Imaging-based Classification Algorithms on Clinical Trial Data with Injected Tumour Responses

Yunpeng Li, Emily Porter, Mark Coates
Dept. of Electrical and Computer Engineering, McGill University, Montréal, Québec, Canada
yunpeng.li@mail.mcgill.ca, emily.porter@mail.mcgill.ca, mark.coates@mcgill.ca

***Abstract*—Current microwave breast cancer imaging algorithms focus primarily on generating an image, and provide little machinery for interpretation of the image. Within-image contrast is commonly used as a performance metric, but a better reflection of the tumour detection capability of an algorithm is the difference between the maximum voxel intensities observed in images from scans of tumour-free and tumour-bearing breasts. This paper extends existing imaging algorithms by incorporating an automatic tumour detection technique that involves classification based on maximum voxel intensities. We compare results obtained from different algorithms on the data collected from healthy breast scans performed during clinical trials of a microwave radar system. We artificially inject tumour signals that are constructed based on the transmission properties of the radar system and the estimated breast tissue properties. The results provide insights into which algorithms are sufficiently robust to handle discrepancies between the real measurement data and the modeling assumptions.**

***Index Terms*—microwave breast cancer detection, clinical trial, imaging algorithms.**

I. Introduction

Microwave breast imaging techniques have been proposed to complement the standard methods of x-ray mammography (for screening), and ultrasound and magnetic resonance imaging (MRI) (for diagnosis). Microwave methods offer potential for comfortable, cost-effective breast scans that do not involve ionizing radiation [1]. The techniques are based on a reported inherent contrast in the dielectric properties of healthy and malignant tissues, over the microwave frequency range.

There are two types of microwave imaging techniques: tomographic and radar. Microwave tomography aims to recover a map of the electrical properties of the breast tissue [2]. It involves solving an ill-conditioned inverse problem that can be computationally expensive, and generally requires a large number of antennas located around the breast in order to enable high definition in the recovered map. For these reasons, we focus only on backscattered radar methods.

The most common radar method, confocal microwave imaging (CMI) [3], and its variant, delay-multiply-and-sum (DMAS) [4], are based on time shifting and summing of backscattered signals. Other well-known radar imaging methods include microwave imaging via space-time beamforming (MIST) [5], robust weighted Capon beamforming (RWCB) [6], and multistatic adaptive microwave imaging [7]. All of these methods focus on generating images; a clinical expert must then examine the images to determine the tumour existence.

Another type of microwave imaging technique is to view the breast cancer imaging challenge as a hypothesis testing problem for each voxel, with the null hypothesis representing the tumour-free case. A generalized likelihood ratio test (GLRT) is formulated to detect the tumour presence on a voxel-to-voxel basis in [8]. Its performance is assessed with numerical and experimental breast phantoms.

Although the GLRT algorithm is formulated as a detection problem, the detection is at the voxel level (is the voxel a scatterer or not?). It does not strive to make a global decision about the presence or absence of a tumour in a breast. In the study presented here, we use the maximum voxel intensity obtained by applying various imaging algorithms to breast scan data as the input to a classifier. We examine the performance of the delay-multiply-and-sum (DMAS) algorithm [4], the MIST algorithm from [5], the weighted Capon beamforming (WCB) [6], and the GLRT algorithm of [8].

The use of a classifier allows us to make a global decision as to whether a breast contains a tumour. We can then compare different algorithms based on their detection performance. Our experiments involve signals obtained using an experimental system applied to volunteers with healthy breasts. The tumour signals are constructed numerically based on the transmission properties of the system and the estimated breast tissue properties. The analysis thus provides insights into which algorithms are sufficiently robust to handle discrepancies between the real measurement data and the modeling assumptions.

The paper is organized as follows. We introduce the problem statement in Section II. The detection algorithm is proposed in Section III. Section IV introduces the clinical data and injected tumour responses. Finally, we present the classifier performance for the different algorithms in Section V.

II. Problem statement

A common setup for a microwave imaging system is that each antenna of the array surrounding the breast sequentially sends an ultra-wide bandwidth (UWB) pulse to the breast and collects the backscattered signal. Thus, there are $M = R(R-1)$ directed antenna pairs for the R antennas. One measurement contains received pulses for each antenna pair.

Suppose that we have access to a set of measurements from different people. We assume that we have multiple measurements from each individual, taken at monthly or yearly intervals. The first scan from each individual is considered

a "baseline"; we assume that it corresponds to a healthy breast state, confirmed through scans using other modalities. In subsequent measurements the breast may remain tumour-free ("nominal"), or a tumour may develop. Now we are provided with measurements from a new patient, from which we have baseline and tumour-free measurements collected as before. We want to decide from her current measurements whether a tumour has developed or not. In making this decision, we need to control the trade-off between the false-alarm rate and detection power.

III. Imaging-based classifier

The microwave breast cancer detection algorithm we propose consists of two stages. In the first stage, the breast region is divided into small voxels. We use different imaging algorithms to generate different image intensities for each voxel. The maximum voxel intensity for each measurement is then used to decide whether a tumour is present.

A. Imaging algorithms

Four well-known microwave breast imaging algorithms are used to generate the classifier input in this paper. The delay-multiply-and-sum algorithm (DMAS) [4] first extracts the backscatters from signals in each channel and estimates their time delays for each voxel position. The time-aligned backscatters are then multiplied pair-wisely before a summing and integration operation outputs the imaging intensities for each voxel. It exploits the fact that the synthetically aligned backscatters with the correct tumour location will be correlated, which results in a higher image intensity for that voxel.

Microwave imaging via space-time beamforming (MIST) [5] compensates for the frequency-dependent time delay by passing the time-aligned signal to finite impulse response (FIR) filters. The filter signals are then summed to produce the beamformer output.

The robust weighted Capon beamforming (RWCB) method [6] has a weighted combination of time-aligned signals in different channels to improve image resolution and suppress interference. The robust part of the approach involves solving an optimization problem for each voxel, which makes the algorithm prohibitively slow for batch processing of a large number of measurements, since we usually construct tens of thousands of voxels for the image generation. Thus we adopt the weighted Capon beamforming (WCB) algorithm.

A generalized likelihood ratio test (GLRT)-based breast cancer imaging algorithm [8] instead considers the imaging as hypothesis testing problems for each voxel, with the null hypothesis representing the tumour-free case. It requires knowledge of the signal template, which is a model for the backscattered signal with respect to each focus position.

B. Classification

Once an image is generated, it is used in the second stage to determine the presence of a tumour. For all measurements other than the first measurement (baseline) of each volunteer, we generate a 3-D image I and use the maximum intensity $\max(I)$ as the classifier score. This is compared with a threshold η to decide whether a tumour is present. We do not discuss the choice of η in this paper; instead we examine performance of the different imaging algorithms as η is varied over a reasonable range of values.

IV. Clinical data and simulated tumour response

A. Clinical trial

We performed breast scans on 12 healthy volunteers through a clinical trial using a time-domain multistatic radar system. A detailed description of the system prototype can be found in [9]. The scans were conducted over 48 patient visits spanning an 8-month period, with each volunteer visiting up to once per month. The data set includes scans from volunteers who visited a minimum of two and a maximum of six times. The volunteers ranged in age from 21 to 77, with bra cup sizes from A to D. Each breast scan contains 240 signals, one from each transmit-receive antenna pair (with 16 antennas overall). The data were collected with a sampling rate of 40GHz and a signal length of 1024 samples.

We treat the measurements of the left breast and the right breast from the same people as independent. There are two measurements for each breast during one visit. The latter one is taken after the volunteer re-positions herself. We use the $48 \times 2 = 96$ sets of the first measurement of each visit to represent the tumour-free data from different breasts, which can be denoted by $\{p_{ij}\}$, where $i \in \{1, 2, \ldots, 24\}$ is the index of breasts, and $j = 1, 2, \ldots, J(i)$ is the index of volunteer visits. For breast i, we can use the first set of tumour-free measurements p_{i1} as the baseline, and generate the differential signals $b_{ij} = p_{i(j+1)} - p_{i1}$ as the tumour-free differential data, for $j = 1, 2, \ldots, J(i) - 1$. Thus, there will be $96 - 24 = 72$ different B_{ij}.

Since we only have clinical data from healthy volunteers, there are no tumour-bearing breast measurements in our original dataset. However, we can simulate the tumour responses r_i for volunteer i, based on the transmitted pulses $g(n)$ and the dielectric properties of breast tissue and tumour. We then add them to the $48 \times 2 = 96$ sets of the measurements taken after volunteers re-positioned themselves, w_{ij}, to form the tumour-bearing measurements $q_{ij} = w_{ij} + r_i$, for $j = 2, \ldots, J(i)$. In this way, we obtain $96 - 24 = 72$ tumour-bearing differential data $t_{ij} = q_{i(j+1)} - p_{i1}$, for $j = 1, 2, \ldots, J(i) - 1$.

A diagram of the number of measurements for each of the 24 breasts is shown in Table I.

TABLE I
Number of tumour-free measurements for each visitor. Odd i corresponds to a volunteer's left breast, while even i corresponds to a volunteer's right breast.

i	1	3	5	7	9	11	13	15	17	19	21	23	Total
	2	4	6	8	10	12	14	16	18	20	22	24	
J(i)	3	3	4	5	2	6	6	4	4	4	3	4	96

B. Injected tumour response

To simulate the tumour response, we need to first simulate the 16-element antenna array system which is used to collect the data. The radome is modeled by a hemisphere $x^2/a^2 + y^2/b^2 + z^2/c^2 = 1$, where $z \geq 0$ is the depth information, $a = b = c = 7.3$cm. The antenna's positions in x and y axis mimic our real multi-static radar system. The breast is assumed to be in a smaller hemisphere, where $a' = b' = c' = 7$cm. The geometry of the breast model is shown here in Figure 1. The space

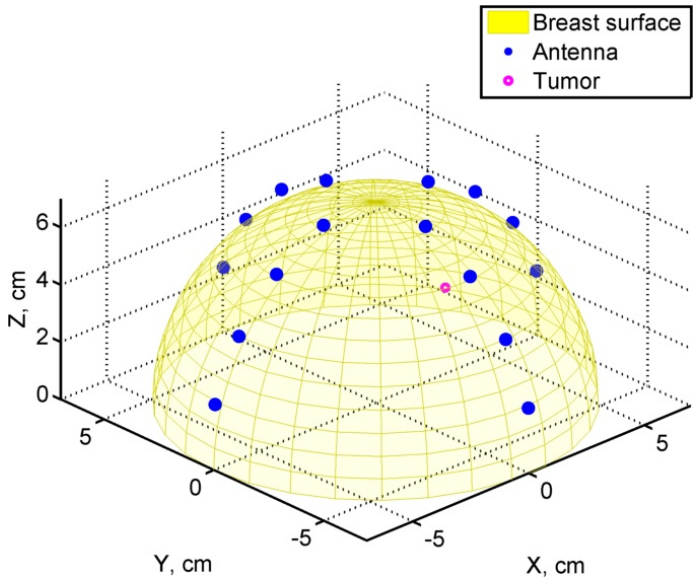

Fig. 1. The breast shape model, the antenna positions, and a tumour at $(1, 1, 1)cm$.

between the antennas and the breast is assumed to be filled with lossless non-dispersive immersion medium with $\epsilon_{rb} = 9.6$ to reduce the artefacts.

We simulate the transmitted pulse $g(n)$ using the sampled pulse fed into each antenna collected from a real multi-static radar system [9]. When a tumour exists at r_0, the frequency domain representation for the noise-free backscattered signal r^m for antenna pair m is denoted as in [5]

$$R^m(\omega) = \Gamma \tilde{R}^m(\omega) = \Gamma G(\omega) H_m(r_0, \omega), \quad (1)$$

where Γ is the portion of the signals propagating through the breast that will be reflected by the tumour. Γ is set to 0.2 which approximates the calculated values from the desired inserted tumour and breast tissue permittivity values. $G(\omega)$ is the Fourier transform of $g(n)$, and $H_m(r_0, \omega)$ is the frequency response of the propagation from the transmitting antenna to the tumour at r_0 then back to the receiving antenna, for antenna pair m. $H_m(r_0, \omega)$ can be represented as

$$H_m(r_0, \omega) = e^{-j(k_b d_b + k_n d_n)}. \quad (2)$$

k_b and k_n are the wavenumber for the coupling medium and the breast tissue, respectively. d_b and d_n are the lengths of the direct path for antenna pair m, in the coupling medium and the breast tissue, respectively. $k_b = 2\pi/\lambda_{rb}(\omega) = \sqrt{\epsilon_{rb}}\omega/c$, and $k_n = 2\pi/\lambda_r(\omega) = \sqrt{\epsilon_r(\omega)}\omega/c$ where ϵ_r is estimated using Debye model

$$\epsilon_r(\omega) = \epsilon_\infty + \frac{\Delta\epsilon}{1 + j\omega\tau} + \frac{\sigma_s}{j\omega\epsilon_0}. \quad (3)$$

We randomly set Debye parameters $(\epsilon_\infty, \Delta\epsilon, \tau, \sigma_s)$ to different values for different breasts according to uniform distributions: $\Delta\epsilon \sim \mathcal{U}(20, 32.08), \sigma_s \sim \mathcal{U}(0.36, 0.52)$. $\tau \sim \mathcal{U}(8.68, 13) \times 10^{-12}$. ϵ_∞ is set to 5.57, and $\epsilon_0 = 8.854 \times 10^{-}12$ is the permittivity of free-space. The range of the uniform distribution is chosen according to clinical experiment results from [10], [11] to produce permittivity values within the range of the values estimated from our experimental data. Different ϵ_r values will lead to tumour responses of different time delay and attenuation. To add more variability to the tumour responses, we also randomly choose different tumour positions for different volunteers. It is unlikely for the tumours to develop in the skin layer, so we randomly sample the tumour positions within a hemisphere of radius 6cm.

Figure 2 shows the clinical experimental data and the simulated tumour-bearing data from Antenna pair A8-A12, for the first two visits of Volunteer 6. Antenna pair A8-A12 is chosen because it has the largest tumour responses among all of the antenna pairs for Volunteer 6.

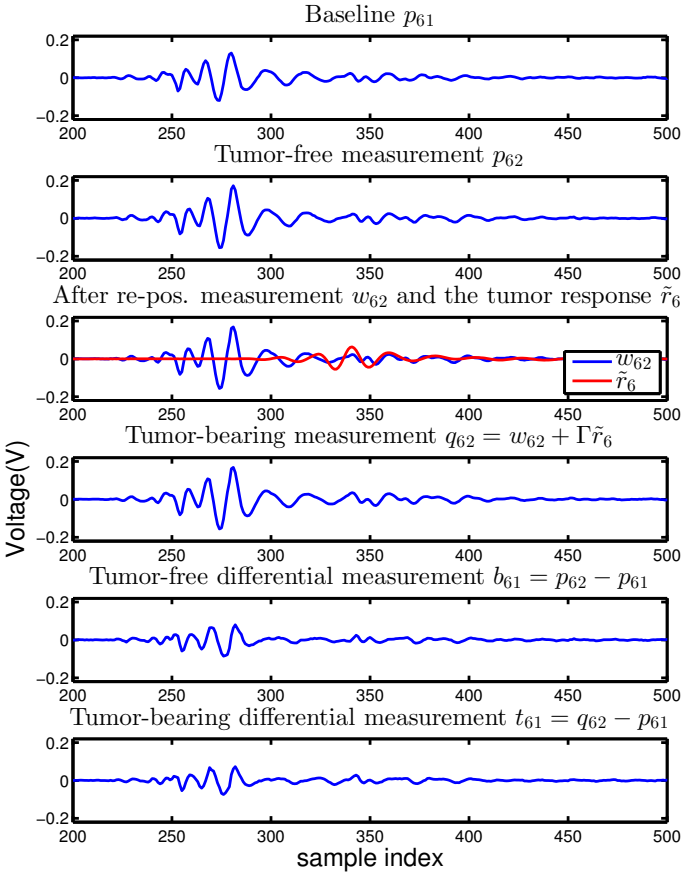

Fig. 2. The tumour-free and tumour-bearing breast measurements of Volunteer 6 from antenna pair A8-A12. The tumour position is $(x, y, z) = (-0.3, -0.1, 2.9)$cm. Γ is set to 0.2.

V. Experiment results

We divide the possible tumour locations (a hemisphere with radius 6cm) into voxels of dimension $2 \times 2 \times 2$mm. This leads to 56552 voxels inside the region.

We estimate average ϵ_r for each breast using the tumour-free measurements to calculate the average propagation speed required in the algorithms. Some of the values are significantly different than the values we used to generate the tumour responses. This mimics the real scenario where we often have limited knowledge about the exact tissue permittivity of the patients. For the same reason, we use a fixed set of Debye parameters $(\epsilon_\infty = 5.57, \Delta\epsilon = 26.04, \tau = 10.84 \times 10^{-12}, \sigma_s = 0.44)$ to generate signal templates for the GLRT algorithm. The FIR filter length in the MIST algorithm is used to control the time alignment, which is set to 10 for computational efficiency. The number of frequency samples used in MIST is 100, uniformly covering the range from 1GHz to 6GHz.

The imaging results of the four algorithms on the first nominal and tumour-bearing breast measurements from the left breast of Volunteer 2 are shown in the top rows of Figure 3–6. We show these images because they represent typical images from those algorithms. The top two images show slices where the maximum voxel intensity occurs in the 3D images.

The maximum image intensity of the tumour-bearing DMAS image is higher than the nominal image. But the tumour-bearing image is noisy, showing clutter of high intensity regions away from the true tumour position. It is very difficult to distinguish between the MIST images from the nominal and tumour-bearing measurements, since there are significant image processing artefacts. The frequency-related compensation of the signal delays does not appear to work since the real signal propagation environment is varying and unknown. The WCB tumour-bearing image does not show any patterns related to the tumours. The WCB algorithm requires a highly accurate description of the tumour response, which is unavailable in the real dataset. The GLRT image for the tumour-bearing measurement clearly indicates the presence of the tumour. Its maximum image intensity is also much higher than that from the nominal measurement.

We can observe from the bottom rows of Figure 3–6 that there is less overlap between the maximum GLRT image intensities from the tumour-bearing measurements and those from the nominal measurements, compared to the results from the other algorithms. This indicates that we can achieve higher predictive power by applying a global threshold to the GLRT outputs, which is confirmed in Figure 7.

We can use maximum image intensities from the 72 tumour-free and 72 tumour-bearing differential measurements as classifier score values to perform classification. We use a leave-8-out approach to generate 18 different ROC curves to obtain the average ROC curve in Figure 7. For the GLRT algorithm, the classifier is able to control the false positive rate at 0.1 while achieving a detection rate of 0.55. The performance of the other algorithms is much poorer; for the false positive rate of 0.1, the detection rate is only 0.1.

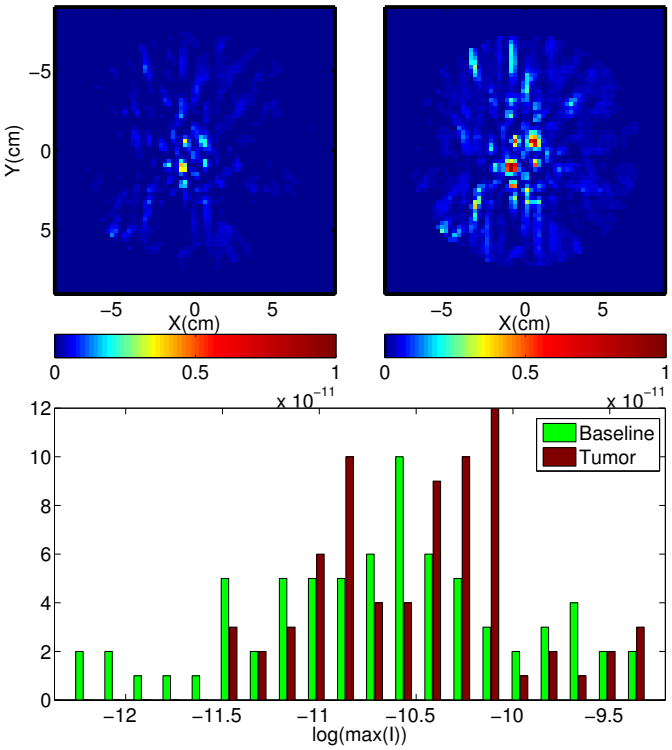

Fig. 3. Top: the DMAS imaging results at the slice where the maximum voxel intensity occurs, from the second visit of tumour-free (left) and tumour-bearing (right) measurements of Volunteer 2 (tumour position $(x, y, z) = (-1.5, -0.5, 2.7)$cm). Bottom: the maximum voxel intensities from the DMAS output of each measurement for all volunteers.

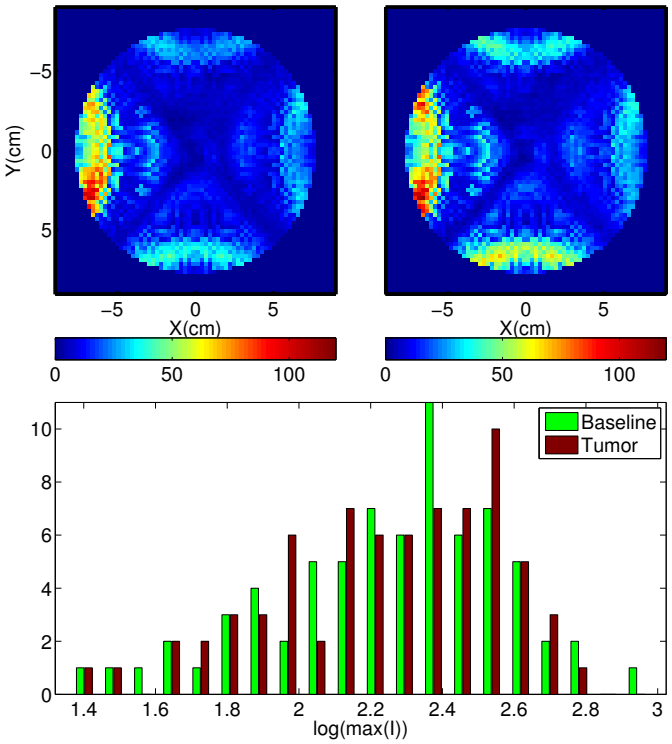

Fig. 4. The MIST results on the same data used in Figure 3.

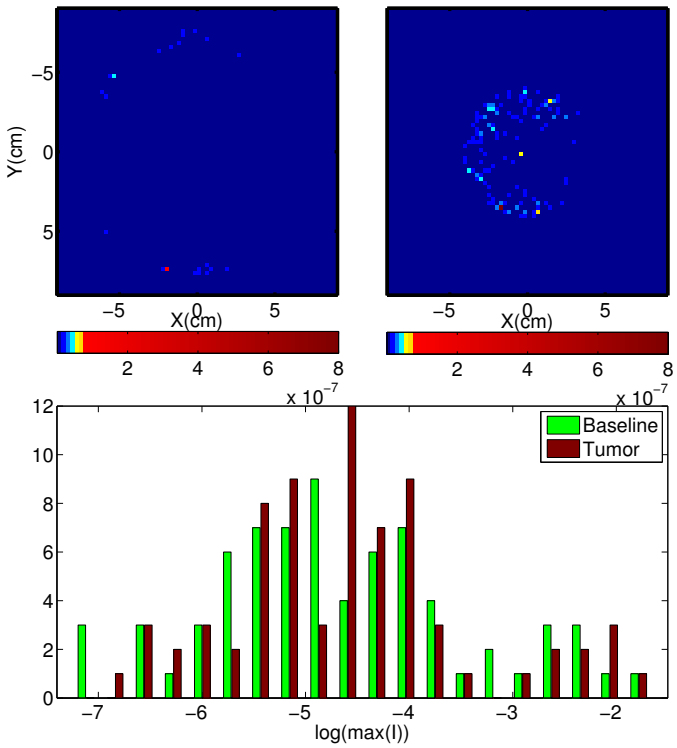

Fig. 5. The WCB results on the same data used in Figure 3.

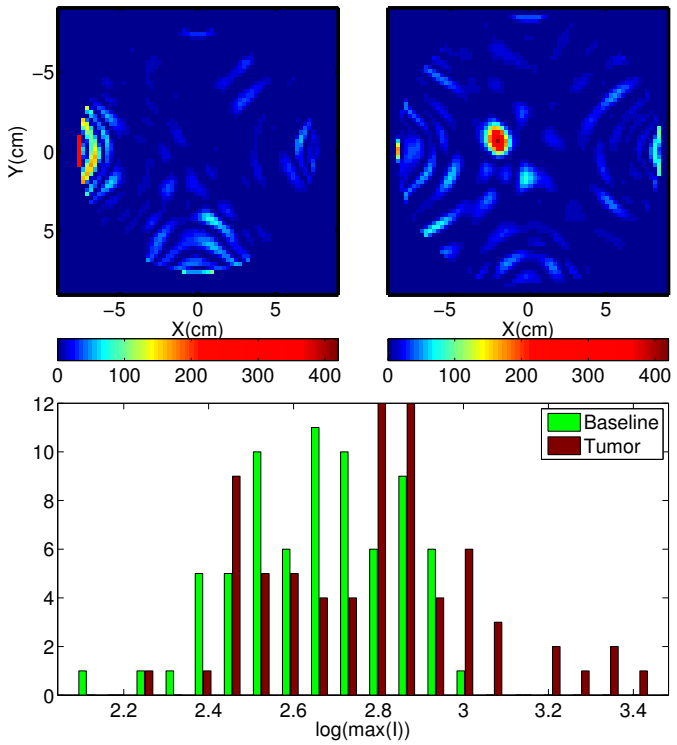

Fig. 6. The GLRT results on the same data used in Figure 3.

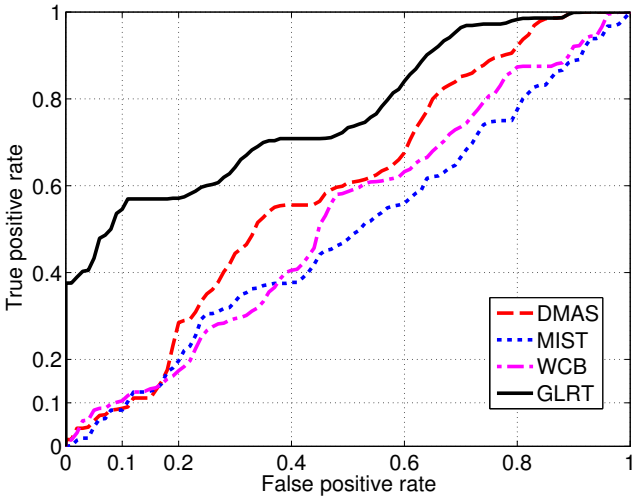

Fig. 7. Performance of classifiers based on different imaging algorithms.

VI. Conclusions

In this paper, we have compared classification techniques based on different imaging algorithms on clinical trial data. Tumour response signals are constructed based on the transmission properties of the system and the estimated breast tissue properties. Current results demonstrate that GLRT-based classification algorithm has better sensitivity and specificity than the other imaging algorithms-based classification. Future research directions include investigating more robust classifiers based on the imaging results.

References

[1] N. Nikolova, "Microwave imaging for breast cancer," *IEEE Microwave*, vol. 12, no. 7, pp. 78–94, Dec. 2011.

[2] T. M. Grzegorczyk et al., "Fast 3-d tomographic microwave imaging for breast cancer detection," *IEEE Trans. Med. Imag.*, vol. 31, pp. 1584–1592, Aug. 2012.

[3] E. C. Fear et al., "Confocal microwave imaging for breast cancer detection: localization of tumors in three dimensions," *IEEE Trans. Biomed. Eng.*, vol. 49, pp. 812–822, Aug. 2002.

[4] H. B. Lim et al., "Confocal microwave imaging for breast cancer detection: Delay-multiply-and-sum image reconstruction algorithm," *IEEE Trans. Biomed. Eng.*, vol. 55, pp. 1697–1704, June 2008.

[5] E. J. Bond et al., "Microwave imaging via space-time beamforming for early detection of breast cancer," *IEEE Trans. Antennas Propagat.*, vol. 51, pp. 1690–1705, Aug. 2003.

[6] B. Guo et al., "Microwave imaging via adaptive beamforming methods for breast cancer detection," *J. Electromagn. Waves and Appl.*, vol. 20, no. 1, pp. 53–63, 2006.

[7] Y. Xie et al., "Multistatic adaptive microwave imaging for early breast cancer detection," *IEEE Trans. Biomed. Eng.*, vol. 53, pp. 1647–1657, Aug. 2006.

[8] S. K. Davis et al., "Ultrawideband microwave breast cancer detection: a detection-theoretic approach using the generalized likelihood ratio test," *IEEE Trans. Biomed. Eng.*, vol. 52, pp. 1237–1250, July 2005.

[9] E. Porter et al., "Time-domain multistatic radar system for microwave breast screening," *IEEE Antennas Wireless Propag. Lett.*, vol. 12, pp. 229–232, 2013.

[10] M. Lazebnik 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, no. 20, pp. 6093–6115, 2007.

[11] ——, "Highly accurate Debye models for normal and malignant breast tissue dielectric properties at microwave frequencies," *IEEE Microw. Wireless Comp. Lett.*, vol. 17, no. 12, pp. 822–824, 2007.