Progress In Electromagnetics Research B, Vol. 23, 311–327, 2010

SUPPORT VECTOR MACHINES FOR THE CLASSIFICATION OF EARLY-STAGE BREAST CANCER BASED ON RADAR TARGET SIGNATURES

R. C. Conceição, M. O'Halloran, M. Glavin, and E. Jones †

Electrical and Electronic Engineering
College of Engineering and Informatics
National University of Ireland Galway
Ireland

Abstract—Microwave Imaging (MI) has been widely investigated as a method to detect early stage breast cancer based on the dielectric contrast between normal and cancerous breast tissue at microwave frequencies. Furthermore, classification methods have been developed to differentiate between malignant and benign tumours. To successfully classify tumours using Ultra Wideband (UWB) radar, other features have to be examined other than simply the dielectric contrast between benign and malignant tumours, as contrast alone has been shown to be insuficient. In this context, previous studies have investigated the use of the Radar Target Signature (RTS) of tumours to give valuable information about the size, shape and surface texture. In this study, a novel classification method is examined, using Principal Component Analysis (PCA) to extract the most important tumour features from the RTS. Support Vector Machines (SVM) are then applied to the principal components as a method of classifying these tumours. Finally, several different classification architectures are compared. In this study, the performance of classifiers is tested using a database of 352 tumour models, comprising four different sizes and shapes, using the cross validation method.

1. INTRODUCTION

Microwave Imaging is an appealing method for early-breast cancer detection as it does not involve the use of ionising radiation (as is the

Received 24 June 2010, Accepted 15 July 2010, Scheduled 25 July 2010
Corresponding author: R. C. Conceicao (raquelcruzconceicao@gmail.com).
† All are also with Bioelectronics Research Cluste, NCBES, National University of Ireland Galway, Ireland.

case of X-Ray mammography). MI is based on the dielectric differences between the constituent tissues of the breast at microwave frequencies, as recently thoroughly studied by Lazebnik et al. [1, 2].

Three different approaches have been proposed in order to image the breast based on these contrasting dielectric properties: UWB Radar Imaging, Microwave Tomography and Time-Reversal FDTD methods. The UWB Radar Imaging involves illuminating the breast with a UWB pulse, and recording the resulting backscattered signals. Subsequently, these recorded signals are processed using a beamforrner to identify the presence and location of significant dielectric scatterers within the breast [3–5], which potentially indicates the presence of tumours. Microwave Tomography involves a full reconstruction of the dielectric profile of the breast, by solving a forward and inverse scattering problem. These algorithms seek to minimise the difference between measured and calculated electric fields [6–11]. Time-Reversal Finite-Difference Time Domain (FDTD) algorithm involves applying time-reversed FDTD equations to all points of the breast grid, with the wave converging at the tumour point [12–14].

Although, historically, many MI studies have looked at methods to help detect and locate a tumour within the breast, only a few studies have looked at the classification of the detected suspicious masses. Chen et al. [15, 16] studied the effect of the morphology of different tumours on the late-time response of backscattered signals obtained from an FDTD breast model. In these studies, 2D models of benign and malignant tumours were created using an approximation to irregular polygons, with an eleptical baseline, which had been used by Rangayyan et al. [17] to circumscribe different breast tumours in X-Ray Mammographies. Chen also developed an algorithm that allows for the location of a tumour and indicates, by means of a correlation, whether the lesion boundary is more or less irregular, possibly indicating whether a tumor is benign or malignant. More recently, in [18], Chen et al. analysed the issue of classification after applying a contrast-agent to lesions in order to increase the contrast between cancerous and normal breast tissue.

In Davis et al. [19] tumours are modelled in a 3D Total-Field/Scattered-Field (TF/SF) FDTD model, so that only the response from the tumour is recorded in the backscattered signals and can be directly analysed, without the need for any more hardware or software than is already used for UWB imaging. In this study the tumours were modelled using Gaussian Random Spheres, first introduced by Muinonen [20, 21], to represent different shapes and sizes of tumours. A full classification algorithm composed of two methods for basis selection followed by a Linear Discriminant Analysis (LDA)

based classifier was applied to a database with targets ranging between four sizes and three shapes.

Previous work by the authors [22] extended the work by Davis et al. [19] by investigatings and evaluating various different classification architectures, adding extra granularity to the tumour classification process by introducing the macrolobulated shape to benign tumours, and introducing the Quadratic Discriminant Analysis (QDA) as an alternative classification method.

This paper has two significant contributions:

- The investigation of Support Vector Machines (SVM) to classify tumours based on their RTS;
- comparison of SVM with existing classification methods, such as LDA and QDA [22].

The remainder of the paper is organised as follows: Section 2 describes the experimental setup, including the numerical tumour models and FDTD simulation; Section 3 presents the feature extraction method used (PCA); Section 4 describes the SVM classification algorithm and the different architectures of classifiers; in Section 5 the results and discussion are presented; finally in Section 6 the conclusions are drawn and future work is suggested. It must be noted that some subjects in Sections 2 and 3 have already been described in [22] but they are included in this paper for completeness.

2. EXPERIMENTAL SET

This section is divided into two subsections: in Section 2.1 the models used for representing the different tumours are detailed and in Section 2.2 describes the FDTD model in which the tumours are placed.

2.1. Tumour Models

For the purpose of classification, the significant features defining the tumour are: size, shape and texture of surface, since these will most significantly influence the RTS of tumours. Benign tumours typically have smooth surfaces and have spherical, oval or at least well-circumscribed contours. Conversely, malignant tumours usually present rough and complex surfaces with spicules or microlobules, and their shapes are typically irregular, ill-defined and asymmetric. In this study, the main concern is the shape classification of early-stage tumours, up to 1 cm in radius, and therefore shape and texture of the surface are the two most important features to help differentiate between a benign and a malignant tumour, whereas size may be

crucial when looking at the development of a tumour over a period of time and may indicate malignancy when the tumour reaches a higher volume [17, 23–25].

In this study, the tumours are modelled using the Gaussian Random Spheres (GRS) method [19–21]. GRS can be created in different sizes and shapes by mathematically modifying the mean radius (α) and the covariance function of the logarithmic radius (or simply logradius). The shape, $r = r(\vartheta, \varphi)$, is described in spherical coordinates (r, ϑ, φ) by the spherical harmonics series for the logradius $s = s(\vartheta, \varphi)$:

$$r(\vartheta, \varphi) = \alpha \exp\left[s(\vartheta, \varphi) - \frac{1}{2}\beta^2\right] \tag{1}$$

$$s(\vartheta, \varphi) = \sum_{l=0}^{\infty} \sum_{m=-l}^{l} s_{lm} Y_{lm}(\vartheta, \varphi) \tag{2}$$

In the equations above, β is the standard deviation of the logradius, s_{lm} are the spherical harmonics coefficients and Y_{lm} are the orthonormal spherical harmonics.

The database of tumours used in this study includes four different models of tumours in four different sizes, similar to [22]. The four different types of models of tumours are spiculated, microlobulated, macrolobulated and smooth GRS, in which the first two represent malignant tumours, and the others represent benign tumours. Microlobulated, macrolobulated and smooth GRS are obtained by varying the correlation angle from low to high. Furthermore, spiculated GRS are obtained by adding 3, 5 or 10 spicules to smooth GRS. The average radius of all types of spheres take discrete values: 2.5, 5, 7.5 or 10 mm [22]. Among all sizes and shapes, the number of tumour models developed was 352. A sample of each of the four shapes of the GRS, with a radius of 5 mm, is shown in Figure 1, which was previously reproduced in [22].

2.2. FDTD Simulation

The breast and the tumour are modelled using a 3D Finite-Difference Time-Domain (FDTD) model in which the dielectric properties of the different types of tissue are incorporated. The FDTD model has a 0.5 mm cubic grid resolution. The backscattered signals are generated through a Total-Field/Scattered-Field (TF/SF) region, in which the tumours are completely embedded in the Total Field (TF) [19, 22].

The TF/SF region has the following dimensions: the Scattered Field (SF) is a square geometric prism with square bases measuring 153.5 mm on the side and the height measuring 137.5 mm. The TF

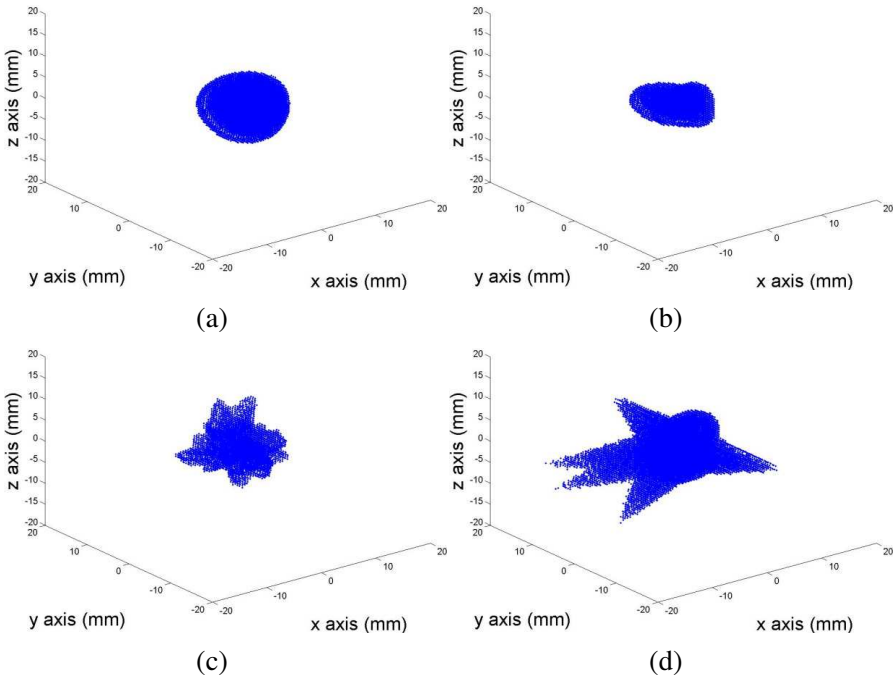

Figure 1. Samples of different Gaussian Random Spheres. From left to right, top to bottom: (a) smooth, (b) macrolobulated, (c) microlobulated and (d) spiculated (5 spicules) models, with an average radius size of 5 mm.

is located at the centre of the SF and is represented by a cube with 50 mm on each side. The origins of the SF and the TF are co-located at $(0, 0, 0)$ mm. The TF/SF region is terminated with a 6 mm Uniaxial Perfectly Matched Layer (UPML) which suppresses any boundary reflections [19, 22, 26].

For all simulations, the whole breast region is modelled with Debye parameters for homogeneous lossy adipose tissue, whereas the tumour areas are modelled with Debye parameters for malignant tissue, as established by Lazebnik et al. [1, 2]. The Debye parameters for malignant tissue are as follows: $\epsilon_\infty = 6.749, \Delta\epsilon = 50.09, \sigma_s = 0.794\,\mathrm{Sm}^{-1}$ and $\tau = 10.50$ ps; whereas for homogeneous lossy adipose tissue they are as follows: $\epsilon_\infty = 3.140, \Delta\epsilon = 1.708, \sigma_s = 0.036\,\mathrm{Sm}^{-1}$ and $\tau = 14.65$ ps.

A pulsed plane wave is transmitted towards the target from four equidistant points, hitting the target at four different angles. The location of the four observation points are at: $(0, 0, -74)$, $(-74, 0, 0)$,

$(0, 0, 74)$ and $(74, 0, 0)$ mm, in (x, y, z) axes. The resulting co-polarised backscatter signals are recorded and analysed from the same four observation points. It must be noted that only co-polarised backscatter is recorded as this was previously found to provide sufficiently high classification performance without the extra computational cost that would involve analysing full polarimetric backscatter [19]. The incident pulse is a modulated Gaussian pulse with centre frequency at 6 GHz and a $1/e$ full temporal width of 160 ps [19]. For the first observer location, $(0, 0, -74)$, the pulse is linearly polarised in the y and x direction and transmitted in the z direction, for the remaining observer locations the pulse is polarised and transmitted accordingly. Each observation point is located in the SF all around the tumour at a constant distance of 74 mm from the centre of the tumour, which is located at the centre of the TF/SF region. The four acquired backscattered recorded signals are then downsampled from 1200 GHz to 75 GHz.

Figure 2, adapted from [22], shows a representation of the TF/SF grid, with the location of the origin of the first incident plane wave and respective observer point (**X**) as well as the position of the tumour. All

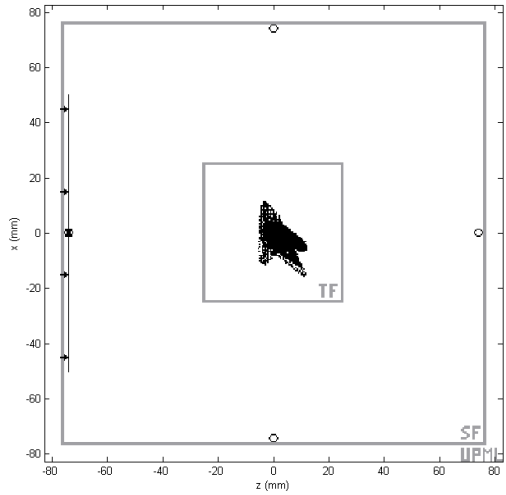

Figure 2. Cross-section of the 3D FDTD space lattice partitioned into Total Field (TF), Scattered Field (SF) and UPML regions. The target, a spiculated tumour located at the centre of the TF, is illuminated by a pulsed plane wave (the arrows indicate the direction of the plane wave and the line in front of the arrows represents its origin) propagating in the $+z$ direction and backscatter is recorded at the first observer location: $(0, 0, -74)$ mm (represented by **X**). All four observation points are represented by small circles in the image.

four observation points are represented by small circles.

3. TUMOUR CLASSIFICATION ALGORITHM

The tumour classification algorithm comprises two steps: the extraction of relevant features from each recorded backscattered signal and the classification method itself. The feature extraction method used is Principal Component Analysis (PCA), similar to [22], which ultimately allows for the extraction of the principal components of the RTS from each tumour. This is followed by a dimensionality reduction in which the more representative Principal Components of the data are selected. Finally, Support Vector Machines (SVM) are used to classify the selected Principal Components. The following two subsections present the two steps of the classification algorithm separately.

3.1. Principal Components Analysis

PCA reduces the dimensionality of multivariate data and reveals simplified structures that are often hidden in the original data set while also disregarding less relevant information such as noise or colinearities in signals [27, 28]. PCA allows for a new representation of the original data in which maximal variance is exposed, so that data may then be better discriminated. Mathematically, the basis that was used to record the original signals is changed, by means of a linear algebraic operation, into a new orthonormal basis that allows for the data to present maximal variance [28]. The resulting principal components are ordered by decreasing variance. It must be noted that PCA is non-parametric so only the data influences the PCA calculation disregarding any prior knowledge on the acquiring system or known labels [19, 28].

A dimensional reduction of the principal components is often completed for computational simplicity, restricting the principal components to a minimum that represents the original data satisfactorily [29].

To obtain the principal components of a matrix $\mathbf{X}$ represented by $(m \times n)$, where m is the number of measurements and n is the number of samples, the mean of the sample for each ith measurement is subtracted and finally the basis vectors h_m, which are the eigenvectors of the covariance matrix $\mathbf{C} = E\{\mathbf{X}\}$ are calculated. The centered data is represented, for each ith measurement, by its Karhunen-Loéve expansion:

$$\mathbf{X} - E\{\mathbf{X}\} = \sum\nolimits_{m=1}^{N_m} \theta_m \mathbf{h}_m, \tag{3}$$

in which θ_m represents each basis expansion coefficient and N_m represents the full dimensionality of the problem [19, 28].

3.2. Classifiers Based on Support Vector Machines

This Section is presented in three steps: the first step describes the SVM method, the second step details the preprocessing of data for application of SVM, and the final step presents the different architectures of classifiers using the SVM implementation.

3.2.1. Support Vector Machines

The SVM learning algorithm is typically used as a method to handle nonlinear relations between the samples input vectors and their labels by succesfully classifying linearly inseparable data into two groups [30–37].

A learning machine is a type of classification algorithm in which the machine learns the system, so the user must provide a labeled training group from which the machine will generalise a decision (hyperplane) to differentiate between the classes of any member within the testing group. For the particular case of the SVM, the input vectors are mapped to a higher-dimension feature space by means of a Kernel (K) [34, 35].

The Kernel used for this study is the Radial Basis Function (RBF), which allows for all input vectors to be non-linearly mapped in an infinite-dimension feature space, typically a Hilbert Space. The decision hyperplane can then be obtained in the feature space and is generically given in the following format:

$$\mathbf{w}x + b = 0 \tag{4}$$

in which $\mathbf{w}$ is the normal to the hyperplane, x is the data and b is the bias.

Knowing that the data can be represented by the inner product $x_i \cdot x_j$ (this is an implication of using an infinite feature space), the equation for the RBF is defined in Equation (5):

$$K(x_i, x_j) = \exp(-\gamma||x_i - x_j||^2), \quad \gamma > 0 \tag{5}$$

in which γ is the scaling factor of the RBF Kernel [30, 31].

It should be noted that the decision hyperplane is supported by two parallel vectors, one on each side of the hyperplane. Each of these support vectors are at the same distance from the hyperplane — margin — and each of them limits either the first or the second labels. A classifier will work better when the value for the margin is maximised, so the concept of a soft margin is introduced, as opposed

to hard margins, as described in [31, 34, 35]. When soft margins are used, it implies that the support vectors are most likely built with supporting samples that represent meaningful samples of the training group; while outlier samples, such as noisy data or unusual data, are ignored for the calculation of the support vectors. If such conditions are met, the learning machine ensures high generalisation [30, 32, 33] and therefore will be able to successfully classify an independent testing group. Knowing that the training set is composed of sample-label pairs (x_i, x_j), in which $i = 1, \ldots, l$ represents each sample, x_i represents the input vectors of each sample and y_i represents the respective label, the soft margins can be calculated by following the mathematical optimisation in Equation (6):

$$\min_{\mathbf{w},b,\xi} \left[\frac{1}{2}(\mathbf{w} \cdot \mathbf{w}) + C \sum\nolimits_{i=1}^{l} \xi_i \right] \tag{6}$$

with the following conditions: $y_i(\mathbf{w} \cdot x_i + b) \geq 1 - \xi_i$, in which the slack variable $\xi_i \geq 1$. While for the hard margin the data is scaled so that the margin equals 1, for the soft margin the margin can be below one as it is given by $1 - \xi_i$. However, this results in the increase of the objective function since the sum of errors, given by $\sum_{i=1}^{l} \xi_i$, is multiplied by C [30, 35]. The function of C is two-fold: it controls the relative weighting to keep $\mathbf{w} \cdot \mathbf{w}$ small (as the size of the margin is maximised) and it ensures that most samples have a functional margin of at least 1 [33].

3.2.2. Data Preprocessing

For application of SVM, the data needs to be preprocessed so that the SVM classification algorithm can be optimised adequately for the samples, as efficiently as possible. These preprocessing steps are as follows:

i) Scaling of the training and data set;
ii) Application of the Kernel function, RBF;
iii) Application of the Cross-Validation method;
iv) Optimisation of the RBF parameters.

In the first step, the input vectors (represented by the selected Principal Components) for each sample in both training and testing group are scaled to the range $[-1, +1]$. This step is important so that attributes in greater numeric ranges do not dominate those in smaller numeric ranges, and also that the computational load of the whole algorithm is restricted [30].

For the second step of the SVM algortihm, the RBF is applied.

Another issue that has to be considered is the fact that the combination of (C, γ) for the RBF Kernel has to be tested on the

training data through a cross-validation, or leave-one-out, method, as this will help to optimise the training accuracy (while minimising the generalisation error, i.e., the ratio of misclassified samples). A cross-validation method allows for outlier samples that represent noise or unusual data to be removed, and as a result, some of the outlier supporting samples may be omitted from the final solution. Effectively, the training data is divided in equally-sized A subsets and, sequentially, each of these subsets is tested against the remaining $(A-1)$ training subsets. And so the cross-validation accuracy is given by the percentage of samples correctly classified, or by the average of the performance of each sub-classification [30, 31].

Finally, the parameters of the chosen Kernel function are adjusted so that the classifier is successful in classifying independent testing groups. For the RBF, the combination of (C, γ) is optimised, C is the penalty parameter of the error term, and as earlier mentioned, γ is the scaling factor of the RBF Kernel. A parameter search, such as the Grid-Search described in [30], is applied to the data set.

3.2.3. Classifier Architectures

Six different classifier architectures are considered, in which there are different combinations of size and/or shape SVM binary classifiers. The different architectures are defined by the size and/or shape granularity, i.e., if there are coarse or fine size and/or shape classifiers. Coarse size classifiers split tumours in one step into two size groups (the first group has 2.5 and 5 mm tumours and the second group has 7.5 and 10 mm tumours) and fine size classifiers further split the RTS into the four subcategories of size (2.5, 5, 7.5 and 10 mm) in two steps. Coarse shape classifiers split tumours into two groups in one step (benign and malignant tumours) and Fine shape classifiers further split benign tumours into smooth and macrolobulated tumours, and malignant tumours into microlobulated and spiculated tumours in two steps. In this study, the following architectures of classifiers are used: Coarse-Shape (CS), Fine-Shape (FS), Coarse-Size-Coarse-Shape (CSCS), Coarse-Size-Fine-Shape (CSFS), Fine-Size-Coarse-Shape (FSCS) and Fine-Size-Fine-Shape (FSFS). For more details on these architectures, the reader is advised to refer to the authors' previous work [22].

It must be emphasised that the use of a coarse or fine size classifier does not influence the decision of whether a tumour is benign or malignant, however it may influence the performance of the following shape classifier. Additionally, it must be emphasised that a coarse shape classifier is used to classify tumours into either malignant or benign tumours, which may be sufficient for most clinical applications. However, the extra granularity available through the

fine shape classifier provides further classification of tumours giving important clinical information on the development stage of a breast tumour, for instance a macrolobulated shape could potentially be an indicator of pre-malignancy and therefore closer surveillance of the patient may be required.

4. RESULTS AND DISCUSSION

For all results, a database of 352 models, with signals recorded from four different angles, was used for training and testing the classifiers, using the Cross-Validation method described in Section 3.2. Similar experimental work was previously carried out by the authors [22] and results indicated that 30 principal components, extracted through PCA, are sufficient as they offer a good compromise between classification accuracy and computational time.

The results for the six different architecture of classifiers are presented in Table 1, which presents the names of the different architectures and the corresponding accuracy results. The accuracy results are shown in three columns which represent:

- the partial accuracy for the size classification;
- the partial accuracy for the shape classification;
- the overall accuracy for the size-then-shape classifier.

The accuracy of the partial size or the partial shape classifier is expressed in terms of the proportion of tumours correctly identified in terms of size or shape, respectively, in isolation. The overall accuracy for the size-then-shape classifier is calculated by multiplying the partial accuracies for the size and shape classifiers and represents

Table 1. Accuracy for size and subsequent shape classifiers and overall size-then-shape classifier using SVM binary classification for six different architectures of classifiers, with γ set to 16 and C set to 256. LDA and QDA results obtained in [22] are added for direct comparison of classification methods (*).

Architectures of classifiers	*Partial size classifier (%)*			*Partial shape classifier (%)*			*Size-then-shape classifier (%)*		
	SVM	LDA*	QDA*	SVM	LDA*	QDA*	SVM	LDA*	QDA*
Coarse-Shape	N/A			89.20	80.90	84.03	N/A		
Fine-Shape	N/A			72.73	58.33	55.90	N/A		
Coarse-Size-Coarse-Shape	94.89	93.05	91.32	90.62	87.15	82.29	85.99	81.10	75.13
Coarse-Size-Fine-Shape	94.89	93.05	91.32	75.00	67.36	62.15	71.16	62.68	56.76
Fine-Size-Coarse-Shape	86.93	79.86	72.22	90.34	86.80	84.72	78.53	69.32	61.19
Fine-Size-Fine-Shape	86.93	79.86	72.22	75.28	69.44	64.58	65.44	55.46	46.64

the percentage of tumours correctly classified in terms of both size and shape.

Examining Table 1, it must be noted that there is an error propagation when a partial coarse size (or shape) classifier is extended to a fine classifier: as any misclassified tumours in the first step of the size (or shape) classification — which corresponds to the results of the coarse classifier — are necessarily misclassified in the second step of size (or shape) classification. This error propagation results in an absolute performance decrease of 7.96% for the partial size classification and between 15.06% and 16.47% for the partial shape classifications.

Examing the six architectures in Table 1, it can be observed that applying a size classifier, both coarse and fine, to the whole group of 352 tumours, results in better performance than applying a shape classifier in isolation, for both coarse and fine classifiers. In absolute terms, the performance of the coarse size classifier is 5.69% higher than the coarse shape classifier, and the fine size classifier is 14.20% higher than the fine shape classifier when applied directly to the whole data set. Given these results it is sensible to investigate the performance of a shape classifier when a size classifier is previously applied, like in architectures CSCS, CSFS, FSCS and FSFS.

It is also observed that the partial shape classifier has slightly higher performance when there is a previous size classifier compared to when there is no previous size classification: the partial coarse shape classifier is highest in CSCS, 90.62%, when there is a previous coarse size classifier, and the partial fine shape classifier is highest in FSFS, 75.28%, when there is a previous fine size classifier. However, it must be noted that the performance of the partial shape classifiers does not vary significantly depending on whether there is a previous size classifier or not, in absolute terms, the performance of the coarse shape classifier in CSCS is 1.42% higher than in CS and the performance of the fine shape classifier in FSFS is 2.55% higher than in FS. Another observation that must be made is that shape classifiers perform significantly better when discriminating between malignant and benign tumours when compared to classifiers that further discriminate between the four shape categories. This difference of performance varies between 15.06 and 16.47% in absolute terms.

In terms of the overall performance of the classifiers it is observed that the more steps of partial size and/or shape classifiers the lower the overall performance. Conversely, the fewer steps of partial size and/or shape classifiers the higher the overall performance. The overall performance drops by 23.76%, in absolute terms, from the architectures CS to FSFS. The performance of the classifiers decreases with the increase of granularity for two specific reasons:

- Firstly, for fine classification (e.g., differentiating between smooth and macrolobulated), the RTSs of the tumours are quite similar, so classification is much more difficult and misclassifications are much more likely to occur;
- Secondly, when classifiers are grouped in architectures such as the ones used in this study, errors can propagate through the multi-stage classifier. For instance, a microlobulated tumour which is first classified as benign will never be classified correctly in a fine shape classifier (as it will automatically be misclassified as a smooth or a macrolobulated tumour).

The misclassified tumours in terms of both size and shape are recorded at each step of the classification architectures. In general it was observed that the number of tumours misclassified for one class was very similar to the number of tumours misclassified for the other class. However, there were two noticeable exceptions:

- The fine classifier misclassified several smooth tumours as macrolobulated and vice-versa despite the type of size pre-classification. This is due to the similarity between these two types of benign tumours.
- There is a significant number of spiculated tumours misclassified as microlobulated and vice-versa for larger tumours (with 7.5 and 10 mm radius). This can be explained by the fact that the spicules of the spiculated tumours are always the same length (independent of the tumour radius). Therefore, in smaller tumours models, the spicules extend further beyond the surface of the tumour compared to larger tumours. In larger malignant tumours, the spicules may not influence the RTS of the tumours as much as for smaller spiculated tumours, and therefore misclassifications between larger spiculated and microlobulated tumours are more likely to occur.

Comparing the SVM results with the LDA and QDA results obtained in [22] for a 288 tumour-database, the SVM outperforms the other two classifiers for all the architectures used for this study: CS, FS, CSCS, CSFS, FSCS and FSFS. The most noticeable improvements are observed in the size-then-shape performance of the FSFS, which is 9.98% and 18.80% higher, in absolute terms, for SVM than for LDA and QDA, respectively. Also, in absolute terms, the partial coarse shape classifier in CSCS is 3.47% and 8.33% higher in SVM than it is for LDA and QDA, respectively; and the partial fine shape classifier in FSFS is 5.84% and 10.70% higher in SVM than for LDA and QDA, respectively. The reason why SVM results outperforms both LDA and QDA classification algorithms is due to the fact that SVM is

able to handle nonlinear relations between measurements and their classes by mapping the measurements in a higher dimension feature space, allowing for an optimised classification whereas the dimension feature space of LDA and QDA is limited by the dimensions of the measurements.

5. CONCLUSIONS AND FUTURE WORK

In this paper, a novel classification scheme based on SVM is used to classify 352 breast tumours in terms of size and shape. The size is correctly classified in coarse and fine groups with an accuracy of 94.89% and 86.93%, respectively. The shape is correctly classified in coarse groups with an accuracy varying between 89.20% to 90.62% between the CS and the CSCS architectures, identifying a tumour as either benign or malignant. The shape is correctly classified in fine groups with an accuracy varying between 72.73% to 75.28% between the FS and the FSFS architectures, indentifying more precisely the stage of development of a tumour. Regarding the size-then-shape accuracy of the classifiers, it can be concluded that as granularity increases, the classification performance decreases.

Future work will include the study of dielectric heterogeneity and its impact on classification algorithms. Spiking Neural Networks will also be investigated as an alternative tumour classification algorithm.

ACKNOWLEDGMENT

This work is supported by Science Foundation Ireland (SFI) under grant number 07/RFP/ENEF420.

REFERENCES

1. Lazebnik, M., L. McCartney, D. Popovic, C. B. Watkins, M. J. Lindstrom, et al., "A large-scale study of the ultrawideband microwave dielectric properties of normal breast tissue obtained from reduction surgeries," *Phys. Med. Biol.*, Vol. 52, 2637–2656, 2007.
2. Lazebnik, M., D. Popovic, L. McCartney, C. B. Watkins, M. J. Lindstrom, 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.

3. Hagness, S. C., 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 T-BME*, Vol. 45, 1470–1479, 1998.
4. Hagness, S. C., A. Taflove, and J. E. Bridges, "Three-dimensional FDTD analysis of a pulsed microwave confocal system for breast cancer detection: Design of an antenna-array element," *IEEE T-AP*, Vol. 47, No. 5, 783–791, 1999.
5. Fear, E. C., X. Li, S. C. Hagness, and M. A. Stuchly, "Confocal microwave imaging for breast cancer detection: Localization of tumors in three dimensions," *IEEE T-BME*, Vol. 49, No. 8, 812–822, 2002.
6. Meaney, P. M., K. D. Paulsen, J. T. Chang, M. W. Fanning, and A. Hartov, "Nonactive antenna compensation for fixed-array microwave imaging: Part II — Imaging results," *IEEE T-MI*, Vol. 18, No. 6, 508–518, 1999.
7. Meaney, P. M., M. W. Fanning, D. Li, S. P. Poplack, and K. D. Paulsen, "A clinical prototype for active microwave imaging of the breast," *IEEE T-MTT*, Vol. 48, No. 11, 1841–1853, 2000.
8. Meaney, P. M., M. W. Fanning, T. Raynolds, C. J. Fox, Q. Fang, et al., "Initial clinical experience with microwave breast imaging in women with normal mammography," *Acad. Radiol.*, Vol. 14, No. 2, 207–218, 2007.
9. Souvorov, A. E., A. E. Bulyshev, S. Y. Semenov, R. H. Svenson, and G. P. Tatsis, "Two-dimensional computer analysis of a microwave flat antenna array for breast cancer tomography," *IEEE T-MTT*, Vol. 48, No. 8, 1413–1415, 2000.
10. Bulyshev, A. E., S. Y. Semenov, A. E. Souvorov, R. H. Svenson, A. G. Nazarov, et al., "Computational modeling of three-dimensional microwave tomography of breast cancer," *IEEE T-BME*, Vol. 48, No. 9, 1053–1056, 2001.
11. Liu, Q. H., Z. Q. Zhang, T. T. Wang, J. A. Bryan, G. A. Ybarra, et al., "Active microwave imaging I — 2-D forward and inverse scattering methods," *IEEE T-MTT*, Vol. 50, No. 1, 123–133, 2002.
12. Kosmas, P. and C. M. Rappaport, "Time reversal with the FDTD method for microwave breast cancer detection," *IEEE T-MTT*, Vol. 53, No. 7, 2317–2323, 2005.
13. Kosmas, P. and C. M. Rappaport, "FDTD-based time reversal for microwave breast cancer detection — Localization in three dimensions," *IEEE T-MTT*, Vol. 54, No. 4, 1921–1927, 2006.
14. Kosmas, P. and C. M. Rappaport, "A matched-filter FDTD-based

time reversal approach for microwave breast cancer detection," *IEEE T-AP*, Vol. 54, No. 4, 1257–1264, 2006.

15. Chen, Y., E. Gunawan, K. S. Low, S.-C. Wang, C. B. Soh, et al., "Effect of lesion morphology on microwave signature in ultra-wideband breast imaging: A preliminary two-dimensional investigation," *IEEE AP-S International Symposium*, 2168–2171, 2007.
16. Chen, Y., E. Gunawan, K. S. Low, S.-C. Wang, C. B. Soh, et al., "Effect of lesion morphology on microwave signature in 2-D ultra-wideband breast imaging," *IEEE T-BME*, Vol. 55, No. 8, 2011–2021, 2008.
17. Rangayyan, R. M., N. M. El-Faramawy, J. E. L. Desautels, and O. A. Alim, "Measures of acutance and shape for classification of breast tumors," *IEEE T-MI*, Vol. 16, No. 6, 799–810, 1997.
18. Chen, Y., I. J. Craddock, and P. Kosmas, "Feasibility study of lesion classification via contrast-agent-aided UWB breast imaging," *IEEE T-BME*, Vol. 57, No. 5, 1003–1007, 2010.
19. Davis, S. K., B. D. V. Veen, S. C. Hagness, and F. Kelcz, "Breast tumor characterization based on ultrawideband microwave backscatter," *IEEE T-BME*, Vol. 55, No. 1, 237–246, 2008.
20. Muinonen, K., "Introducing the Gaussian shape hypothesis for asteroids and comets," *Astronomy and Astrophysics*, Vol. 332, 1087–1098, 1998.
21. Muinonen, K., *Light Scattering by Stochastically Shaped Particles, in Light Scattering by Nonspherical Particles: Theory, Measurements, and Applications*, M. I. Mishchenko, J. W. Hovenier, and L. D. Travis (eds.), Chapter 11, Academic Press, 2000.
22. Conceição, R. C., M. O'Halloran, E. Jones, and M. Glavin, "Investigation of classifiers for early-stage breast cancer based on radar target signatures," *Progress In Electromagnetics Research*, Vol. 105, 295–311, 2010.
23. Conceição, R. C., D. Byrne, M. O'Halloran, M. Glavin, and E. Jones, "Classification of suspicious regions within ultrawideband radar images of the breast," *16th IET ISSC 2008*, Vol. 1, 60–65, Galway, Ireland, 2008.
24. Nguyen, T. M. and R. M. Rangayyan, "Shape analysis of breast masses in mammograms via the fractial dimension," *Engineering in Medicine and Biology 27th Annual Conference*, 3210–3213, Shangai, China, 2005.
25. Guliato, D., R. M. Rangayyan, J. D. Carvalho, and S. A. Santiago, "Polygonal modeling of contours of breast tumors with the

preservation of spicules," *IEEE T-BME*, Vol. 55, No. 1, 14–20, 2008.

26. Taflove, A. and S. C. Hagness, *Computational Electrodynamics: The Finite-difference Time-domain Method*, 2nd edition, Artech House, Boston, 2000.
27. Wold, H., "Estimation of principal components and related models by iterative least squares," *Multivariate Analysis*, K. R. Krishnaiah (ed.), 391–420, Academic Press, New York, 1996.
28. Shlens, J., "A tutorial on principal component analysis," Mar. 25, 2003, Available from: http://www.cs.princeton.edu/picasso/mats/PCA-Tutorial-Intuition_jp.pdf.
29. Bartholomew, D. J., F. Steele, I. Moustak, and J. I. Galbraith, "The analysis and interpretation of multivariate data for social scientists," *Statistical Science*, Chapman & Hall/CRC, USA, 2002.
30. Hsu, C.-W., C.-C. Chang, and C.-J. Lin, "A practical guide to support vector classification," Apr. 3, 2010, Available from: www.csie.ntu.edu.tw/ ∼ cjlin/papers/guide/guide.pdf.
31. Boser, B. E., I. M. Guyon, and V. N. Vapnik, "A training algorithm for optimal margin classifiers," *Proc. of the 5th Annual Workshop on Computational Learning Theory*, 144–152, Pittsburgh, Pennsylvania, 1992.
32. Cortes, C. and V. Vapnik, "Support-vector networks," *Machine Learning*, Vol. 20, No. 3, 273–297, 1995.
33. Ng, A., "Support vector machines (Part V of CS229 machine learning course materials)," May 13, 2010, Available from: http://www.stanford.edu/class/cs229/notes/cs229-notes3.pdf.
34. Bennett, K. P. and C. Campbell, "Support vector machines: Hype or hallelujah?" *ACM SIGKDD Explorations Newsletter*, Vol. 2, No. 2, 1–13, 2000.
35. Campbell, C., "Introduction to support vector machines," Feb. 5, 2008, Available from: http://videolectures.net/epsrcws08_campbell_isvm/.
36. Donelli, M., F. Viani, P. Rocca, and A. Massa, "An innovative multiresolution approach for DOA estimation based on a support vector classification," *IEEE T-AP*, Vol. 57, No. 8, 2279–2292, 2009.
37. Viani, F., P. Meaney, P. Rocca, R. Azaro, M. Donelli, et al., "Numerical validation and experimental results of a multi-resolution SVM-based classification procedure for breast imaging," *IEEE AP-S International Symposium, APSURSI*, 1–4, Charleston, USA, 2009.