Progress In Electromagnetics Research, Vol. 144, 45–57, 2014

Investigation of Classifiers for Tumor Detection with an Experimental Time-Domain Breast Screening System

Adam Santorelli*, Emily Porter, Evgeny Kirshin, Yi Jun Liu, and Milica Popović

Abstract—In this work we examine, for the first time, the use of classification algorithms for early-stage tumor detection with an experimental time-domain microwave breast screening system. The experimental system contains a 16-element antenna array, and testing is done on breast phantoms that mimic breast tissue dielectric properties. We obtain experimental data from multiple breast phantoms with two possible tumor locations. In this work, we investigate a method for detecting the tumors within the breast but without the usual complexity inherent to image-generation methods, and confirm its feasibility on experimental data. The proposed method uses machine learning techniques, namely Support Vector Machines (SVM) and Linear Discriminant Analysis (LDA), to determine whether the current breast being scanned is tumor-free. Our results show that both SVM and LDA methods have promise as algorithms supporting early breast cancer microwave screening.

1. INTRODUCTION

Microwave breast screening and imaging modalities have been the focus of many studies over the past decade. Microwave methods offer several advantages over the standard techniques of ultrasound, x-ray mammography, and magnetic resonance imaging (MRI); (i) scans are painless and without breast compression, (ii) there is no use of potentially harmful ionizing radiation, and, (iii) microwave methods promise cost-effective screening [1]. In particular, research on microwave breast cancer detection, or imaging, can be subdivided into tomography and radar systems. Tomography aims to reproduce a full map of the dielectric properties of tissues inside the breast (for example, in [2–4]), whereas radar systems look to plot only regions of high dielectric scatters (e.g., the interface between healthy and malignant tissues) [5–7]. Further, each type of system can record measurements either in the frequency-domain (typically using a vector network analyzer (VNA)) [5, 8], or in the time-domain (using a pulse generator and an oscilloscope) [7, 9, 10].

The primary focus of many of these above mentioned systems is to perform breast cancer screening and detection through imaging, either by reconstruction of tissue dielectric properties or by identifying regions of electromagnetic scatterers. Numerous methods have been proposed and investigated for recreating such images, for instance, [8, 11, 12]; however, such methods can require high levels of mathematical and computational complexity. In the study presented here, we aim to avoid image reconstruction altogether in favor of classification techniques.

Various classification methods have been used successfully in related fields. In [13], mammogram images were analyzed to classify the breast tissue density in accordance with the BIRADS system using a k-Nearest Neighbors (kNN) classifier after feature extraction. In [14], the data obtained from electrical impedance measurements (400 kHz up to 1 MHz) of breast tissue samples was classified between non-carcinoma and carcinoma samples using Principle Component Analysis (PCA) to extract selected features and Linear Discriminant Analysis (LDA) to classify the tissue samples. Specifically, they observed that using a multi-stage classification algorithm comprising of PCA and LDA, data can first

Received 7 November 2013, Accepted 6 December 2013, Scheduled 16 December 2013

* Corresponding author: Adam Santorelli (adam.santorelli@mail.mcgill.ca).

The authors are with the Department of Electrical and Computer Engineering, McConnell Building, McGill University, 3480 University, Montreal H3A 0E9, Canada.

be classified between fatty and non-fatty tissues, and the non-fatty tissues can be then further separated between carcinoma and non-carcinoma tissues. Additionally, the classifier is designed to minimize the false-negative rate as it represents the most dangerous diagnosis [14].

Recent literature has shown that these and other similar classification techniques can be extended to microwave systems for breast cancer detection [6, 15–18], stroke detection [19], and bladder monitoring [20]. To date, the use of classifiers for breast cancer detection has focused on microwave systems using a numerical analysis; most recently the first experimental study, using frequency-domain measurements, was reported in [21]. Major differences between this paper and [21], aside from the use of time-domain measurements, are that we use realistic breast phantoms with inherent dielectric variability, and we attempt to classify between phantoms with and without a tumor.

One of the primary classification methods investigated for breast cancer detection has been Support Vector Machines (SVM) [6, 16–18]. The analysis in [21] was concerned with classifying tumors by shape and size, and demonstrated that the size and shape category of a tumor can be determined correctly in more than 86% of cases. In [18], it was found that the SVM classification was successful in determining whether the breast was healthy or not at least 70% of the time, and in many cases more than 80% of the time. The authors further showed that, in the series of tests they conducted, the SVM classifier always outperformed Linear Discriminant Analysis (LDA) [18].

In this paper, we present the application of machine learning classification techniques for breast screening on experimental signals obtained from breast phantoms with a range of realistic dielectric properties. In particular, we examine the accuracy of the LDA and SVM classifiers in detecting the presence of a tumor embedded in two possible locations within the breast phantom. Our system utilizes a time-domain microwave radar set-up, but instead of recreating a full image of breast scatterers we here analyze the scattered signals in such a way that the output is a numerical likelihood of whether the breast contains malignant tissue or not. This likelihood can, if desired, then be converted into a binary determination, i.e., "healthy" or "not-healthy". We use time-domain measurements over frequency-domain measurements because they offer the potential for faster scan times and improved cost-effectiveness [9]. This allows for improved accessibility for the general population, an important aspect for screening applications.

In the here presented analysis, we aim to emulate the clinical application of our system. The data required to train the classification algorithms is first obtained from a specific set of breast phantoms. Testing data is then obtained from newly fabricated phantoms; thus, we simulate the procedure of first training the classifier with a known entity and then testing the trained classifier with an unknown data set (new patient). This procedure ensures that there is no overlap in the data source distribution, testing the robustness of the system.

The paper is organized as follows. Section 2 describes in detail our measurement system, procedure and the breast phantoms. Section 3 explains the classification algorithms and the testing methodology. Finally, Section 4 presents the results, emphasizing the rates of successful tumor detection and a comparison between SVM and LDA.

2. MEASUREMENT DESCRIPTION

This section details the breast cancer detection system, the phantoms that are used for testing, and the measurement procedure.

2.1. System Set-up

Our time-domain system has been described previously in [10]. The main component of the system is an antenna array held on the exterior of a hollow, hemi-spherical radome. The array is composed of 16 ultrawideband elements; the antennas were specifically designed for biological sensing applications and are meant to operate in a region with dielectric properties similar to those of breast tissues. Description and characteristics of the Travelling-Wave Tapered and Loaded Transmission-Line Antennas (TWTLTLA) can be found in [22]. The breast phantom under test is positioned inside of the bowl-shaped radome, and in this series of experiments, fills the radome completely.

The measurement system operates as follows. An impulse generator is triggered by a clock operating at 25 MHz; this pulse is then shaped using a Synthesized Broadband Reflector (SBR, as in [23]) such that its frequency content is concentrated from 2–4 GHz. The pulse is then amplified (+35 dB gain), passes through an automated 16×2 switching matrix that selects each transmitting antenna in turn, and propagates into the breast. The scattered response is picked up by the selected receive antenna and recorded by a sampling oscilloscope using an equivalent-time sampling rate of 80 GSa/s (also triggered by the 25 MHz clock). For each breast scan, we record a total of 240 bistatic signals. A photograph of the experimental system, including the radome and antennas, is shown in Figure 1.

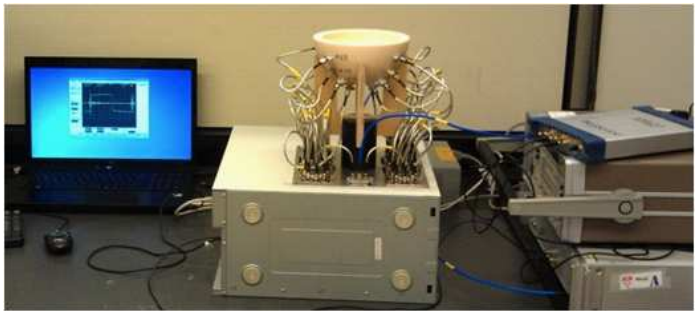

Figure 1. A photograph of the experimental system. The radome is the hemispherical bowl in the center of the image, with the antennas attached to its exterior surface. Also visible is the clock, pulse generator, switching matrix, oscilloscope, and the recording computer.

2.2. Breast Phantoms

2.2.1. General Phantom Description

We use tissue phantoms that are designed to have dielectric properties matched to those of the actual tissues. The procedure for making the phantoms can be found in [24]; they are based on the actual tissue measurements reported in [25], and the chemical combinations proposed in [26]. In particular, we use homogeneous phantoms that mimic adipose tissue. However, measurements have shown that, despite following the same recipe and fabrication procedure, there is significant variation in the dielectric properties of the tissue phantoms. The relative permittivity and conductivity can vary from phantom to phantom, as well as within a singular phantom.

To confirm the level of variation within the tissue phantoms, we made multiple fat and tumor phantoms, ensuring consistency in mixing procedures. We then used a dielectric probe (Agilent HP85070E) connected to a microwave network analyzer (Agilent N5242A PNA-X) to record the dielectric properties. The properties were recorded everyday over the span of a week at four locations on each of four phantoms. Figures 2 and 3 show the measurement results for the healthy and tumor phantoms, respectively. The plots show the average relative permittivity, ε_r, over all tissue phantoms of that type, along with the maximum and minimum recorded values. From Figure 2, it is seen that, over our 2–4 GHz range of interest, the relative permittivity of the healthy phantom varies from approximately 5.5 to 19, with an average of 11.5 (at 3 GHz), while the conductivity varies from 0.06 S/m to 1.2 S/m with an average value of 0.45 S/m. The relative permittivity for the tumor phantom, over the same frequency range, varies from 41 to 73 with an average of almost 50, as shown in Figure 3. The measured conductivity of the tumor phantom has an average value of 2.1 S/m, a maximum of 4.5 S/m, and a minimum of 0.94 S/m.

These results demonstrate that, even if each breast phantom is composed of 100% fat-mimicking material, the properties of this fat phantom can vary widely both between samples and within an individual sample. This adds to the complexity of experimental testing, and to the challenge of applying a signal classification algorithm.

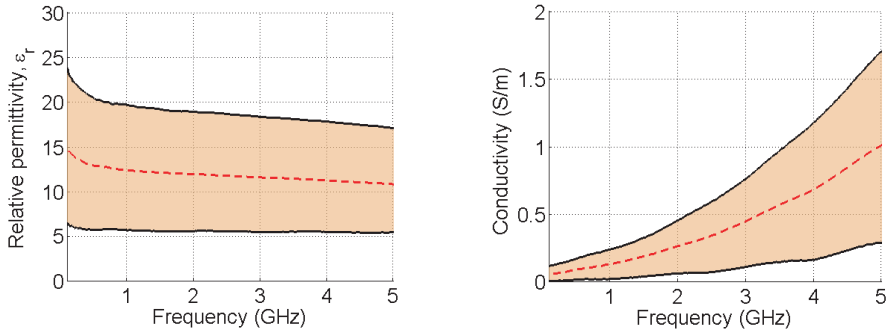

Figure 2. Plot of the measured variation in relative (a) permittivity and (b) conductivity for the healthy tissue phantoms. The red dashed line in the middle is the average while the black solid outer bounds of the shaded region represent the maximum and minimum measured values.

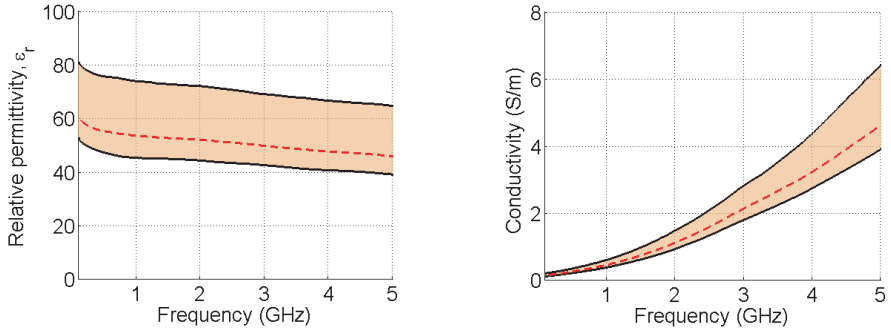

Figure 3. Plot of the measured variation in relative (a) permittivity and (b) conductivity for the tumor phantom. The red dashed line in the middle is the average while the black solid outer bounds of the shaded region represent the maximum and minimum measured values.

2.2.2. Phantom Test Scenario

The breast phantoms used in this study are composed of adipose-mimicking tissue that fills the radome entirely. Tumor phantoms are made separately, then are manually carved into the desired shape and size. In particular, we use tumors that are approximately spherical with 1 cm radius. The carved tumors are embedded into a healthy breast phantom at the chosen location to create the tumorous breast phantom. In this study, two tumor locations are tested, one centred at a depth of about 2 cm positioned on one side of the phantom ("Position A"), and the other centred at a depth of about 3 cm on the other side of the phantom ("Position B"). An illustration depicting the two possible tumor locations within the breast phantom is shown in Figure 4.

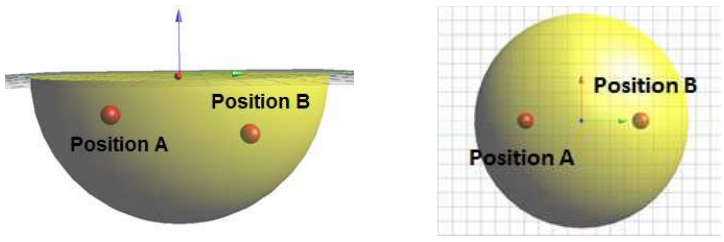

Figure 4. An illustration showing the breast phantom (yellow), and the two possible tumor locations, denoted by the red spherical tumors. A (a) side and (b) top view is shown.

2.3. Measurement Procedure

Measurements were performed on both "healthy" and "tumorous" breast phantoms. As discussed in Section 2.2.2, two types of phantoms with malignancies were fabricated, each with a unique tumor location. For "healthy" breast measurements, no tumor is placed into the fat phantom.

In this study, a total of 230 breast scans from 15 different breast phantoms were performed over a period of eight weeks; of which 200 breast scans, from 10 breast phantoms, were used to train the classification algorithms. Scans were performed daily for each phantom over a week-long period to collect the training and test data. Each breast scan consists of 240 signals, recorded with a sampling period of 12.5 ps. The training data set consisted of an equal number (100 scans) of healthy and tumorous breast scans. Additionally, the tumorous breast scans were equally divided (50 scans) into recordings from both tumor positions. The 30 test data sets were obtained from a completely different set of five newly fabricated phantoms, with an equal number of recordings, 10 scans each, from healthy phantoms, tumor Position A, and tumor Position B.

3. METHODOLOGY AND DETECTION ALGORITHM

In this section, we describe the steps necessary to successfully implement the proposed classification method for breast cancer detection with our microwave time-domain system. The experimental data must first be processed to compensate for the effects of measurement uncertainty that arise from clock and trigger jitter. We then apply a feature extraction method, Principal Component Analysis (PCA), to the entire data set. The SVM and LDA classifiers are applied to selected principal components, and are optimized using a cross-validation method. The following subsections describe these steps in detail. The procedure is illustrated in Figure 5 with a flow-chart block diagram.

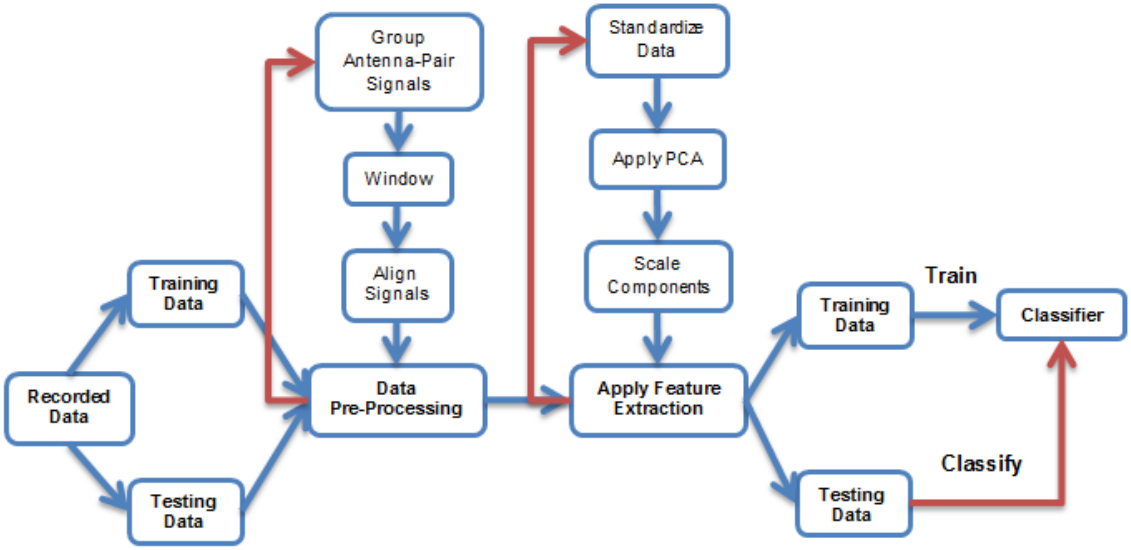

Figure 5. A flow-chart showing the steps followed to implement the detection algorithm.

3.1. Data Pre-processing

The first step in the data pre-processing algorithm is to group similar signals and to window these signals in the region of interest. We sort signals recorded by the same antenna pairs (i.e., we group all the signals recorded with Antenna X as transmitter and Antenna Y as receiver, for all possible antenna pairs) for all of the breast scans. Thus, we create 240 separate data matrices to represent the recorded data for each antenna pair. This procedure allows us to investigate the effects of including or removing specific antenna pairs on the accuracy of the classification algorithms. The data recorded from the

oscilloscope is a length of 4096 samples. We window the data, based on the longest possible path for a wave to travel within the radome, to only include the 400 samples (first 5 ns) immediately after the start of the direct pulse. Thus, the entire training data set is represented by 240 matrices with dimension of 200×400. Each of these matrices is referred to hereafter as $\tilde{\mathbf{X}}_i$, where $i = 1, \ldots, 240$.

The second step of the pre-processing algorithm is to mitigate sources of measurement uncertainties. The two primary sources of measurement uncertainty in our experimental system are due to antenna and jitter effects. These effects can cause the backscattered electromagnetic (EM) energy recorded by the antennas to be buried within the noise of the system. Additionally, the horizontal noise associated with jitter introduces a large amount of phase variability in the recorded signals; in order to implement a machine learning algorithm to classify signals recorded from similar phantoms (i.e., be able to decide if the signals recorded are from a healthy or tumorous phantom), we ideally want to minimize the variability between signals of the same class and maximize the variability between classes. By performing the data pre-processing, we remove the variability associated with measurement uncertainty. Here, we present a summary of the steps followed to mitigate these effects (a complete analysis can be found in [27]) and to prepare the data for the classification algorithms.

The primary concern with the antenna effects is due to slight positional differences between the antennas in different scans; these changes in position can alter the phase and amplitude of the recorded signals. This concern is addressed by solidly fixing the antenna locations within the radome [27].

Jitter, inconsistent deviations in the timing of signals, is a source of horizontal noise that arises in both the reference clock signal and the oscilloscope trigger signal. These sources of jitter cause the misalignment of subsequent time-domain recordings; the signals recorded from various breast scans, will be out of phase with respect to each other. As in [27], we mitigate the effects of jitter by implementing a correlation alignment procedure; signals are shifted by some ΔT that maximizes the cross-correlation between the signals.

In Figure 6 we show the result of the data pre-processing; we compare a plot (only showing a relevant sub-section) of all the collected signals, from both the healthy and tumorous phantoms, for a specific antenna pair (Antenna 1 transmits, Antenna 2 receives), before and after performing the aforementioned noise mitigation procedure. From this comparison it is clear that the phase difference and the horizontal noise between recordings has been greatly reduced; we observe that the signals are significantly better aligned after the data pre-processing. The remaining variation between the signals is due to the differences between phantoms, the inclusion of the tumor, and the difference in the tumor locations.

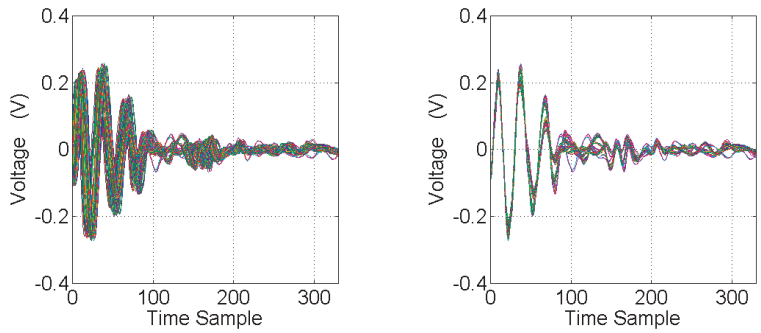

Figure 6. Received signals (a) before and (b) after noise mitigation for a sample antenna pair (Antenna 1 is the transmitter, Antenna 2 receives).

3.2. Feature Extraction

The method proposed in this paper follows a similar procedure to that in [14, 17, 18]; we apply PCA to the dataset and then apply the classification algorithms to the extracted features. Applying feature extraction to the recorded data allows us to reduce the dimensionality of the data, thus we can minimize

the processing of the input data by extracting only the essential components, and using these components for analysis with the classifier. In [28] it was shown that PCA was the best feature extraction method on average; additionally, when combined with the SVM classifier, it yielded the best results compared to the combination of other classifier and feature extraction methods. Thus, we adopt a similar methodology in this paper.

Prior to feature extraction the data set is first standardized on a signal-to-signal basis; the mean of the standardized data set is 0 and the standard deviation is 1. The standardized data set can be defined as:

$$\mathbf{X}_i^m = \frac{\tilde{\mathbf{X}}_i^m - \mu_i^m}{\sigma_i^m}, \quad m = 1, \ldots, 200, \tag{1}$$

where $\mathbf{X}_i$ and $\tilde{\mathbf{X}}_i$ are the standardized data set and the original data set, respectively, for a specific antenna pair i, where $i = 1, \ldots, 240$, m is a specific row of the matrices, and μ_i^m and σ_i^m are the mean and the standard deviation, respectively, of a specific row m of matrix $\tilde{\mathbf{X}}_i$. Thus, resulting standardized data sets $\mathbf{X}_i$ are the same size as $\tilde{\mathbf{X}}_i$.

PCA is then applied to the standardized data set, $\mathbf{X}_i$, in order to extract the most important features. The extracted data feature set is stored in $\mathbf{x}_i$, with i defined as above. The use of PCA reduces the dimensionality of the data by performing an orthogonal linear transformation to express the data in a new vector space such that each component that makes up this new coordinate system represents a diminishing amount of variance of the data (the projection of the data onto the first principle component represents the most variance), thus not all principle components need to be kept. This procedure diminishes the influence of the inherent noise of the experimental system. In Section 3.4, we discuss the procedure followed to determine the ideal number of principle components for use in this investigation; at this point only the first 50 components are saved. The training data is now represented by 240 matrices, $\mathbf{x}_i$, with dimension of 200×50.

3.3. Classifier Implementation

Previous applications of classification algorithms for breast cancer detection with a microwave system have demonstrated that the SVM method outperforms other classification methods [17, 18]. While these studies were focused on data obtained numerically, we attempt to confirm that these findings are relevant for an experimental system by applying the SVM classifier to our data. The LDA classifier is used as a benchmark; from [17, 18, 28], it is expected that the use of an optimized SVM classifier will improve the detection accuracy of the proposed method. The classifiers are trained with the $\mathbf{x}_i$ matrices found from feature extraction, such that classification is performed for each transmit-receive antenna pair independently (total 240 pairs).

3.3.1. Support Vector Machines

SVM is a machine learning algorithm that can be used when the data to be classified can be divided into exactly two groups (classes). In order to properly train the SVM classifier it is necessary to provide sample-label pairs for the training data (the classifier must know the specific group that the training data belongs to). In this study the two groups correspond to the healthy and tumorous breast phantoms. The SVM classifier will first map the training data to a higher dimensional space in order to determine the support vectors. A hyperplane that correctly separates the two classes is then created and the algorithm searches for the best hyperplane. The best hyperplane is one that maximizes the distance (margin) between the hyperplane and the nearest data points of each class (support vectors) [29]. A hard margin represents a slab region that is parallel to the hyperplane, the interior of which does not contain any of the training data points. In certain cases a separating hyperplane cannot be found, this can be due to the presence of outliers in the data. A soft margin, which separates the majority of the training data but not all, can be used in this case [30]. In this paper, we implement a soft-margin SVM. Additionally, it is possible to employ a more complex mathematical formulation to create a non-linear hyperplane to more accurately separate the training data. More details about the mathematical formulations and expressions governing the computation of the support vectors and the classification algorithm can be found in [29, 30].

3.3.2. Linear Discriminant Analysis

In this study we apply a LDA classifier as a means for comparison with the performance of the SVM classifier. It is a method that is simple to implement and to apply to our application. The use of LDA as a benchmark for studying the effectiveness and performance of SVM was shown in [17, 18]. LDA is similar to PCA in that it also tries to find linear combinations of variables that best represent the data [31]; however, LDA directly characterizes the difference between the various data groups. Specifically, an LDA classifier is ideally suited to situations in which the groups to be classified have multivariate normal distributions [17, 18], and involves the calculation of an in-group covariance for each subset of test data which is used to first obtain the discriminant function and second to perform the classification.

3.4. Classifier Optimization

A procedure to optimize classifier performance was proposed in [32]. Based on this procedure, we adopt a similar method to optimize our SVM classifier, summarized as follows:

(i) Choose the appropriate SVM kernel function.
(ii) Scale the training data to ensure that it is in the $[-1, +1]$ range. Apply the same scaling factor to the test data set.
(iii) Use cross-validation to find the optimal number of principle components, and ideal SVM parameters (C, the box constraint, and γ, kernel scaling factor).
(iv) Train the SVM classifier with the optimal settings solved for using cross-validation.

As suggested in [32], we implement a Radial Basis Function (RBF) kernel to non-linearly map the data to some higher dimension vector space. Unlike the linear (base) kernel function, the RBF kernel is well suited to handle classification problems where a simple separating hyperplane, defined by a linear function in the vector space, cannot be used as a useful separating criterion. The RBF kernel allows for the discriminant hyperplane to be defined by a non-linear function, thus allowing for a more robust and precisely defined separating hyperplane.

Once we have chosen the appropriate kernel function for the SVM classifier we linearly scale the training data set to the $[-1, +1]$ range before training the classifier; this process ensures that training data in higher numeric ranges do not dominate the classifier [32]. A linear scaling is applied to each principle component to determine the necessary scaling factor; this factor is then applied to the test data.

In order to optimize the performance of the SVM classifier we aim to identify the optimal number of principle components, n, to use in our analysis (previous work, [17, 18], has shown promising results with between 30 and 50 components), and to identify a good C and γ pair; this can be achieved by training and testing the classifier using only the training data set (known recorded data). The aim of this optimization is to find a (C, γ) pair that improves the classifier performance when predicting unknown (test) data. To accurately reflect this scenario, we implement 10-fold cross-validation to search for n, and (C, γ). In 10-fold cross-validation the training data set is equally divided into 10 randomly chosen subsets. Each subset is then sequentially tested with a classifier trained with the 9 remaining data subsets; thus, each subset of the training data is tested exactly once. The cross-validation accuracy is the mean of the correctly classified data across the 10 tests.

As suggested in [32], a coarse-then-fine grid search is used to determine the optimal (C, γ) values. The results for the coarse- and fine-grid search are shown in Figures 7(a) and 7(b), respectively, as a contour plot. We compare the mean classifier accuracy from the 10-fold cross-validation, denoted by the color intensity, with variation of exponentially incremental (C, γ) values. The fine-grid search was conducted in the region within the most accurate region of the coarse grid search.

A simple linear search, ranging from 1 to 50, was used to determine the number, n, of principle components to use for the SVM classifier; additionally, a similar search was used to find the optimal value of n for the LDA classifier. In both cases, a 10-fold cross-validation was used to find n. Results of this search are shown in Figure 8; we plot a comparison of the mean cross-validation accuracy for the LDA and SVM classifiers as the number of principle components is varied. Additionally, we compare the computational cost for both the LDA and SVM classifiers as the number of principle

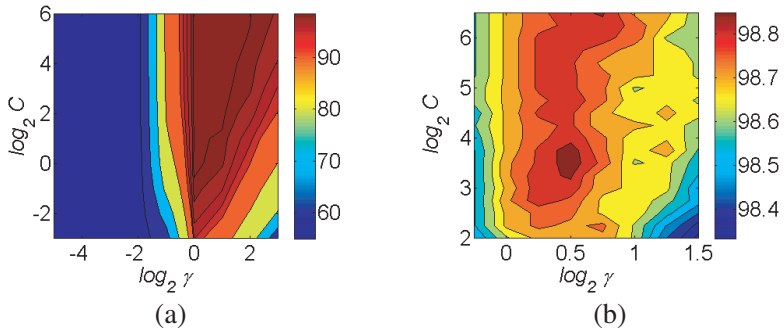

Figure 7. Results of grid search to find the optimal C and γ. (a) A coarse grid search and (b) a fine grid search are shown. The color intensity represents the mean classifier accuracy for 10-fold cross-validation.

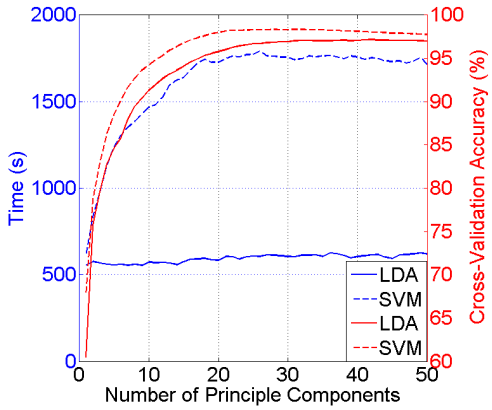

Figure 8. Comparison of computational cost (blue, left-hand axis) and accuracy (red, right-hand axis) versus the number of principle components for SVM (dashed) and LDA (solid).

components is increased.We observe that the cross-validation accuracy and the computational cost for the SVM classifier follow similar trends; additionally, there is minimal change in both accuracy and computation time for $n > 20$. For the LDA classifier the computational cost is relatively unchanged as the number of principle components is increased. Since the computational cost for both classifiers is relatively unchanged when n is increased, we can choose the optimal number of principle components based solely on the cross-validation accuracy without concern for increasing computational costs. The optimal number of principle components, based on the maximum cross-validation accuracy, is $n = 31$ and $n = 42$ for the SVM and LDA classifiers, respectively. It is clear that the SVM classifier is more computationally expensive than the LDA classifier; however, it also outperforms the LDA classifier.

4. RESULTS

The final classifier is trained with the optimal values of (n, C, and γ) found from the optimization procedure listed in Section 3.4. The trained classifier is then tested with signals collected from newly fabricated phantoms, thus ensuring that there is minimal correlation between the training and testing data sets; furthermore, this serves to emulate a clinical setting in which the classifier is trained with known data sets and then tested with new patients of unknown health status. As stated in Section 2.3, the test data set is made up of signals recorded from 30 new breast scans.

Due to the large antenna beamwidth [22], the low attenuation in the homogeneous phantoms, and the multi-path reflections that waves undergo while travelling through the radome, phantom, and back through the radome again, we expect the presence of a tumor (if there is one) to be identifiable above the noise level in signals from all 240 transmit-receive antenna pairs. Thus, we assume here that

correct classification implies that each individual signal recorded from healthy and tumor phantoms, respectively, should be classified as such.

We compare the performance of the optimized SVM and LDA classifiers in Table 1. We examine the overall detection performance as well as the percentage of correctly classified signals for each of the three phantom types (healthy, tumor position A, tumor position B). We confirm that the SVM algorithm improves the detection accuracy (defined as the percentage of correctly classified signals) of the system when compared with the LDA method; however, the difference is not as pronounced as was seen in [18]. Additionally, the prediction accuracy of the LDA classifier, here, outperforms the one presented in [18], perhaps suggesting that the LDA classifier, and not just the SVM classifier, is also suited for experimental analysis. While a direct comparison with the results presented in [18] is impossible, as we do not recreate the scenario exactly (different tumor sizes, shapes, material properties), we observe that the overall detection accuracy for both classifiers is decreased in our experimental scenario. There are several contributing reasons for this decrease, many of which are typical of experimental measurements; (i) the dielectric properties of both the healthy and tumor phantoms vary over a large range of values, (ii) we test, as in a real-life scenario, on a phantom with dielectric properties that the classifier has not been trained with, and, (iii) there are significant sources of noise that affect the signals, which are mitigated here but not eliminated. All of these factors make classification with a realistic experimental scenario more complex.

Table 1. Detection accuracy of each classification algorithm. We compare the overall detection accuracy, as well as the accuracy for the three phantom types tested.

	Healthy	**Tumor**		**Overall**
		Position A	**Position B**	
LDA (%)	70.00	70.45		70.30
		62.90	78.00	
SVM (%)	67.48	76.71		73.64
		70.48	82.95	

The overall detection accuracy for the LDA and SVM classifiers, respectively, is 70.30% and 73.64%. It is interesting to note that the while the LDA classifier slightly outperforms the SVM classifier in correctly identifying signals from healthy breast phantoms, with a detection accuracy of 70.00% for the LDA classifier compared to 67.48% for the SVM classifier, the SVM classifier is much more successful at identifying tumors, correctly identifying signals from the tumorous phantoms with an accuracy of 76.71% as opposed to only 70.45% for LDA. We also notice that regardless of the choice of classifier, the system is more adept at classifying signals recorded with the tumor embedded in Position B (an increase of over 15% and 12% for the LDA and SVM classifiers, respectively).

For the results described above, the LDA and SVM classifiers work on a signal-to-signal basis; each of the 240 signals making up a test data set is classified as being a signal derived from a healthy or tumorous phantom. Now, instead, we select the most relevant antenna clusters to apply our classification to. As was shown in our previous work [10, 23], transmit-receive antenna pairs that are located on the same side of the radome tend to better pick up the scattered tumor response. Thus, these clusters of antennas can be used for more reliable detection, particularly when the phantom is heterogeneous (when attenuation is higher, and tumor responses from far away transmit-receive antenna pairs may be lost below the noise level). The antenna array is composed of four antennas lined up in each of four quadrants of the radome; therefore, the four antennas of each quadrant make up an antenna cluster, for a total of four antenna clusters.

In Table 2 we compare the detection accuracy for both the LDA and SVM classifiers, for each phantom type, when using only signals from the four antenna clusters. Using these selected signals significantly improves the tumor detection accuracy, from 70.48% to 83.50% for a tumor in Position A, and from 82.95% to 91.75% for a tumor in Position B, for the SVM classifier. For LDA, an improvement from 62.90 % to 71.04% and from 78.00% to 83.75% are observed for Positions A and B, respectively.

These values suggest that further investigation into the effect of antenna groupings on detection may be worthwhile to improve detection in scenarios with complex tissue composition. In particular, we hypothesize that transmit-receive antenna pairs that are in the quadrant closest to the tumor will be the most useful for confirming the presence of the tumor.

As shown in Table 1 and Table 2, both classifiers perform poorly when attempting to classify signals from new healthy breast phantoms. The variation in the conductivity of the healthy phantom (Figure 2) can lead to very large changes in signal attenuation (amplitude and phase shift), due to the long paths travelled. This means that there is a large variation in the collected baseline signals. The tumor, although it also has a large variation of dielectric properties, is small relative to the rest of the breast tissue and thus it does not have as much of an impact on the overall signal. In particular, the antenna clusters, which were shown to best pick up on differences in phantom properties, suffer the most when the healthy tissue properties vary so significantly (as shown in Table 2). We suspect that increasing the number of signals (and phantoms) used to train the classifiers can account for these variations and would improve classifier performance.

The results from Table 2 suggest that very few tumors will be misclassified at the cost of many misclassified healthy tissues; this situation likely does not represent the best trade-off between false-positive and false-negative rates. Thus, future work will include a limit on the number of acceptable false alarms permissible in the classification algorithm, as shown in [14]. The changes seen in Table 2 towards better tumor detection, but inferior healthy detection as compared to Table 1, also suggests that taking antenna clusters shifts the false alarm rate, potentially due to a decrease in the overall number of signals used to train the classifiers. Our near-future investigations will question whether optimizing the SVM classifier with only this subset of antenna pairs can decrease the false-alarm rate and maintain the improved detection accuracy.

Table 2. The detection accuracy for SVM and LDA when only signals from the four antenna clusters are used.

	Healthy	**Tumor**	
		Position A	**Position B**
LDA (%)	54.79	77.40	
		71.04	83.75
SVM (%)	57.00	87.63	
		83.50	91.75

5. CONCLUSION

In this paper we implement machine learning classification algorithms to detect the presence of breast cancer with a time-domain microwave radar system. Specifically, we demonstrate that both the LDA and SVM classifiers can be used to detect the presence of tumors with experimental data obtained from dielectrically realistic breast phantoms. The classifiers are trained with data obtained from a set of phantoms and tested with data obtained from new phantoms, representing new patients. Detection accuracy is assessed by correct classification of the test signals. We demonstrate that the SVM classifier outperforms the LDA classifier, based on overall detection accuracy. This confirms the trend seen in past numerical studies in the literature. Additionally, we briefly examine the role that antenna clusters play in improving detection accuracy.

Our future work will focus on using classification algorithms to monitor changes within the breast. Recent literature has suggested that using differential scans for monitoring can greatly improve the detection accuracy. We will apply similar analysis using heterogeneous breast phantoms, where both the glandular content and the tumor development can change over time. We plan to implement a classification algorithm for successful differentiation between healthy and unhealthy changes within the breast tissue.

ACKNOWLEDGMENT

This work was supported by the Natural Sciences and Engineering Research Council of Canada (NSERC), le Fonds québécois de la recherche sur la nature et les technologies (FQRNT), and Partenariat de Recherche Orientée en Microélectronique, Photonique et Télécommunications (PROMPT). The authors are extremely grateful to the McGill University Photonics Systems Group for allowing us use of their lab space and the long-term loan of key measurement components. We would like to thank Professor Mark Coates for assistance with fine-tuning and documenting the classification algorithms.

REFERENCES

1. Nikolova, N., "Microwave imaging for breast cancer," *IEEE Microwave Magazine*, 78–94, Dec. 2011.
2. Fhager, A., S. K. Padhi, and J. Howard, "3D image reconstruction in microwave tomography using an efficient FDTD model," *IEEE Antennas Wireless Propag. Lett.*, Vol. 8, 1353–1356, 2009.
3. Guardiola, M., S. Capdevila, J. Romeu, and L. Jofre, "3-D microwave magnitude combined tomography for breast cancer detection using realistic breast models," *IEEE Antennas Wireless Propag. Lett.*, Vol. 11, 1622–1625, 2012.
4. Meaney, P. M., et al., "Microwave tomography in the context of complex breast cancer imaging," *Proceedings of the 32nd Annual International Conference of IEEE EMBS*, 3398–3401, Buenos Aires, Argentina, Aug. 31–Sept. 4, 2010.
5. Klemm, M., I. J. Craddock, J. A. Leendertz, A. Preece, and R. Benjamin, "Radar-based breast cancer detection using a hemispherical antenna array — Experimental results," *IEEE Trans. Ant. Propag.*, Vol. 57, No. 6, Jun. 2009.
6. Byrne, D., M. O'Halloran, M. Glavin, and E. Jones, "Breast cancer detection based on differential ultrawideband microwave radar," *Progress In Electromagnetics Research M*, Vol. 20, 231–242, 2011.
7. Lai, J. C. Y., C. B. Soh, E. Gunawan, and K. S. Low, "UWB microwave imaging for breast cancer detection — Experimentals with heterogeneous breast phantoms," *Progress In Electromagnetics Research M*, Vol. 16, 19–29, 2011.
8. Flores-Tapia, D. and S. Pistorius, "Real time breast microwave radar image reconstruction using circular holography: A study of experimental feasibility," *Med. Phys.*, Vol. 38, No. 10, 5420–5431, Oct. 2011.
9. Zeng, X., A. Fhager, P. Linner, M. Persson, and H. Zirath, "Accuracy investigation of an ultra-wideband time domain microwave imaging system," *Proceedings of the 5th European Conference on Antennas and Propagation (EUCAP)*, 1928–1932, Rome, Italy, Apr. 11–15, 2011.
10. Porter, E., E. Kirshin, A. Santorelli, M. Coates, and M. Popović, "Time-domain multistatic radar system for microwave breast screening," *IEEE Antennas Wireless Propag. Lett.*, Vol. 12, 229–232, 2013.
11. Byrne, D., "Ultrawideband radar for the early detection of cancer within the heterogeneous breast," Ph.D. Diss., 2012.
12. Agarwal, K., L. Pan, Y. K. Leong, M. Han, O. Y. Chan, X. Chen, and S. P. Yeo, "Practical applications of multiple signal classification," *International Journal of RF and Microwave Computer-aided Engineering*, Vol. 22, No. 3, 359–369. 2012.
13. Arnau, O., J. Freixenet, R. Martí, and R. Zwiggelaar, "A comparison of breast tissue classification techniques," *Medical Image Computing and Computer-Assisted Intervention, MICCAI*, 872–879, Springer Berlin Heidelberg, 2006.
14. Prasad, D. P., C. Quek, and M. K. H. Leung, "A hybrid approach for breast tissue data classification." *TENCON 2009-2009 IEEE Region 10 Conference*, 1–4, IEEE, 2009.
15. Davis, S. K., et al., "Breast tumor characterization based on ultrawideband microwave backscatter," *IEEE Transactions on Biomedical Engineering*, Vol. 55, No. 1, 237–246, 2008.
16. Kerhet, A., M. Raffetto, A. Boni, and A. Massa, "A SVM-based approach to microwave breast cancer detection," *Engineering Applications of Artificial Intelligence*, Vol. 19, 807–818, 2006.

17. Conceição, R. C., M. O'Halloran, M. Glavin, and E. Jones, "Support vector machines for the classification of early-stage breast cancer based on radar target signatures," *Progress In Electromagnetics Research B*, Vol. 23, 311–327, 2010.
18. Byrne, D., M. O'Halloran, E. Jones, and M. Glavin, "Support vector machine-based ultrawideband breast cancer detection system," *Journal of Electromagnetic Waves and Applications*, Vol. 25, No. 13, 1807–1816, 2011.
19. Fhager, A., Y. Yu, T. McKelvey, and M. Persson, "Stroke diagnostics with a microwave helmet," *Proceedings of the 7th European Conference on Antennas and Propagation (EUCAP)*, 845–846, Gothenburg, Sweden, Apr. 8–12, 2013.
20. O'Halloran, M., F. Morgan, M. Glavin, E. Jones, R. C. Conceição, and D. Byrne, "Bladder-state monitoring using ultra wideband radar," *Proceedings of the 7th European Conference on Antennas and Propagation (EUCAP)*, 624–627, Gothenburg, Sweden, Apr. 8–12, 2013.
21. Conceicao, R. C., H. Medeiros, M. O'Halloran, D. Rodriguez-Herrera, D. Flores-Tapia, and S. Pistorius, "Initial classification of breast tumour phantoms using a UWB radar prototype," *2013 International Conference on Electromagnetics in Advanced Applications (ICEAA)*, 720–723, IEEE, 2013.
22. Kanj, H., and M. Popović, "A novel ultra-compact broadband antenna for microwave breast tumor detection," *Progress In Electromagnetics Research*, Vol. 86, 169–198, 2008.
23. Santorelli, A., et al., "Experimental demonstration of pulse shaping for time-domain microwave breast imaging," *Progress In Electromagnetics Research*, Vol. 133, 309–329, 2013.
24. Porter, E., J. Fakhoury, R. Oprisor, M. Coates, and M. Popovic, "Improved tissue phantoms for experimental validation of microwave breast cancer detection," *Proceedings of the 4th European Conference on Antennas and Propagation (EUCAP)*, Barcelona, Spain, Apr. 12–16, 2010.
25. Lazebnik, M., et al., "A large-scale study of the ultrawideband microwave dielectric properties of normal, benign and malignant breast tissues obtained from cancer surgeries," *Phys. Med. Biol.*, Vol. 52, 6093–6115, 2007.
26. Lazebnik, M., E. Madsen, G. Frank, and S. Hagness, "Tissue-mimicking phantom materials for narrowband and ultrawideband microwave applications," *Phys. Med. Biol.*, Vol. 50, 4245–4258, 2005.
27. Porter, E., E. Kirshin, A. Santorelli, and M. Popovic, "Microwave breast screening in the time-domain: Identification and compensation of measurement-induced uncertainties," *Progress In Electromagnetics Research B*, Vol. 55, 115–130, 2013.
28. Conceicao, R. C., et al., "Evaluation of features and classifiers for classification of early-stage breast cancer," *Journal of Electromagnetic Waves and Applications*, Vol. 25, No. 1, 1–14, 2011.
29. Boser, B. E., I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," *Proceedings of the Fifth Annual Workshop on Computational Learning Theory*, 144–152, ACM, 1992.
30. Cortes, C. and V. Vapnik, "Support-vector networks," *Machine Learning*, Vol. 20, No. 3, 273–297, 1995.
31. Martinez, A. and A. Kak, "PCA versus LDA," *IEEE Trans. Pattern Anal. Mach. Intell.*, Vol. 23, No. 2, 228–233, Feb. 2001.
32. Hsu, C.-W., C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," Tech. Rep., Department of Computer Science, National Taiwan University, 2003, http://www.csie.ntu.edu.tw/cjlin/papers.html.