialized for the detection of tumors located outside the fibro-glandular tissues by working on the input data that are extracted from the cleaned radar signals measured by using the DGP at 2 GHz. Because the previous analyses were performed considering a limited scenario and ideally cleaned radar signals, in this section, we present a robustness assessment study of such a network.

First, we will consider a larger data set for which the tumor is positioned without the constraint before used (see Section III.A). For this dataset we introduced also the case when the anomaly is positioned in contact with the outer surface of the fibro-glandular tissues. Moreover, an analysis of the sensitivity, when varying both the depth and dimensions, is presented.

As the second step, we assess the performance in the case when the skin response is suppressed by using a realistic model-based skin artifact removal technique.

Finally, the network is tested by using the data ideally cleaned but measured on a realistic 3D breast model.

a) Generic testing data without constraints on the tumor position

In order to test the ANN in a more general way, we generated 2000 new realistic geometries, 1000 T and 1000 NT, with the anomaly positioned outside the fibro-glandular tissues but without any constraint. An example of such a cancerous geometry is shown in Fig. 8. The cleaned radar signals were obtained by applying the ideal cleaning technique discussed in Section II.B.

Table V shows the results in terms of the confusion matrix. The table shows that the network detected the tumor with a sensitivity of 81%, a specificity of 79%, and an accuracy of 80%.

Fig. 9 shows a histogram that summarizes the sensitivity when considering, among the 1000 T cases used for Table V, different single classes according to the depth and dimension of the tumor. The depth is measured from the outer surface of the skin. The figure shows that in the case where the tumor's depth ranges between 0.5 cm and 1.5 cm, the sensitivity is 71% in the case of tumors diameter between 2 mm and 4 mm, 94% for diameters between 4 mm and 6 mm, 88% for diameters between 6 mm and 8 mm, and 79% for diameters between 8 mm and 10 mm. In contrast, for tumor depths greater than 1.5 cm, the sensitivity is 77% for tumor diameters between 2 mm and 4 mm, 71% for 4-mm–6-mm diameters, 74% for 6-mm–8-mm diameters, and 9% for 8-mm–10-mm diameters. It worth noting that, if the analysis is limited to tumors with depths ranging between 1,5 cm and 2.5 cm, the sensitivity reaches the 70%. Moreover, if among the 1000 T cases used for Table V we exclude the tumors with depth greater than 2.5 cm, the global sensitivity increases to 83%.

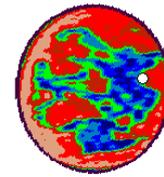


Fig. 8. Example of a cancerous geometry used for the general test with 2000 examples. A tumor of diameter 4 mm is positioned at a depth of 1 cm from the skin surface.

TABLE V. Confusion matrix, ANN-4-8-1-Text-2-GHz, 2000 testing examples, no constraints on the T location, and ideal cleaning technique. The overall accuracy is 80%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	1000	810	190	81	/
NT	1000	214	786	/	79

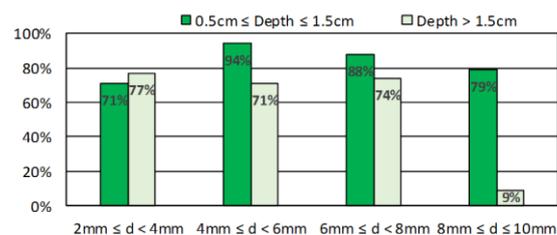


Fig. 9. The sensitivity for different classes of the depth and dimension of the tumor. The 1000 cancerous geometries used for Table V were here considered.

b) Application of a model-based skin artifact removal technique

In the previous sections, ANN-4-8-1-Text-2-GHz was assessed by using input data measured on ideally cleaned radar signals. The analysis of the network's performance when applied to an ideal scenario is important in order to provide a baseline performance of the algorithm. However, a review of the literature on the topic of skin suppression techniques [22, 23, 24, 25], suggests that effective skin response suppression is a challenging task. In fact, in a realistic scenario, the skin response suppression algorithm may suppress the skin response, but significant residual effects may be present after the algorithm is applied. Moreover, the skin response suppression algorithm may alter the tumor response and degrade the performance of the detecting technique. In this section, we present an analysis of the performance of our ANN-based approach in the case where the backscattered fields are processed with a realistic model-based skin artifact removal algorithm, already proposed by the authors [26].

This technique includes the use of a reference cleaning model to obtain the backscattered signal to be used as the cleaning signal. This last is subtracted from the total real one in order to obtain the cleaned signal. The reference model consists of a bi-layered cylinder (skin and adipose tissue) characterized by suitable dielectric Debye parameters for each of these two tissues. Other important parameters for its characterization are the skin thickness, the radius, and the distance between the skin interface and the radar antenna.

In [26], the introduced signal distortions were found to be minimized in the case where a reference cleaning model with a radius of 11 cm is used by placing it at the same real distance between the skin and the antenna. Moreover, the best

results were obtained under the hypothesis of knowing the real values of the skin thickness and the static dielectric permittivity ( $\epsilon_s$ ) of both the skin and adipose tissue, while the other dielectric parameters ( $\epsilon_\infty$ ,  $\sigma$ , and  $\tau$ ) are fixed at average standard values. These last were obtained by averaging the range of values provided in the literature [27].

Respecting these constraints, we configured the reference cleaning models and applied the model-based skin artifact removal algorithm to the 2000 realistic models used and described in Section III.C.a. After the cleaned signals were obtained, the amplitudes and arrival times of the first two maxima/minima were measured and tested on the ANN-4-8-1-Text-2-GHz. The resulting confusion matrix is reported in Table VI, and it shows that the network detected the tumor with a sensitivity of 78%, a specificity of 74%, and an accuracy of 76%.

Moreover, in [26], satisfactory results were achieved also in the case where more generic reference cleaning models were used. These last are characterized by average standard values for all the Debye dielectric parameters ( $\epsilon_s$ ,  $\epsilon_\infty$ ,  $\sigma$ , and  $\tau$ ) of both the skin and adipose tissue. The standard values are obtained by averaging the range of values provided in [27]. Because better results are obtained where the network is trained using training data measured on radar signals cleaned by using the generic reference cleaning models [26], we trained a new network, named ANN-4-8-1-Text-2-GHz-Generic-Model-Based, by using the generic reference cleaning models made up of dielectric standard values. Table VII shows the results obtained by testing the new network by using the same 2000 realistic breast models described in Section III.C.a, and the generic model-based cleaning technique.

Table VII shows that the tumor is detected with a sensitivity of 74%, a specificity of 73%, and an accuracy of 74%.

This last result is highly interesting because it means that a satisfactory accuracy can be reached also in the case when a skin suppression algorithm is applied without any a priori knowledge of both the dielectric and geometric properties of the skin and the adipose tissue, as happens in practical situations. The only parameter that was assumed to be known is the skin thickness.

c) *Test with a 3D realistic cancerous breast model*

In this section, we present the results obtained by testing ANN-4-8-1-Text-2-GHz with data obtained in the case of a realistic 3D breast model available on the UWCEM database. In particular, in our simulations we used the breast phantom, identified in such a database by ID 010204, which belongs to the so called scattered fibro-glandular class. The Debye parameter values used for the dielectric characterization of the healthy tissues are reported in Table VIII [27].

TABLE VI. Confusion matrix, ANN-4-8-1-Text-2-GHz, 2000 testing examples, no constraints on the T location, and model-based cleaning technique. The overall accuracy is 76%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	1000	784	216	78	/
NT	1000	257	743	/	74

TABLE VII. Confusion matrix, ANN-4-8-1-Text-2-GHz-Generic-Model-Based, 2000 testing examples, no constraints on the T location, and generic model-based cleaning technique. The overall accuracy is 74%.

Type	Test Number	ANN Output		ANN Performance (%)	
		Yes	No	Sensitivity	Specificity
T	1000	742	258	74	/
NT	1000	267	733	/	73

The tumor was dielectrically characterized with a static relative permittivity ( $\epsilon_s$ ) of 61.6, a relative permittivity at infinite frequency ( $\epsilon_\infty$ ) of 14.5, a conductivity ( $\sigma$ ) of 0.7 S/m, and a relaxation time ( $\tau$ ) of 13 ps.

The network was tested for different tumor situations obtained by positioning the anomaly always at a depth of 1.5 cm from the skin surface but at three different distances from the chest (2 cm, 5 cm, and 8 cm). For each position, we considered tumors of different diameters: 2 mm, 4 mm, 6 mm, 8 mm, and 10 mm. Fig. 10 shows a sagittal section of the 3D model in the case where a tumor with a diameter of 4 mm is positioned at a depth of 1.5 cm from the skin and at a distance of 5 cm from the chest.

In order to suppress the skin reflections, we applied the ideal technique described in Section II.B. Starting from the ideally cleaned radar signal, we extracted both the amplitudes and arrival times of the first two maxima/minima, and we used them to test the network previously trained and tested for 2D geometries (ANN-4-8-1-Text-2-GHz).

From all the obtained results, we found that this network correctly detected tumors positioned at a depth of 1.5 cm and characterized by a diameter as small as 2 mm.

D. *Diagnostic Criterion Assessment*

In this section, we report some numerical results that were obtained by applying the diagnostic criterion described in Section II.E. Because our aim is to preliminarily assess the improvements that can be reached, we choose to present the performance introduced in the case when the algorithm is applied to an ideal scenario, namely, when the cleaned radar signals are obtained by using the ideal cleaning technique described in Section II.B.

In this context, we considered the case of tumors positioned outside the fibro-glandular tissues when the DGP at 2 GHz is used, and we analyzed 10 2D realistic breast models (five T and five NT) chosen from the 2000 models described in Section III.C.a. For each of these we collected 72 different radar signals by positioning the mono-static radar antenna in 72 different angular positions around the breast geometry. Starting from the ideally cleaned and equalized radar signals, the amplitudes and arrival times of the first two maxima/minima were measured, and ANN-4-8-1-Text-2-GHz was tested. In fact, applying the diagnostic criterion on the outcomes provided by the network, we obtained highly satisfactory results.

In the case of the NT geometries, even if the network wrongly detected the anomaly as malignant in isolated angular positions, by applying the diagnostic criterion the definitive outcomes of our ANN-based detection approach furnished only true-negative results. Moreover, for the T geometries, the obtained results prove that the ANN produced Yes answers in

some isolated angular positions but also consecutively inside a cone of approximately 30° centered on the tumor's location.

TABLE VIII. Values of the Debye parameters used for dielectric characterization of healthy breast tissues.

Breast Tissue	$\epsilon_s$	$\epsilon_\infty$	$\sigma$ [S/m]	$\tau$ [ps]
Fibrogland-1	54.690	14.200	0.824	13.00
Fibrogland-2	49.360	13.810	0.738	13.00
Fibrogland-3	37.390	12.990	0.397	13.00
Intermediate	22.461	8.4890	0.239	13.00
Adipose-1	7.532	3.987	0.080	13.00
Adipose-2	4.708	3.116	0.050	13.00
Adipose-3	3.952	2.848	0.005	13.00
Skin	39.760	15.930	0.831	13.00

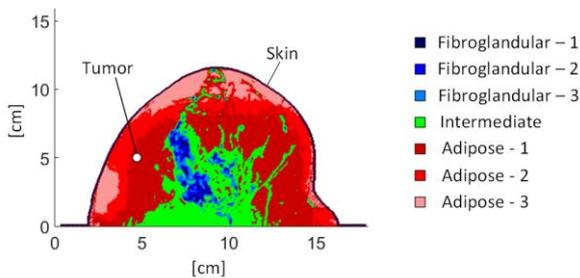


Fig. 10. Sagittal section of the used 3D model. A tumor of 4-mm diameter is positioned at a depth 1.5 cm from the skin and at a distance 5 cm from the chest.

Because the true-positives were detected under angles of approximately 30° of consecutive malignant positive detections, and the true-negatives were detected under larger angles of consecutive non-malignant negative detections, we found it acceptable to give credibility only to arcs of at least 30° of consecutive Yes, or at least containing a percentage of them greater than a fixed threshold. For the analyzed cases, such a percentage was always at least 70% for angles of 30°. As examples, in Fig. 11, 12, and 13, we show three of the 10 studied geometries.

In particular, Fig. 11 shows an NT geometry on which, for all the angular positions, the outcomes of the ANN are reported. The isolated false-positives are highlighted in red as Yes points. The correct diagnosis is obtained by giving credibility only on the arcs of negative detections (the black points).

Fig. 12 shows the outcomes obtained by testing a cancerous geometry. All the malignant detections are highlighted as Yes points. The figure shows that the tumor is detected as malignant under an angle of approximately 30° of consecutive Yes points.

Finally, Fig. 13 shows the case of a T geometry for which, inside the arc of Yes, there is a single isolated No answer. In this case, by giving more credibility to the arc of Yes, we avoid failure in the case of a false-negative result.

The above considerations highlight the importance of applying the diagnostic criterion in order to provide high values of sensitivity and specificity, reaching for the cases here presented an accuracy of 100%, but also the possibility of correctly localizing the angular position of the detected tumor.

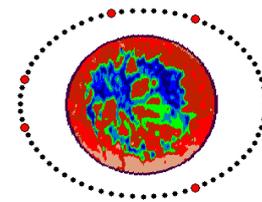


Fig. 11. Representation of a healthy breast geometry for which the ANN detects the presence of a malignant tumor only on single isolated angular positions (marked in red).

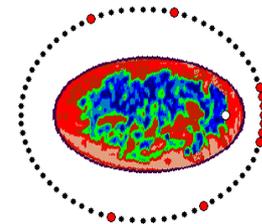


Fig. 12. Representation of a cancerous breast geometry for which the ANN identifies the anomaly as malignant in a cone of different positions centered on the location of the T and in other single isolated positions (marked in red). The tumor is correctly detected under a 30° angle of 100% consecutive single malignant detections.

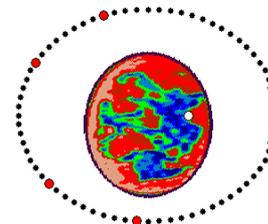


Fig. 13. Representation of a cancerous geometry for which there is one isolated incorrect result inside the arc of Yes detections. The tumor is correctly detected under a 30° angle with 85% of single malignant detections (marked in red).

#### IV. CONCLUSIONS

In this paper, we presented a new UWB radar technique for breast cancer detection based on the use of artificial neural networks. Our purpose was not to achieve imaging of cancerous breasts, but to provide a Yes/No diagnostic tool with the highest possible accuracy. The first key step is the ANN processing of any single radar trace recorded around the breast to determine the presence or absence of a tumor. Then, a diagnostic criterion is applied on the basis of a collective evaluation.

The numerical assessments were conducted by using 2D and 3D healthy and cancerous breast geometries derived from the models of the UWCEM database. A mono-static radar configuration was assumed to measure the simulated backscattered signals along a circular line around the breast and at different distances from the chest.

First, by using realistic 2D breast models and an ideal skin artifact removal technique, we studied the cases of tumors located both outside and inside the fibro-glandular tissues. Moreover, two different UWB incident pulses were assessed. The best results were obtained by using the differentiated

Gaussian pulse with a central frequency of 2 GHz: in the case of tumors located outside the fibro-glandular tissues, an overall accuracy of 91% on 100 testing data was achieved, whereas for the tumors positioned inside the fibro-glandular tissues the accuracy was 68%.

Because of the successful results obtained in the case of tumors located outside the fibro-glandular tissues, we focused on this topic and presented a robustness assessment study by using a more generic and realistic scenario of testing data. First, the ANN was tested by considering a larger and more generic set of examples consisting of 2000 2D breast geometries. The results showed the ability to detect tumors with a sensitivity of 81%, a specificity of 79%, and an overall accuracy of 80%. Moreover, when excluding tumors with depth from the outer skin surface greater than 2.5 cm, the sensitivity increases to 83%. Second, we assessed the performance in the case where a realistic model-based skin artifact removal technique is used in order to remove the strong skin reflections. In this situation, the tumor was detected with a sensitivity of 74%, a specificity of 73%, and an accuracy of 74%. Moreover, the ANN was tested by using a realistic 3D breast model. In this realistic case, our proposed technique correctly detected tumors characterized by diameters as small as 2 mm, located at different distances from the chest, and at a depth of 1.5 cm from the skin surface.

Finally, we applied the proposed diagnostic criterion to 10 realistic 2D geometries. For the cases here considered, we found that it reduces incorrect diagnostic responses completely to zero, furnishing in the same time the exact angular position of the tumor.

REFERENCES

[1] American Cancer Society, *Cancer Facts & Figures 2015*, Atlanta, GA, USA, American Cancer Society, 2015.

[2] L. Tabár and P. B. Dean, "A new era in the diagnosis and treatment of breast cancer," *The Breast Journal*, vol. 16, pp. S2–S4, 2010.

[3] A. Berrington de González and S. Darby, "Risk of cancer from diagnostic X-rays: Estimates for the U.K. and 14 other countries," *The Lancet*, vol. 363, no. 9406, pp. 345–351, 2004.

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

[5] W. A. Berg, L. Gutierrez, M. S. Ness-Aiver, W. B. Carter, M. Bhargavan, R. S. Lewis, and O. B. Ioffe, "Diagnostic accuracy of mammography, clinical examination, U.S., and MR imaging in preoperative assessment of breast cancer," *Radiology*, vol. 233, no. 3, pp. 830–849, 2004.

[6] A. M. Hassan and M. El-Shenawee, "Review of electromagnetic techniques for breast cancer detection," *IEEE Reviews in Biomedical Engineering*, vol. 4, pp. 103–118, 2011.

[7] S. C. Hagness, A. Taflove, and J. E. Bridges, "Two-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Fixed-focus and antenna-array sensors," *IEEE Transactions on Biomedical Engineering*, vol. 45, no. 12, pp. 1470–1479, 1998.

[8] B. Bocquet, J. C. Van de Velde, A. Mamouni, Y. Leroy, G. Giaux, J. Delannoy, and D. Delvalet, "Microwave radiometric imaging at 3 GHz for the exploration of breast tumors," *IEEE Transactions on Microwave Theory and Techniques*, vol. 38, no. 6, pp. 791–793, 1990.

[9] F. Bardati and S. Iudicello, "Modeling the visibility of breast malignancy by a microwave radiometer," *IEEE Transactions on Biomedical Engineering*, vol. 55, no. 1, pp. 214–221, 2008.

[10] L. V. Wang, X. Zhao, H. Sun, and G. Ku, "Microwave induced acoustic imaging of biological tissues," *Review of Scientific Instruments*, vol. 70, no. 9, pp. 3744–3748, 1999.

[11] J. D. Shea, B. D. Van Veen, and S. C. Hagness, "A TSVD analysis of microwave inverse scattering for breast imaging," *IEEE Transactions on Biomedical Engineering*, vol. 59, no. 4, pp. 936–945, 2011.

[12] Q. Fang, P. M. Meaney, and K. D. Paulsen, "Viable three-dimensional medical microwave tomography: Theory and numerical experiments," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 2, pp. 449–458, 2010.

[13] M. Klemm, I. J. Craddock, J. A. Leendertz, A. Preece and R. Benjamin, "Radar-based breast cancer detection using a hemispherical antenna array—experimental results," *IEEE Transactions on Antennas and Propagation*, vol. 57, no. 6, pp. 1692–1704, 2009.

[14] M. Klemm, J. A. Leendertz, D. Gibbins, I. J. Craddock, A. Preece, and R. Benjamin, "Microwave radar-based differential breast cancer imaging: Imaging in homogeneous breast phantoms and low contrast scenarios," *IEEE Transactions on Antennas and Propagation*, vol. 58, no. 7, pp. 2337–2344, 2010.

[15] E. C. Fear, J. Bourqui, C. Curtis, D. Mew, B. Docktor, and C. Romano, "Microwave breast imaging with a monostatic radar-based system: A study of application to patients," *IEEE Transactions on Microwave Theory and Techniques*, vol. 61, no. 5, pp. 2119–2128, 2013.

[16] E. C. Fear, X. Li, S. C. Hagness and M. A. Stuchly, "Confocal microwave imaging for breast cancer detection: Localization of tumors in three dimensions," *IEEE Transactions on Biomedical Engineering*, vol. 49, no. 8, pp. 812–822, 2002.

[17] Y. Xie, B. Guo, L. Xu, J. Li, and P. Stoica, "Multistatic adaptive microwave imaging for early breast cancer detection," *IEEE Transactions on Biomedical Engineering*, vol. 53, no. 8, pp. 1647–1657, 2006.

[18] Y. Chen, I. J. Craddock, P. Kosmas, M. Ghavami, and P. Rapajic, "Multiple-input multiple-output radar for lesion classification in ultrawideband breast imaging," *IEEE Journal of Selected Topics in Signal Processing*, vol. 4, no. 1, pp. 187–201, 2010.

[19] E. A. Marengo, F. K. Gruber, and F. Simonetti, "Time-reversal MUSIC imaging of extended targets," *IEEE Transactions on Image Processing*, vol. 16, no. 8, pp. 1967–1984, 2007.

[20] M. D. Hossain, A. S. Mohan and M. J. Abedin, "Beam-space time-reversal microwave imaging for breast cancer detection," *IEEE Antennas and Wireless Propagation Letters*, vol. 12, pp. 241–244, 2013.

[21] D. Byrne, M. O'Halloran, M. Glavin, and E. Jones, "Data independent radar beamforming algorithms for breast cancer detection," *Progress In Electromagnetics Research*, vol. 107, pp. 331–348, 2010.

[22] T. Yin, F. H. Ali, and C. C. Reyes-Aldasoro, "A robust and artifact resistant algorithm of ultrawideband imaging system for breast cancer detection," *IEEE Transactions on Biomedical Engineering*, vol. 62, no. 6, pp. 1514–1525, 2015.

[23] W. Zhi and F. Chin, "Entropy-based time window for artifact removal in UWB imaging of breast cancer detection," *IEEE Signal Processing Letters*, vol. 13, no. 10, pp. 585–588, 2006.

[24] A. Maskooki, E. Gunawan, S. B. Soh, and K. S. Low, "Frequency domain skin artifact removal method for ultra-wideband breast cancer detection," *Progress In Electromagnetics Research*, vol. 98, pp. 299–314, 2009.

[25] B. Maklad and E. C. Fear, "Reduction of skin reflections in radar-based microwave breast imaging," in *Proceedings 30th Annual International Conference of the IEEE Engineering in Medicine and Biology Society*, pp. 21–24, Vancouver, BC, August 2008.

[26] S. Caorsi and C. Lenzi, "Skin artifact removal technique for breast cancer radar detection," *Radio Science*, vol. 51, no. 6, pp. 767–778, 2016.

[27] E. Zastrow, S. K. Davis, M. Lazebnik, F. Kelcz, B. D. Van Veen, and S. C. Hagness, "Database of 3D Grid-Based Numerical Breast Phantoms for use in Computational Electromagnetics Simulations," *Department of Electrical and Computer Engineering University of Wisconsin-Madison*, <http://uwcem.ece.wisc.edu/MRI/database/InstructionManual.pdf>.

[28] M. Lazebnik, M. Okoniewski, J. H. Booske, and S. C. Hagness, "Highly accurate Debye models for normal and malignant breast tissue dielectric properties at microwave frequencies," *IEEE Microwave and Wireless Components Letters*, vol. 17, no. 12, pp. 822–824, 2007.

[29] M. Lazebnik, D. Popovic, L. McCartney, et al., "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign, and malignant breast tissues obtained from cancer surgeries," *Physics in Medicine and Biology*, vol. 52, no. 20, pp. 6093–9115, 2007.



- [30] A. Giannopoulos, "Modelling ground penetrating radar by GprMax," *Construction and Building Materials*, vol. 19, no. 10, pp. 755–762, 2005.
- [31] X. Li and S. C. Hagness, "A confocal microwave imaging algorithm for breast cancer detection," *IEEE Microwave and Wireless Components letters*, vol. 11, no. 3, pp. 130–132, 2001.
- [32] D. Uduwawala, "Gaussian vs differentiated gaussian as the input pulse for ground penetrating radar applications," in *Proceedings International Conference on Industrial and Information Systems (ICIIS 2007)*, pp. 199–202, Penadeniya, Sri Lanka, August 2007.
- [33] S. Haykin, *Neural Networks: A Comprehensive Foundation*, Prentice Hall, New York, NY, USA, 1994.