

Journal of Biomedical Optics

BiomedicalOptics.SPIEDigitalLibrary.org

Pulsed terahertz imaging of breast cancer in freshly excised murine tumors

Tyler Bowman
Tanny Chavez
Kamrul Khan
Jingxian Wu
Avishek Chakraborty
Narasimhan Rajaram
Keith Bailey
Magda El-Shenawee

Pulsed terahertz imaging of breast cancer in freshly excised murine tumors

Tyler Bowman,^{a,*} Tanny Chavez,^a Kamrul Khan,^b Jingxian Wu,^a Avishek Chakraborty,^b Narasimhan Rajaram,^c Keith Bailey,^d and Magda El-Shenawee^a

^aUniversity of Arkansas, Bell Engineering Center, Department of Electrical Engineering, Fayetteville, Arkansas, United States

^bUniversity of Arkansas, Science and Engineering Building, Department of Mathematical Sciences, Fayetteville, Arkansas, United States

^cUniversity of Arkansas, Bell Engineering Center, Department of Biomedical Engineering, Fayetteville, Arkansas, United States

^dOklahoma State University, Oklahoma Animal Disease Diagnostic Laboratory, Stillwater, Oklahoma, United States

Abstract. This paper investigates terahertz (THz) imaging and classification of freshly excised murine xenograft breast cancer tumors. These tumors are grown via injection of E0771 breast adenocarcinoma cells into the flank of mice maintained on high-fat diet. Within 1 h of excision, the tumor and adjacent tissues are imaged using a pulsed THz system in the reflection mode. The THz images are classified using a statistical Bayesian mixture model with unsupervised and supervised approaches. Correlation with digitized pathology images is conducted using classification images assigned by a modal class decision rule. The corresponding receiver operating characteristic curves are obtained based on the classification results. A total of 13 tumor samples obtained from 9 tumors are investigated. The results show good correlation of THz images with pathology results in all samples of cancer and fat tissues. For tumor samples of cancer, fat, and muscle tissues, THz images show reasonable correlation with pathology where the primary challenge lies in the overlapping dielectric properties of cancer and muscle tissues. The use of a supervised regression approach shows improvement in the classification images although not consistently in all tissue regions. Advancing THz imaging of breast tumors from mice and the development of accurate statistical models will ultimately progress the technique for the assessment of human breast tumor margins. © 2018 Society of Photo-Optical Instrumentation Engineers (SPIE) [DOI: 10.1117/1.JBO.23.2.026004]

Keywords: terahertz; medical imaging; biomedical optics; breast cancer; mice; statistical modeling; morphing.

Paper 170747RR received Nov. 18, 2017; accepted for publication Jan. 23, 2018; published online Feb. 14, 2018.

1 Introduction

Terahertz (THz) imaging technology has been a growing area of interest for biomedical applications.¹ For example, pulsed THz systems have been used to conduct imaging of liver cirrhosis,² osseous tissue damage,³ and differentiation of cancerous tissues of the brain,⁴ stomach,⁵ colon,⁶ and breast,^{7–10} among others. However, imaging of breast and other cancers has been mostly limited to formalin-fixed, paraffin-embedded (FFPE) tissue to date, with only a few studies working with freshly excised tissue due to the challenges of obtaining fresh tissue outside of a surgical setting.^{4,10–12} One alternative to fresh human tissue, as used in this work, is tumors from mice. Mice have notably been used to generate xenograft brain tumors for THz imaging of fresh and FFPE tissue using either 9L/lacZ glioma cells⁴ or 6C glioma cells.¹³ They have also been used to obtain FFPE xenograft liver tumors from H22 liver cancer cells for THz imaging¹⁴ and breast cancer tumors grown from MDA MB 231 cells imaged subcutaneously *in vivo* and fresh *ex vivo*.¹⁵ This work makes use of E0771 mouse-derived breast adenocarcinoma to generate xenograft tumors for excision and THz imaging. Additionally, a high-fat diet is used to provide a sufficient fatty background for xenograft tumors to more closely resemble human excised tissue.¹⁶

While a significant amount of THz imaging has been able to qualitatively compare produced images of tissue to pathology sections,^{8,9} the quantification of THz image accuracy has

been limited so far. This process generally requires some automatic classification or thresholding of THz image data and digitization of pathology information for comparison. Digitization of pathology in this work is obtained using a morphing algorithm, which enables a pixel-by-pixel comparison between the THz images and digitized pathology results.^{7,17} Of classification methods used for THz imaging of fresh tissue, the use of support vector machines (SVM) and principal component analysis (PCA) reported up to 92% accuracy for breast cancer when combined.¹⁸ The techniques used separately showed a 96% sensitivity and 87% specificity for SVM and 92% sensitivity and 87% specificity for PCA of normal versus dysplastic tissue in colon cancer imaging.¹⁹ SVM also attained a 72% discrimination in 1.89-THz continuous wave imaging of breast cancer.⁷ In applications not handling fresh tissue, SVM was used for FFPE tissue imaging and spectroscopy.^{5,20,21} Only a few other methods have been used for fresh tissue imaging, such as PCA for murine brain glioma.⁴ However, several classification methods have been applied to FFPE tissue applications and/or spectroscopy of tissues, including wavelet transformation for osteosarcoma,²² orthogonal signal correction and fuzzy rule-bending expert system for cervical cancer,²³ multispectral classification of FFPE basal cell carcinoma,²⁴ and PCA for potentially malignant skin nevi²⁵ and FFPE liver cancer.^{26,27}

This work investigates a Bayesian mixture model utilizing a Markov chain Monte Carlo (MCMC) scheme for THz image classification of both fresh and FFPE murine breast tumors.²⁸

*Address all correspondence to: Tyler Bowman, E-mail: tcbowman@uark.edu

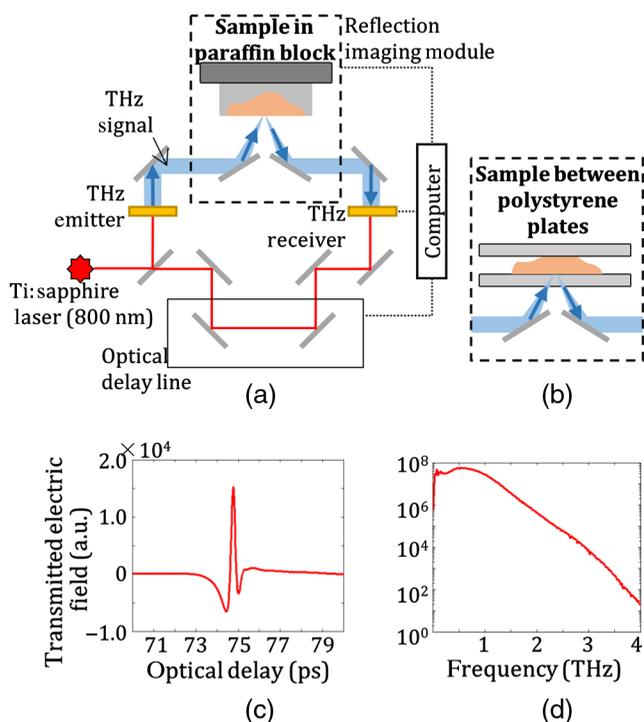


Fig 1. Diagram of THz imaging system in reflection mode (a) for FFPE tissue in paraffin block or (b) for fresh tissue between polystyrene plates. (c) Incident time-domain signal and (d) incident spectrum following Fourier transform.

This work is different from the authors' previous work where they performed qualitative THz imaging and characterization of FFPE breast cancer tissue^{8,9} as well as THz imaging and image processing of three-dimensional FFPE breast cancer tissue and characterization of carbon nanoparticles for THz contrast enhancement.^{29,30} Preliminary results using the methods in this work on a frozen tissue sample have been published in conference proceedings.^{31,32} To the authors' knowledge, this is the first time the Bayesian mixture model has been applied to THz imaging applications, and it is also the first time that the E0771 breast adenocarcinoma cells have been used in THz imaging.

This work is organized as follows: methodology including mice tumor sample preparation, image acquisition, and the statistical model in Sec. 2; the results of THz image correlation with pathology in Sec. 3; and discussion and concluding remarks in Sec. 4.

2 Methodology

2.1 Terahertz Imaging Setup

This work makes use of the TPS Spectra 3000 THz pulsed imaging and spectroscopy system (TeraView, Ltd., United Kingdom) at the University of Arkansas. A diagram of the system can be seen in Fig. 1. The THz emitter is a biased bowtie antenna on a GaAs substrate excited by a femtosecond Ti:sapphire laser pulse to generate the THz signal. The signal is then directed using mirrors to the sample as shown in Fig. 1(a). For this work, two imaging setups are observed. For FFPE samples, the tissue block surface is imaged directly as shown in Fig. 1(a). The setup in Fig. 1(b) is used in fresh tissue samples, where a polystyrene plate is placed between the tissue and the system to

provide a consistent imaging plane and to prevent fluid leakage into the system optics. A second polystyrene plate is placed on top of the tissue to provide light pressure to keep the tissue interface flat with the imaging window. The incident THz time-domain signal is shown in Fig. 1(c), and the Fourier transform of the signal is shown in Fig. 1(d), demonstrating a frequency range from 0.1 to 4 THz. To generate the THz image, the stage holding the sample is scanned in steps (smallest step is 50 μm with 200 μm used in all results in this work). The reflected THz pulse is measured at each pixel. For FFPE samples, the THz image is obtained from the peak value of the reflected signal normalized against the incident signal peak obtained using a gold mirror. For fresh tissue samples, better image clarity was observed by taking the spectral power across a selected range, f_1 to f_2 , as¹¹

$$\text{spectral power} = \int_{f_1}^{f_2} |E_{\text{samp}}(f_{\text{THz}})|^2 / |E_{\text{ref}}(f_{\text{THz}})|^2 df_{\text{THz}}, \quad (1)$$

where E_{samp} is the magnitude of the reflected sample signal following Fourier transform, E_{ref} is a reference signal from the polystyrene plate interface with air at the same plane as the fresh tissue (to keep the signal focus the same), and f_{THz} is the frequency in THz. This calculation is only used for fresh tissue because the FFPE tissue is low-absorption and, therefore, subject to multiple reflections,²⁹ which can cause frequency-domain oscillations.

2.2 Mice Tumor Sample Preparation

E0771 murine breast adenocarcinoma cells were grown in RPMI 1640 media supplemented with 10% fetal bovine serum, 1% L-glutamine, and 1% penicillin-streptomycin. These cells were kept in a humidified incubator (5% CO_2 and 37°C) and cultured when the cells reached 75% to 90% confluence. This culturing involved passaging the cells by collecting and redistributing them into new Petri dishes to prevent overcrowding and cell death. E0771 cells in this work were injected within the first 10 passages to prevent any deterioration in cell viability. For example, one mouse could be injected with cells after three passages, whereas another could be injected after nine passages, depending on cell availability when each mouse reached its target weight. A group of 10 C57BL/6J mice were maintained on a high-fat diet (D12492 from Research Diets, Inc.) until they reached a target weight of 35 g.¹⁶ The mice were then injected in the flank with a subcutaneous bolus (3 million cells suspended in serum- and media-free saline) of E0771 murine breast cancer cells to grow tumor xenografts. Once tumors reached a diameter of 1 cm, tumors and adjacent fat were excised from the mice under anesthesia for THz imaging.

Excised samples were transferred in phosphate-buffered saline for transfer from the excision site to the THz system and imaged within an hour of excision. Tumors were bisected to have a cross-section with both cancer and fat at a flat surface for imaging. Following THz imaging, the samples were placed in 10% buffered formalin and shipped to the Oklahoma Disease Diagnostic Laboratory for histopathology assessment.

Of the 10 injected mice, 9 grew sufficient tumors for this study while 1 mouse did not grow any tumors even after

multiple injections. For the first four tumors, only one half of the bisected tumor was imaged as a fresh sample, but as handling improved, the latter five tumors had both bisected halves imaged, designated as sections A and B in the results presented in Sec. 3. Some bisections were discarded, for example, the sample designated as 6A was found to have an excessive amount of deterioration in the pathology assessment and was, therefore, not investigated further. This leaves a total of 13 samples used in this study. All animals received care in compliance with the guidelines outlined in the Guide for the Care and Use of Laboratory Animals. The procedures were approved by the University of Arkansas Institutional Animal Care and Use Committee.

2.3 Correlation Process of THz and Pathology Images

The goal of this work is to quantitatively assess the accuracy of THz images with respect to the pathology results. A pixel-by-pixel comparison is proposed for the generation of receiver operating characteristic (ROC) curves.⁷ Since a pathology photograph has inherently higher resolution than the THz imaging, two processes are implemented to compare the two images. First, a pathology morphing algorithm is used to digitize the histopathology slides and generate a pathology classification at the same resolution and orientation as the THz image, known as the morphed pathology mask. Second, a statistical model is implemented on the THz image to generate a probability-based classification of the different tissue regions in the sample. These two processes are then combined to obtain pixel-by-pixel correlation. The details of these schemes are given in the following sections.

2.3.1 Pathology morphing and pathology mask generation

To make pixel-by-pixel comparison possible, the image morphing in this work reshapes a pathology image according to the external contour of the THz image to match their external shapes.¹⁷ Although this technique is often used to create a sequence of intermediate images between the source and the target, in this case, it is used only to match the pathology to the external contour of the THz image. The morphing algorithm is performed in MATLAB[®] using the following five steps, which are demonstrated in Fig. 2: (i) masking: a THz image mask [Fig. 2(b)] is established using the external contour of the original THz image [Fig. 2(a)]. The algorithm uses only the external contour so that the morphed pathology is not affected by features in the THz image. (ii) Classification: low-power microscope images of the pathology slide are stitched using a panorama editor, and the stitched image is converted into the hue, saturation, and value color model [Fig. 2(c)]. Fat tissue is identified by differences in saturation while fibrous, muscle, and cancer regions are identified by different hue thresholds scaled to the brightness and contrast of the original pathology image. Each identified region is then assigned a value for the pathology results [Fig. 2(d)]. (iii) Rotation: to account for differences in orientation between the THz image and pathology, the pathology mask is temporarily assigned values of 0 (outside) and 1 (inside) and is rotated in 1-deg iterations. For each rotation, the edges of the pathology and the THz masks are cropped, and the cropped pathology resolution is temporarily reduced to that of the THz mask. The cross correlation between the two images is performed using the sum-product

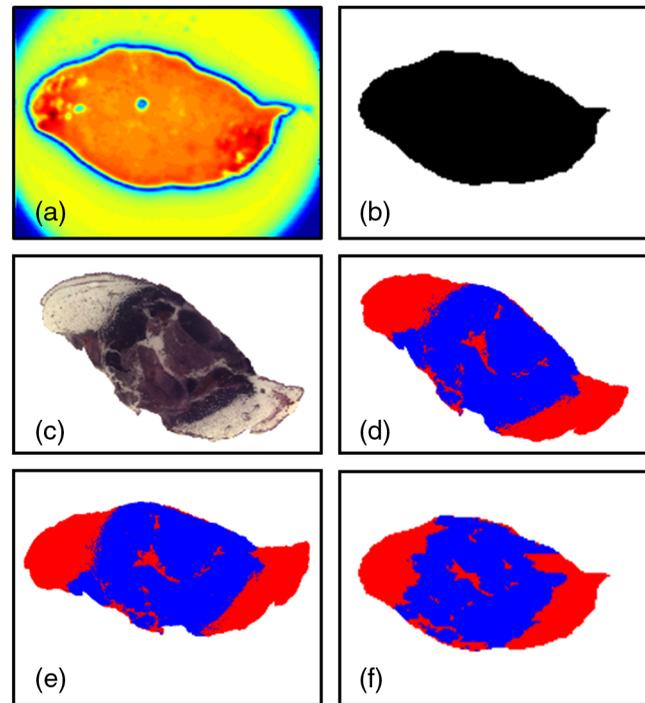


Fig 2. Pathology morphing process for mouse tumor 3: (a) THz reflection image of fresh tissue, (b) resulting THz image mask, (c) the low-power pathology image, (d) pathology mask, (e) rotated pathology mask, and (f) final morphed pathology mask after resizing and reshaping with respect to the image mask in (b).

operation,³³ and the original pathology mask is rotated to the angle with the highest cross correlation [Fig. 2(e)]. (iv) Resizing: the pathology mask is downsampled to match the THz mask resolution. (v) Reshaping: a reshaping operation is performed on each row or column of the pathology mask using cubic spline interpolation for stretching or downsampling for shrinking until it matches the dimensions of THz mask [Fig. 2(f)].³⁴

2.3.2 Statistical Bayesian mixture model

A statistical Bayesian mixture model is used to classify the THz scan data in a way that it can be compared to pathology. This method produces a vector of probabilities that each pixel in the THz image belongs to a region of tissue. This method provides a twofold advantage over simple thresholding techniques: (i) the pixels that would be borderline between two different tissue regions can be represented by probability between those two regions and (ii) a decision rule can be applied to classify each pixel into individual regions. This work uses a modal class decision rule simply based on which region has the largest probability for each pixel, but additional decision rules can be applied without the need to modify the probability model. This work explores both unsupervised (e.g., Gaussian and *t*-distribution) and supervised (e.g., regression model) approaches.

The unsupervised approach considers the pixel-wise summarized intensities of the THz image as random variables that are independent but not identically distributed. Pixels from the same tissue region are assumed to follow the same probability distribution while pixels of different regions may have different probability distributions. Hence, the probability distribution for any pixel in the image can be thought of as a mixture model, a

weighted sum of parametric probability distributions or mixture components. This work utilizes a Bayesian framework for implementing mixture models. An MCMC scheme is used to generate a large number of samples from the posterior joint distribution of mixture parameters and compute empirical summaries for these parameters.²⁸ To carry out the MCMC, the data augmentation technique in Ref. 35 is used. For a THz image of n pixels with x_i denoting the intensity of pixel i for $i = 1, 2, \dots, n$, a hierarchical structure for mixture distribution for x_i is proposed

$$[x_i | \theta, z_i] \stackrel{\text{ind}}{\sim} f(\theta_{z_i}), \quad z_i \sim \text{Mult}(q_1, q_2, \dots, q_k),$$

$$\pi(\theta) = \prod_{j=1}^k \pi(\theta_j), \quad \pi(q_{1:k}) = \text{Dir}(\alpha_{1:k}), \quad (2)$$

where z_i is a latent variable corresponding to pixel i that would indicate the tissue region of the pixel. If the image has k distinct tissue regions, $z_i \in \{1, 2, \dots, k\}$. The marginal model for x integrating out the z is a k -component mixture model with mixture weights given by $q_{1:k}$. Mult and Dir are the abbreviations for multinomial and Dirichlet probability distributions, respectively. θ_j , the parameters for the probability distribution (mean, standard deviation, etc.) for pixels from region j , has a prior distribution $\pi(\theta_j)$ for $j = 1, 2, \dots, k$. $\alpha_{1:k}$ denotes the parameters in the Dirichlet prior for $q_{1:k}$. One can produce different types of mixture model by altering the choice of the family of distributions for f and the associated set of parameters θ . The two specifications of f that used in this work are described below. Our first proposal is a Gaussian mixture model (normal distribution):

$$[x_i | z_i = j] = \mathcal{N}(\mu_j, \sigma_j^2), \quad \mu_j \sim \mathcal{N}(\mu_{0j}, c_{0j}), \quad \sigma_j^2 \sim \frac{1}{\sigma_j^2}, \quad (3)$$

where μ_j and σ_j are the location and scale parameters of the normal-distribution Gaussian function \mathcal{N} for component j . μ_{0j} and c_{0j} are the hyperparameters and assumed to be known. In absence of any reliable prior information, one can arbitrarily choose the c_{0j} to be large to diffuse the prior probability. The prior choice for σ_j^2 is known as the noninformative prior and is improper as well. Our second proposal is a t -mixture model (t -distribution)

$$[x_i | z_i = j] = t_{d_j}(\mu_j, \sigma_j^2), \quad \mu_j \sim \mathcal{N}(\mu_{0j}, c_{0j}), \quad \sigma_j^2 \sim \frac{1}{\sigma_j^2}, \quad (4)$$

where $d_j \in \mathbb{N}$, the degrees of freedom for the j 'th component, characterizes the heaviness of the tail. A small value for d_j will produce a heavy tailed t -density curve whereas higher values of d_j will make the tails lighter and will eventually approach Gaussian tails. Hence, instead of using fixed d_j values, the model can be more flexible by learning the possible values of d_j from the data. However, employing the Gibbs sampling technique directly on a t -mixture density is difficult, so the t -density is instead viewed as a scale mixture of Gaussian density.³⁶ Here, the Gaussian and t -mixture models from Eqs. (3) and (4) are employed on each sample. Additionally, skewed mixture models were also investigated as in Ref. 37,

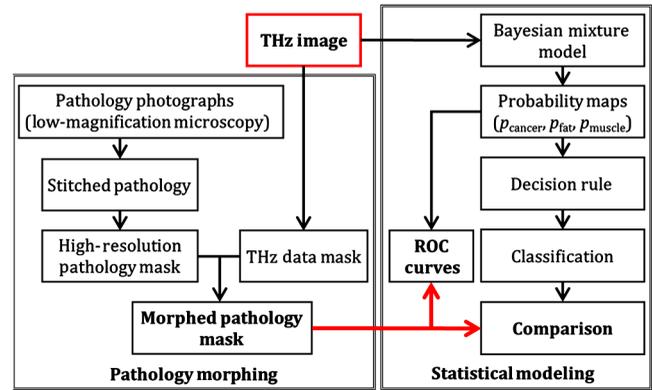


Fig 3. Correlation process flowchart for comparing THz images to pathology.

but the implementation of these models is more complex, and no case was found where they outperform the nonskewed models.

Following each iteration of the model, posterior updates of μ_j , σ_j^2 , $q_{1:k}$, and z_i (and d_j for t -mixture) were applied prior to the next iteration. Post-MCMC, posterior draws were collected for each of z_1, z_2, \dots, z_n . The probability vectors p_i representing the classification uncertainty are empirically calculated as

$$p_i(j) = \frac{\#z_i = j}{\#\text{Iterations}}, \quad i = 1, 2, \dots, n \text{ and } j = 1, 2, \dots, k. \quad (5)$$

For this work, the number of iterations was set at 20,000, and only the latter 10,000 were considered for calculating the probability vectors to let the model converge. One important point to note is that the mixture model for the data as mentioned in Eq. (2) stays the same for any permutation of $\{\theta_1, \theta_2, \dots, \theta_k\}$, i.e., there is no natural ordering between mixture components. Thus, they must be labeled according to some criterion.³⁸ Here, the mean intensity of the included points in each component is used to number them.

Three assumptions are required for this approach to work: (i) the value of k , the number of regions present in a tumor, is assumed prior to running the MCMC algorithm, (ii) the ordering of different regions in terms of increasing intensity values is also assumed, and (iii) each component has a unimodal distribution. However, a multimodal distribution for a region will work as well if the number of modes is known.

For cases where the assumptions for the classification model may not be met, a supervised stochastic learning model is investigated, which builds the model utilizing one or more training THz images and their corresponding pathologies.³⁹ For any pixel, the classification from pathology is considered a categorical response corresponding to the pixel intensity. As there is prior knowledge about the ordering of different regions based on pixel intensities, the problem is treated as an ordinal regression by numbering the categories in ascending or descending order. In a Bayesian framework, the data augmentation approach of Ref. 39 is used.

The correlation schemes discussed above are summarized in the flowchart of Fig. 3. The morphed pathology is used as the reference point for the statistical model in two ways. The first method provides a visual comparison by taking the

Table 1 Selected mice samples.

Mouse No.	2	4	7A	7B	8A	8B	9A
Details	Two regions: cancer and fat	Three regions: cancer, muscle, and fat	Three regions: cancer, muscle, and fat				

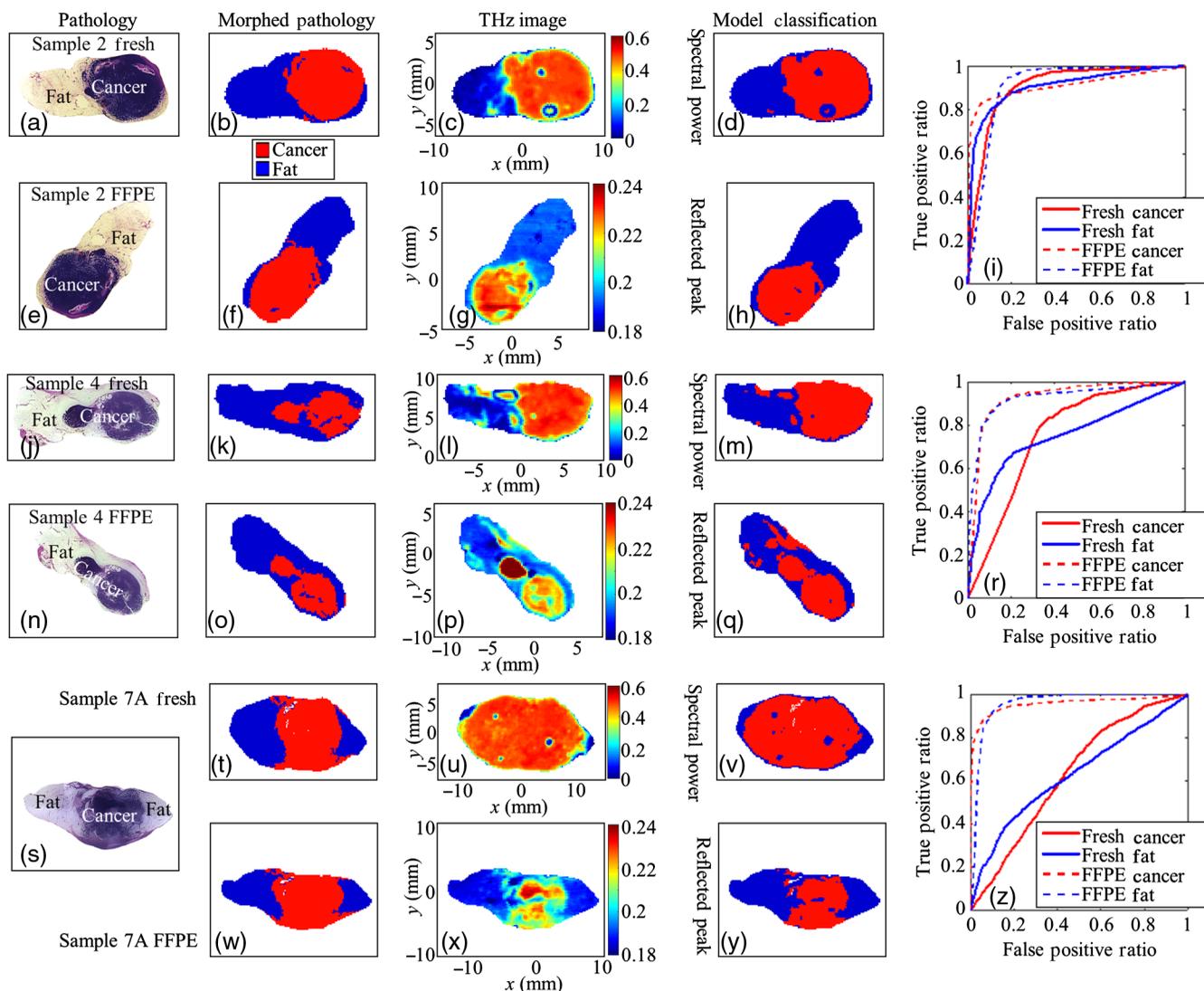


Fig 4. Correlation results for tumor samples 2, 4, and 7A. For tumor sample 2, images (b–d) for freshly excised tumor tissue and images (f–h) for FFPE tissue: (a) pathology image, (b) morphed pathology mask, (c) THz image, and (d) t -distribution model classification; (e) pathology image [same as in (a)], (f) morphed pathology mask, (g) THz image, and (h) t -distribution model classification; and (i) the ROC curves. For tumor sample 4, images (k–m) for fresh tissue and images (o–q) for FFPE tissue: (j) pathology image, (k) morphed pathology mask, (l) THz image, and (m) t -distribution model classification; (n) pathology image, (o) morphed pathology mask, (p) THz image, and (q) t -distribution model classification; and (r) the ROC curves. For tumor sample 7A, images (t–v) for fresh tissue and images (w–y) for FFPE tissue: (s) pathology image, (t) morphed pathology mask, (u) THz image, and (v) t -distribution model classification; (w) morphed pathology mask, (x) THz image, and (y) t -distribution model classification; and (z) the ROC curves.

classification image following the decision rule applied to the probability maps of the model. The second method plots an ROC curve of the true and false positive ratios for different threshold values assigned to the probability maps. For some

probability threshold of a given map (e.g., cancer or fat), a number of pixels will be assigned to the region in question. For example, the true positive ratio of the cancer region indicates the ratio between the pixels correctly assigned as cancer and

the total number of pixels actually belonging to cancer in the morphed pathology. As another example, the false positive ratio of fat is the ratio between the pixels incorrectly assigned as fat and the total number of nonfat pixels in the pathology. A statistically effective technique should achieve a relatively high true positive ratio while maintaining a low false positive ratio.

3 Experimental Results

3.1 Classification of Samples with Two Tissue Regions

In this work, 13 samples obtained from 9 murine xenograft breast cancer tumors are handled. Seven mice samples are selected here to compare the THz imaging, morphed pathology, and statistical model results, as summarized in Table 1.

The images of fresh tissue samples are acquired as well as the images of the FFPE samples of the same tissues following histopathology. The results from three samples (samples 2, 4, and 7A) are shown in Fig. 4. The results of the fresh and FFPE tissues are presented for mouse tumor 2 in Figs. 4(a)–4(i), for mouse tumor 4 in Figs. 4(j)–4(r), and for mouse tumor 7A in Figs. 4(s)–4(z). The first row of each tumor sample corresponds to the fresh tissue, where Fig. 4(a) shows the pathology for mouse tumor 2 (rotated to the fresh tissue orientation) with regions of fat and cancer indicated. The morphed pathology mask in Fig. 4(b) shows the digitized assignment of the two regions with red color designating cancer and blue color designating fat. The THz reflected spectral power image calculated using Eq. (1) with $f_1 = 0.5$ THz and $f_2 = 1$ THz in all results is shown in Fig. 4(c). Here, it can be seen that the cancer region shows a distinctly high reflection compared to the fat region. The result of the statistical model is shown in Fig. 4(d) for the t -distribution and is observed to correlate well to both the THz and the pathology images. The same process is shown for the FFPE tissue with the morphed pathology in Fig. 4(f), THz time-domain peak image in Fig. 4(g), and the model-based classification in Fig. 4(h). Here, the THz image can be seen to closely agree with the pathology, which is supported by the similarity between the morphed pathology and statistical model. The ROC curves for both fresh and FFPE tissue cases are given in Fig. 4(i). Here, it can be seen that the detection of cancer and fat regions is very good for both cases of tumor sample 2, with a greater area under the ROC curve indicating better correlation.

Similarly, good indication of the THz imaging can be seen for tumor sample 4. Figures 4(k)–4(m) show the results for the fresh tissue imaging. Here, it should be noted that the pathology in Fig. 4(j) and mask in Fig. 4(k) reveal two primary regions of cancer: a larger area with cell density similar to other tumor cancer regions in this work and a smaller area [appearing as a circle in Fig. 4(j)] with very dense cancer cells undergoing rapid growth (based on pathology assessment). This second cancer region is not seen in the fresh tissue image of Fig. 4(l) primarily due to the histopathology process that slices into the tissue 100 μm or more to get a good cross section of the tissue. This difference between the imaging surface and the slice taken for pathology can result in different features being visible between fresh tissue and pathology. The FFPE tissue images in Figs. 4(n)–4(q) show this second area of cancer clearly, with particularly high reflection due to the high density of the cancer cells. It should be noted that the pathology slice is taken from the same surface at which the FFPE tissue scan is

performed. The ROC curves in Fig. 4(r) show good detection from the THz imaging of the FFPE tissue, with slightly less detection in the fresh tissue due to the difference in imaging surface discussed above.

For tumor sample 7A, both fresh and FFPE tissues had similar orientation during imaging, so a single pathology image in Fig. 4(s) shows the central cancer tissue with fat on each side. In this case, the morphed pathology in Fig. 4(t) designates much larger areas of fat than is seen in the THz image of the fresh tissue in Fig. 4(u) and subsequent classification in Fig. 4(v). For this sample, there was excess fluid pooled under the tissue when it was freshly excised. Additionally, the tumor had grown large enough to undergo some necrosis in the center, leading to the loose cellular tissue and blood being spread across the tissue surface when bisecting the tumor. As a result, a higher reflection from this liquid was spread over the areas that would normally be fat. On the other hand, the images from the FFPE tissue in Fig. 4(x) and classification in Fig. 4(y) show much better agreement with the pathology mask in Fig. 4(w), because this fluid was removed in the process. Figure 4(x) of THz image also shows the varying reflections in the cancer region, with more dense cancer tissue reflecting more, agreeing with the pathology in Fig. 4(s). Though due to some necrosis, captured in the THz image, the correlation with morphed pathology is not as consistent as in tumors 2 and 4. Figure 4(z) shows that while the classification of the FFPE tissue is accurate, the fresh tissue poses a challenge due to the excess fluid. Note that THz images show clear visual differentiation among tissue regions especially in the FFPE cases, in agreement with previous work.^{9,29–30}

Another challenge is that some tissue samples showed more significant shape distortion between THz images of fresh tissue and pathology fixation. Examples of this distortion are shown in Fig. 5 for tumor sample 7B [Figs. 5(a)–5(h)] and tumor sample 8A [Figs. 5(i)–5(p)]. For sample 7B, the tissue not only has some distortion due to the histopathology process but is also seen to have several gaps as detailed in Fig. 5(a). These gaps primarily arise due to the necrosis on the interior of larger tumors that fill with fluid during the fresh tissue imaging and paraffin for the FFPE tissue. These gaps are considered a separate region during pathology morphing to maintain their size and shape but considered background for the sake of correlation such that the points are disregarded from comparison. The morphed pathology for the fresh tissue outline in Fig. 5(b) shows these gaps clearly in the cancer, with some small regions of fat in the lower corners. These gaps are also imposed on the model results in each case to exclude those points from comparison. Shape distortion between the fresh tissue and pathology can be seen in the THz image and model classification in Figs. 5(c) and 5(d), where the lower reflection regions in the corners of the tissue appear approximately the same size as the fat regions in the pathology but slightly shifted. This results in some challenge for lining up the pathology with the fresh tissue images. In contrast, for FFPE tissue, the THz image in Fig. 5(f) shows very good agreement with the morphed pathology in Fig. 5(e) and the model classification in Fig. 5(g) as well. The THz image of the FFPE tissue in Fig. 5(f) also shows the gaps of the pathology as areas of low reflection due to filling the gaps with paraffin. The image demonstrates good comparison with the original pathology in Fig. 5(a). This is due to the pathology being taken from the surface of the FFPE tissue block as discussed previously. The ROC curves in Fig. 5(h) then show that the THz imaging of the FFPE tissue has good detection

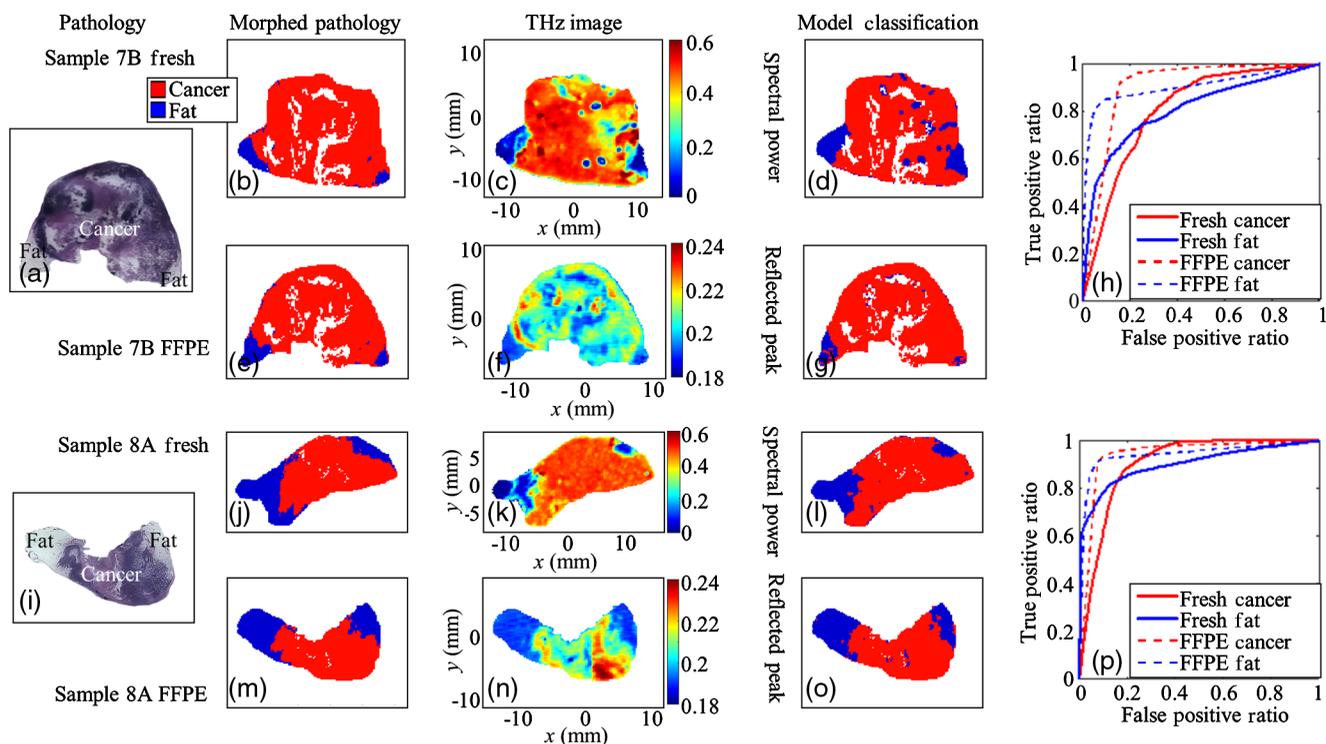


Fig 5. Correlation results for samples 7B and 8A. For tumor sample 7B, images (b–d) for freshly excised tumor and images (e–g) for FFPE tissue: (a) pathology image, (b) morphed pathology mask, (c) THz image, and (d) t -distribution model classification; (e) morphed pathology mask, (f) THz image, and (g) t -distribution model classification; and (h) the ROC curves. For tumor sample 8A, images (j–l) for freshly excised tissue and images (m–o) for FFPE tissue: (i) pathology image, the (j) morphed pathology mask, (k) THz image, and (l) t -distribution model classification; (m) morphed pathology mask, (n) THz image, and (o) t -distribution model classification; and (p) the ROC curves.

of tissue regions while the fresh tissue imaging classification is diminished due to the tissue distortion.

The ability for the current image morphing to account for shape change in the tissue is dependent on the specific tumor sample. As seen for tumor sample 8A, the pathology in Fig. 5(i) undergoes a significant change in shape compared to the morphed pathology for the fresh tissue in Fig. 5(j). However, the THz image in Fig. 5(k) shows good agreement with the pathology mask in Fig. 5(j) and with the statistical model in Fig. 5(l). Meanwhile, the results for the FFPE tissue in Figs. 5(m)–5(o) show good correlation similar to other FFPE tissue cases. The ROC curves in Fig. 5(p) indicate that both fresh and FFPE tissues have reasonable detection for the two regions of tissue.

For comparing the results across several samples, there are a few ways to quickly describe the ROC curve. These include the 5% or 10% false positive sensitivity (the true positive ratio when the false positive ratio is 0.05 or 0.1) or the area under the curve. To relate these two values, an ROC curve with a true positive

ratio of 0.9 when the false positive ratio is 0.1 would generally have an area under the curve of 0.8 to 0.9, so an ROC area above 0.8 would be considered good correlation. The ROC area for all tumor samples with two regions is shown in Fig. 6. The FFPE tissue THz imaging can be seen to have a curve area above 0.8 in all cases, which is expected due to the pathology being taken from the exact surface imaged in the FFPE tissue blocks. THz imaging of fresh tissue can also be seen to have relatively good detection in most cases. There were some cases where the detection was lower, such as in tumor sample 1 where the tissue had been frozen. This was done in an optimal cutting temperature (OCT) medium that could not be cleared for imaging. Other cases of low detection were tumor samples 3 and 7A, where significant fluid had accumulated under the tissue, and tumor sample 4, where there was a difference in the scanned surface between the pathology and fresh tissue image (see Fig. 4). In general, THz imaging shows good detection between distinct regions of fat and cancer tissue in the freshly excised mice tumors, and the challenges in cases with lower correlation

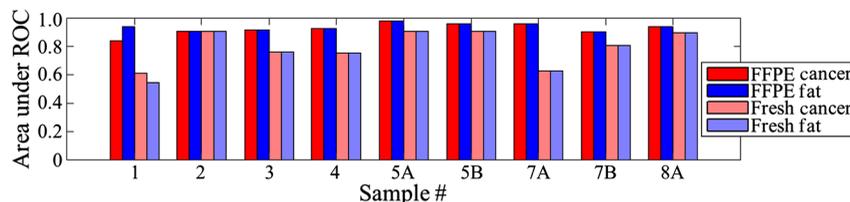


Fig 6. Area under ROC curve for nine samples with two regions.

are clearly identified. For most cases, the tissue morphing was able to resolve differences in tissue shape between the fresh tissue and pathology, even for the severe case in sample 8A, though not all cases were corrected and there is potential to improve the pathology correlation with a more robust morphing algorithm.

3.2 Classification of Samples with Three Tissue Regions

The excision of some larger tumors included muscle tissue from the abdominal wall of the mouse. Muscle is not anticipated in human breast cancer excisions but arose in these samples due to the limited space for the tumors to grow in the fat deposits of the mice. As such they are examined here to test the statistical model. In this section, results of unsupervised (nonregression) and supervised (regression) approaches are shown in Fig. 7. The regression model used data from tumor sample 9B as an arbitrary training sample. The training used the regions defined in the morphed pathology to collect intensity distributions for each tissue type for building the model. The images of sample 9B are not shown here but its statistics are shown later in Fig. 11.

For tumor sample 8B, the pathology in Fig. 7(a) shows that the cancer is mostly along the center and the upper right part of the tissue. This sample was unique in that the center of the cancer had mostly fatty tissue with some cancer mixed in. The other fat deposits and muscle can be seen on the lower edge. These regions translate more or less directly into the morphed pathology in Fig. 7(b), with a few small gaps represented in white color that are imposed to the morphed pathology and model classification. However, for the THz image of the fresh tissue in Fig. 7(c), the fat region at the core of the tissue cannot be seen, and there is a spot of very high reflection over the fat deposit on the left side. The latter is most likely fluid pooled beneath the fat, and the inability to see the cancer distributed through the fat could be from fluid and loose cells being distributed on the surface during bisection. There is also a possibility that the surface changed significantly between the fresh THz image and pathology. The classification model results using the normal distribution in Fig. 7(d) show a generally good assignment of fat and cancer regions, but for muscle regions, it focuses on areas of borderline reflection between high and low regions rather than a separate broad region. The classification image is improved somewhat using the regression model in Fig. 7(e). Here, all higher reflections are

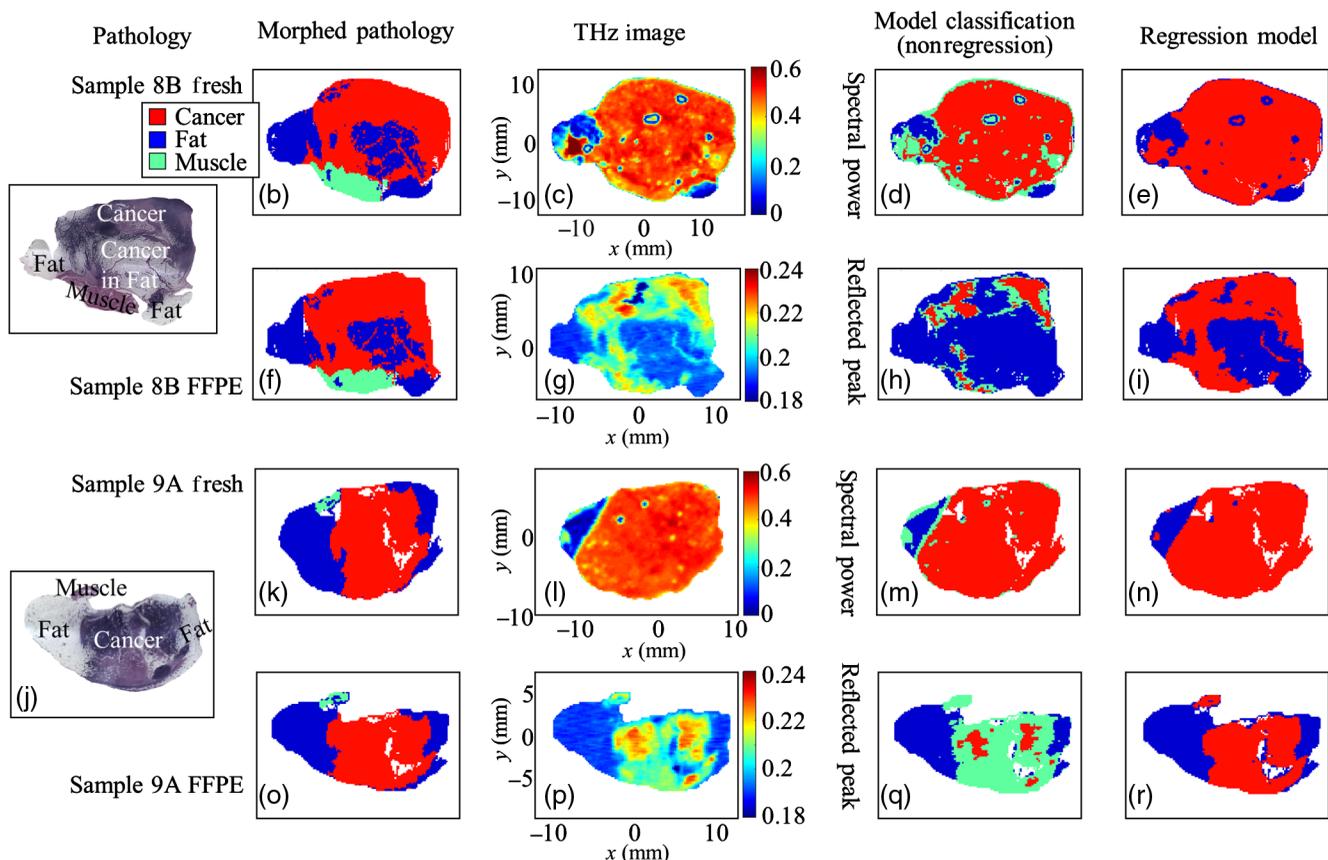


Fig 7. Correlation results for samples 8B and 9A. For tumor sample 8B, images (b–e) for freshly excised tissue and images (f–i) for FFPE tissue: (a) pathology image, (b) morphed pathology mask, (c) THz image, (d) normal-distribution model classification, and (e) regression model classification; and the (f) morphed pathology mask, (g) THz image, (h) normal-distribution model classification, and (i) regression model classification. For tumor sample 9A, images (k–n) for freshly excised tissue and images (o–r) for FFPE tissue: (j) pathology image, (k) morphed pathology mask, (l) THz image, (m) *t*-distribution model classification, and (n) regression model classification; and the (o) morphed pathology mask, (p) THz image, (q) *t*-distribution model classification, and (r) regression model classification.

considered cancer while the region of fat is mostly unchanged. However, the region of muscle tissue in the pathology is still not indicated by the model. Therefore, while there does appear to be some small difference in the THz image along the lower edge where the muscle is in pathology, it is not detected for correlation.

The FFPE tissue results show a similar challenge. The morphed pathology in Fig. 7(f) and THz image in Fig. 7(g) both show very close agreement with the pathology in Fig. 7(a). However, this sample involves both muscle tissue and necrotic cancer tissue (seen as faded areas in the cancer in pathology or low-reflection spots in the cancer in the THz image), which both have reflection in between the cancer and fat and tend to overlap. The result in Fig. 7(h) indicates that the classification model classified almost all of the lower reflection areas as fat, with only the spots of highest reflection being classified as cancer and the small amount of midrange values considered as muscle. The regression model in Fig. 7(i) shows significant improvement in this case with accurate detection of the fat and mostly accurate detection of cancer, though once again the muscle tissue is not classified by the model.

Similar challenges are seen for tumor sample 9A, as shown in Figs. 7(j)–7(r). In this sample, the region of muscle is relatively small in the upper left part of the pathology image in Fig. 7(j), leaving primarily cancer in the center and fat on the left and right sides. The morphed pathology in Fig. 7(k) is consistent with the pathology image in Fig. 7(j). Meanwhile, the THz image of the fresh tissue in Fig. 7(l) once again shows the challenge of clearing fluid from under the tissue. Here, the high reflection of fluid in the tissue decreases the expected area of fat on the left and completely covers the fat on the right. As with the previous case, the classification (nonregression) model in Fig. 7(m) classifies the tissue regions of high and low reflection as cancer and fat, respectively, whereas the transition tissue region between them

is classified as the muscle tissue. The regression model in Fig. 7(n) did not improve the results of Fig. 7(m). The FFPE tissue results in Fig. 7(o)–7(r) are similar to tumor sample 8A. Here, the reflection of the muscle and some of the necrotic cancer tissue overlap in the THz image in Fig. 7(p), resulting in most of the cancer region being classified as muscle in the classification (nonregression) model results in Fig. 7(q). The regression model in Fig. 7(r) correctly defines most of the cancer region but does not classify any noticeable muscle region. In both samples shown here, the THz images show differentiation in heterogeneous regions of tissue that can be directly compared to the original pathology. However, this heterogeneity is lost when reducing the morphed pathology and classification model to just three regions.

The ROC curves in Fig. 8 show the challenge for the current statistical model to classify the tissue when three regions are present (cancer, fat, and muscle). For the fresh tissue images of tumor sample 8B, the ROC curves in Fig. 8(a) show relatively lower areas under the curves (0.6551 for cancer, 0.5967 for muscle, and 0.6244 for fat). This is due to the fact that even if a classification image meets the decision rule for the three regions, the certainty of the classified tissue type in the probability maps of the model may still be low. Meanwhile, the FFPE tissue results in the same figure show that the cancer and fat are both reasonably detected while the muscle is not, which agrees with the visual correlation of the tissue. Using the regression model for sample 8B provides the ROC curves in Fig. 8(b). The regression model is shown to not significantly change the results of the fresh tissue (ROC areas of 0.616 for cancer, 0.6276 for muscle, and 0.6061 for fat), while it improves the classification of cancer and fat in the FFPE (increased from 0.7507 to 0.7793 for cancer and 0.796 to 0.8238 for fat). The classification of muscle tissue is decreased (down from 0.5849 to 0.4624). For tumor sample 9A, the ROC curves in

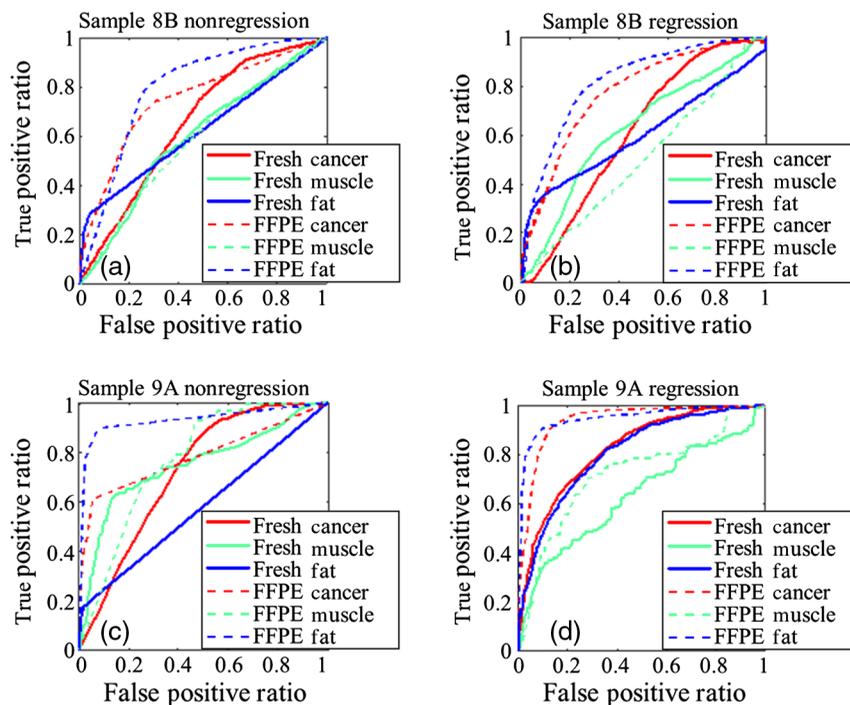


Fig 8. ROC curves for sample 8B using (a) normal-distribution model classification and (b) regression model classification, and for sample 9A, using (c) *t*-distribution model classification and (d) regression model classification.

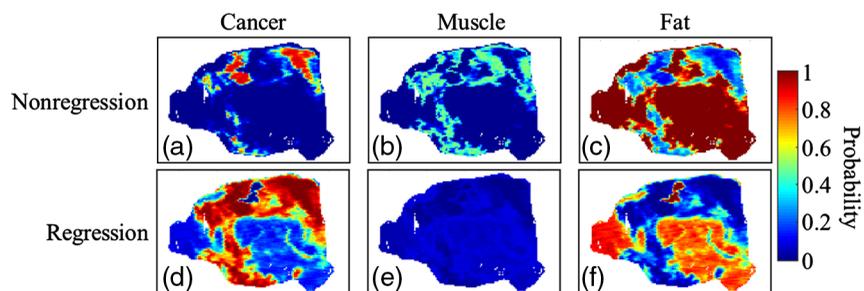


Fig 9. Probability maps generated by the statistical model for sample 8B FFPE. Maps using normal distribution for regions of (a) cancer, (b) muscle, and (c) fat and using regression model for regions of (d) cancer, (e) muscle, and (f) fat.

Fig. 8(c) denote some effective detection of the tissue regions in the fresh tissue but with low certainty, whereas the FFPE tissue has slightly better detection. The regression model results in Fig. 8(d) show notable improvement in the cancer and fat tissue regions for the fresh tissue imaging (areas increased from 0.7253 to 0.8272 for cancer and 0.5808 to 0.8121) with some decrease in muscle classification (down from 0.7512 to 0.6365). For FFPE tissue, the classification for cancer increases significantly (up from 0.7874 to 0.9431) while classification of muscle decreases (ROC area down from 0.7762 to 0.7266) and fat marginally increases (from 0.9318 to 0.9531). As such, while the regression model improves some cases its usefulness is not consistent across all samples.

In all implementations of the regression model observed here, the classification of cancer and fat tended to increase or stay the same while the classification of muscle tended to decrease. This is investigated more directly by observing the probability maps for sample 8B in Fig. 9. Although these probability maps are generated for every sample, only one is shown here for the sake of space. For each image, the regions with higher values (dark red) indicate a larger probability that pixel being in the described region. For the normal-distribution model in Figs. 9(a)–9(c), a high probability of cancer [Fig. 9(a)] is assigned to the areas of highest reflection in the earlier THz image, whereas an increased probability of muscle [Fig. 9(b)] is given to the next highest reflections from muscle and necrotic cancer. It should be noted that the probability of muscle here is not particularly high, but the regions of higher values in Fig. 9(b) still possess a greater probability of muscle than cancer or fat and are, therefore, assigned as such in the classification image. Meanwhile, the probability of fat in Fig. 9(c) shows good indication of the fat regions. By informing the regression model with existing data of cancer, muscle, and fatty tissue, the probability maps in Figs. 9(d)–9(f) are generated. Here, the full region of cancer is now shown to have a high probability for cancer in Fig. 9(d), lending to the improvement seen in the ROC area of Fig. 8(d). However, there is very little probability of muscle anywhere in Fig. 9(e). This may be due to using only one image for training the model, the muscle in all cases being small, and/or the refractive index and absorption coefficients of muscle are close to cancer tissue in THz frequency (Fig. 10).^{40,41} This same effect is the cause for the decreased ROC area for muscle in the samples in Fig. 8. Meanwhile, the fat classification in Fig. 9(f) remains mostly unchanged.

The areas under the ROC curves for the four samples with three tissue regions are given in Fig. 11. In all cases, the FFPE tissue continued to show good classification of cancer and fat areas, though the presence of the muscle region did make the

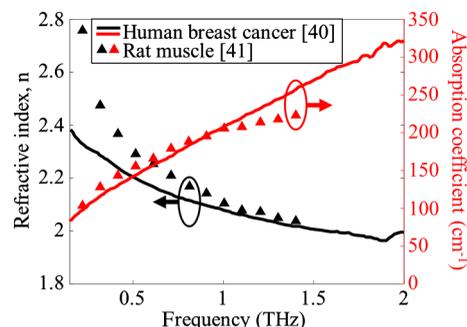


Fig 10. Published refractive index and absorption coefficient values for human breast cancer⁴⁰ and rat muscle.⁴¹

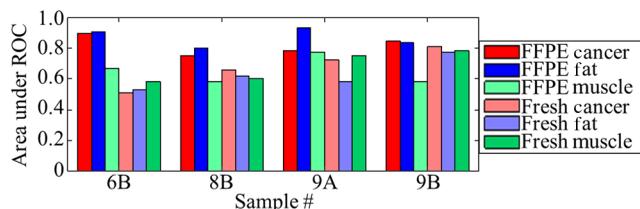


Fig 11. Area under ROC curve for four samples with three regions.

detection of each region less accurate, where the detection of muscle in all cases was the lowest. For fresh tissue, all samples investigated showed some challenge in detecting the three regions except for sample 9B where all ROC areas neared 0.8 except for the detection of muscle in the FFPE tissue (images not shown here). The accuracy across all fresh tissue samples was primarily affected by necrosis inside the larger tumors, which caused higher reflections in the THz imaging due to excess fluid and loose tissue.

A summary of the areas under the ROC curves for all cases discussed in Figs. 6 and 11 is given in Table 2. Here, the strong correlation can be shown in all FFPE cases, except the cases that involved muscle or moderate to advanced necrosis. The primary challenge in fresh tissue is clearing the fluid from the imaging surface, which is often compounded in cases with necrosis.

4 Discussion and Conclusion

Over the 13 samples handled in this work, THz images are seen to have statistically good discrimination between tissues with two regions (cancer and fat), with a few exceptions due to factors outside of imaging performance. The primary challenges in classifying THz images of tissue in this work are (i) the presence

Table 2 Area under ROC curve results for all cases in Figs. 6 and 11.

Mouse No.	FFPE			Fresh			Notes
	Cancer	Fat	Muscle	Cancer	Fat	Muscle	
1	0.84	0.94	N/A	0.61	0.55	N/A	Frozen in OCT before fresh imaging
2	0.91	0.91	N/A	0.91	0.91	N/A	—
3	0.92	0.92	N/A	0.76	0.76	N/A	Excess fluid and minor necrosis
4	0.93	0.93	N/A	0.76	0.76	N/A	Difference in pathology surface
5A	0.98	0.98	N/A	0.91	0.91	N/A	—
5B	0.96	0.96	N/A	0.91	0.91	N/A	—
6B	0.90	0.91	0.67	0.51	0.53	0.58	Excess fluid and minor necrosis
7A	0.96	0.96	N/A	0.63	0.63	N/A	Excess fluid and minor necrosis
7B	0.90	0.90	N/A	0.81	0.81	N/A	Advanced necrosis
8A	0.94	0.94	N/A	0.90	0.90	N/A	Significant shape change in pathology
8B	0.75	0.80	0.58	0.66	0.62	0.60	Excess fluid and advanced necrosis
9A	0.79	0.93	0.78	0.73	0.58	0.75	Excess fluid and moderate necrosis
9B	0.85	0.84	0.58	0.81	0.78	0.78	Moderate necrosis

of fluids, including water and necrotic blood, on the underside of the fresh tissue, (ii) the correlation with pathology obtained after fixing the tissue in formalin and embedding it in paraffin, and (iii) the morphed pathology (digitized pathology) using an interpolation algorithm to reduce the resolution of the original pathology. The presence of fluid occurred to some extent in samples 3 and 7 and in all three-tissue region samples except sample 9B. This resulted in some cases of fat tissue being classified as cancer in the statistical model. In correlating the pathology to the THz images of fresh tissue, a change in tissue shape occurred going through histopathology processing. These concerns can be addressed in future work by drying the tissue thoroughly using a lint-free filter paper, using tissue marking ink for orienting tissue between imaging and pathology, and mounting tissue sections on a rigid surface (i.e., cardboard) for formalin fixation. Additional challenges in shape comparison of THz images of fresh tissue to pathology can be improved with more rigorous morphing techniques, such as adopting a mesh-based morphing (ongoing research) instead of the interpolation used in this work. However, there is still a need for a true comparison against THz imaging of fresh tissue to determine accuracy, and future work will look into other common imaging techniques (e.g., CT, radiography, etc.) to have a direct comparison for fresh tissue imaging.

For the 4 samples where muscle tissue is present, overlap between muscle and cancer tissue reflections in the THz image creates some challenge in correctly classifying these regions. While muscle is unlikely to be present in surgical sections of human breast cancer, other kinds of fibrous tissue may be present and thus it requires investigating more advanced models for three tissue regions. Ongoing research is focusing on spontaneously generated breast cancer tumors from transgenic mice, which have natural tumor structures and fibrous

tissue more comparable to human tissue for more accurate assessment of THz imaging. Another area in which the approach could be improved is the classification model. The statistical models used here show success when handling samples with two regions but tend to miss areas of a third region of tissue. In general, the models presented here can classify three tissue regions if the reflection from the regions is distinct. It was observed in human tumors that the model was not able to fully classify cancer and fibroglandular tissue for the same reason where cancer and fibrous tissue have close properties.⁴² Future work will focus on detection methods and advanced models to address this challenge. The obtained results demonstrate promise for THz imaging of freshly excised tumors and shed the light on the main challenges that need to be resolved before it can be implemented on human tissue.

Disclosures

The authors have no financial interests or conflicts of interests to disclose regarding this paper.

Acknowledgments

This work was funded by the National Institutes of Health Award No. R15CA208798. It was also funded in part by the National Science Foundation (NSF) Award No. 1408007. Funding for the pulsed THz system was obtained through NSF/MRI Award No. 1228958. The authors would like to thank Dr. L. Norian from the University of Alabama, Birmingham, for feedback on the high-fat diet protocol for the mice. The authors would also like to thank the students M. Harper, D. Lee, and K. Alhallak in the Functional Optical Imaging and Spectroscopy laboratory for assistance in cell growth, injection, and tumor excision.

References

- S. Fan et al., "The growth of biomedical terahertz research," *J. Phys. D: Appl. Phys.* **47**, 374009 (2014).
- S. Sy et al., "Terahertz spectroscopy of liver cirrhosis: investigating the origin of contrast," *Phys. Med. Biol.* **55**, 7587–7596 (2010).
- W. E. Baughman et al., "Observation of hydrofluoric acid burns on osseous tissues by means of terahertz spectroscopic imaging," *IEEE J. Biomed. Health Inf.* **17**, 798–805 (2013).
- S. J. Oh et al., "Study of freshly excised brain tissues using terahertz imaging," *Biomed. Opt. Express* **5**, 2837–2842 (2014).
- H. A. Kashanian et al., "Gastric cancer diagnosis using terahertz imaging," *Majlesi J. Multimed. Process.* **4**, 1–7 (2015).
- F. Wahaia et al., "Detection of colon cancer by terahertz techniques," *J. Mol. Struct.* **1006**, 77–82 (2011).
- B. St. Peter et al., "Development and testing of a single frequency terahertz imaging system for breast cancer detection," *IEEE J. Biomed. Health Inf.* **17**, 785–797 (2013).
- T. Bowman, M. El-Shenawee, and L. K. Campbell, "Terahertz transmission vs reflection imaging and model-based characterization for excised breast carcinomas," *Biomed. Opt. Express* **7**, 3756–3783 (2016).
- T. C. Bowman, M. El-Shenawee, and L. K. Campbell, "Terahertz imaging of excised breast tumor tissue on paraffin sections," *IEEE Trans. Antennas Propag.* **63**, 2088–2097 (2015).
- A. J. Fitzgerald et al., "Terahertz pulsed imaging of human breast tumors," *Radiology* **239**, 533–540 (2006).
- C. B. Reid et al., "Terahertz pulsed imaging of freshly excised human colonic tissues," *Phys. Med. Biol.* **56**, 4333–4353 (2011).
- P. Doradla et al., "Detection of colon cancer by continuous-wave terahertz polarization imaging technique," *J. Biomed. Opt.* **18**, 090504 (2013).
- S. Yamaguchi et al., "Brain tumor imaging of rat fresh tissue using terahertz spectroscopy," *Sci. Rep.* **6**, 1–6 (2016).
- C. H. Zhang et al., "Terahertz imaging on subcutaneous tissues and liver inflamed by liver cancer cells," *Terahertz Sci. Technol.* **5**, 114–123 (2012).
- H. Chen et al., "High-sensitivity in vivo THz transmission imaging of early human breast cancer in a subcutaneous xenograft mouse model," *Opt. Express* **19**, 21552–21562 (2011).
- S. K. Boi et al., "Obesity alters immune and metabolic profiles: new insight from obese-resistant mice on high-fat diet," *Obesity* **24**, 2140–2149 (2016).
- M. T. Rahman et al., "A novel approach of image morphing based on pixel transformation," in *Proc. of 2007 10th Int. Conf. on Computer and Information Technology, ICCIT* (2007).
- A. J. Fitzgerald et al., "Classification of terahertz-pulsed imaging data from excised breast tissue," *J. Biomed. Opt.* **17**, 016005 (2012).
- L. H. Eadie et al., "Optimizing multi-dimensional terahertz imaging analysis for colon cancer diagnosis," *Expert Syst. Appl.* **40**, 2043–2050 (2013).
- N. Qi et al., "Terahertz time-domain spectroscopy combined with support vector machines and partial least squares-discriminant analysis applied for the diagnosis of cervical carcinoma," *Anal. Methods* **7**, 2333–2338 (2015).
- D. Hou et al., "Terahertz spectroscopic investigation of human gastric normal and tumor tissues," *Phys. Med. Biol.* **59**, 5423–5440 (2014).
- D. Ng et al., "Classification of osteosarcoma T-ray responses using adaptive and rational wavelets for feature extraction," *Proc. SPIE* **6802**, 680211 (2007).
- N. Qi et al., "Terahertz time-domain spectroscopy combined with fuzzy rule-building expert system and fuzzy optimal associative memory applied to diagnosis of cervical carcinoma," *Med. Oncol.* **32**, 383 (2015).
- E. Berry et al., "Multispectral classification techniques for terahertz pulsed imaging: an example in histopathology," *Med. Eng. Phys.* **26**, 423–430 (2004).
- K. I. Zaitsev et al., "Terahertz spectroscopy of pigmentary skin nevi in vivo," *Opt. Spectrosc.* **119**, 404–410 (2015).
- F. Formanek, M.-A. Brun, and A. Yasuda, "Contrast improvement of terahertz images of thin histopathologic sections," *Biomed. Opt. Express* **2**, 58–64 (2010).
- S. Nakajima et al., "Terahertz imaging diagnostics of cancer tissues with chemometrics technique," *Appl. Phys. Lett.* **90**, 41102 (2007).
- W. R. Gilks, S. Richardson, and D. J. Spiegelhalter, Eds., *Markov Chain Monte Carlo in Practice*, Chapman and Hall, London (1995).
- T. Bowman et al., "Terahertz imaging of three-dimensional dehydrated breast cancer tumors," *J. Infrared Millimeter Terahertz Waves* **38**, 766–786 (2017).
- T. Bowman et al., "A phantom study of terahertz spectroscopy and imaging of micro- and nano-diamonds and nano-onions as contrast agents for breast cancer," *Biomed. Phys. Eng. Express* **3**, 55001 (2017).
- T. Bowman et al., "Terahertz imaging and segmentation of freshly excised xenograft mouse tumors," in *Proc. of XXXII Int. Union of Radio Science 2017 URSI General Assembly and Scientific Symp. (K7P Electromagnetic Biomedical Imaging)* (2017).
- T. Bowman et al., "Terahertz imaging of freshly excised breast cancer using mouse model," in *Proc. of 42nd Int. Conf. on Infrared, Millimeter and Terahertz Waves* (2017).
- K. Sharma and A. Goyal, "Classification based survey of image registration methods," in *Proc. of 2013 Fourth Int. Conf. Computing, Communications and Networking Technologies*, pp. 1–7 (2013).
- W. Dos Passos, *Numerical Methods, Algorithms and Tools in C#*, CRC Press, Hoboken, New Jersey (2009).
- J. Diebolt and C. P. Robert, "Estimation of finite mixture distributions through Bayesian sampling," *J. R. Stat. Soc. Ser. B* **56**, 363–375 (1994).
- C. Fernández and M. F. J. Steel, "Bayesian regression analysis with scale mixtures of normals," *Econ. Theory* **16**, 80–101 (2000).
- S. Frühwirth-Schnatter and S. Pyne, "Bayesian inference for finite mixtures of univariate and multivariate skew-normal and skew-t distributions," *Biostatistics* **11**, 317–336 (2010).
- M. Stephens, "Dealing with label switching in mixture models," *J. R. Stat. Soc. Ser. B Stat. Methodol.* **62**, 795–809 (2000).
- J. H. Albert and S. Chib, "Bayesian analysis of binary and polychotomous response data," *J. Am. Stat. Assoc.* **88**, 669–679 (1993).
- P. C. Ashworth et al., "Terahertz pulsed spectroscopy of freshly excised human breast cancer," *Opt. Express* **17**(15), 12444–12454 (2009).
- G. J. Wilmlink et al., "Development of a compact terahertz time-domain spectrometer for the measurement of the optical properties of biological tissues," *J. Biomed. Opt.* **16**(4), 047006 (2011).
- M. El-Shenawee et al., "Statistical signal processing for quantitative assessment of pulsed terahertz imaging of human breast tumors," in *Proc. of 42nd Int. Conf. on Infrared, Millimeter and Terahertz Waves*, Cancun, Mexico (2017).

Tyler Bowman received his BS and MS degrees in electrical engineering from the University of Arkansas, Fayetteville, Arkansas, in 2012 and 2014, respectively. He is currently a PhD candidate at the University of Arkansas with a research focus on terahertz imaging and characterization of cancer in the Terahertz Imaging and Spectroscopy Laboratory. He served as a NSF GK-12 fellow from 2012 to 2013 and was the recipient of NSF Graduate Research Fellowship in 2013.

Tanny Chavez is a second-year master's student in electrical engineering at the University of Arkansas. In 2015, she received her bachelor's degree in electronics and telecommunications engineering from Escuela Superior Politecnica del Litoral in Guayaquil, Ecuador. Her current research involves the statistical signal processing for breast cancer detection in terahertz imaging.

Kamrul Khan is a graduate research assistant in statistics and analytics (STAN) MS program at the University of Arkansas, Fayetteville. He received his BS and MS degrees in statistics from the University of Dhaka, Dhaka, Bangladesh, in 2011 and 2013, respectively. He is a recipient of 2017 Lawrence J. J. Jr. EF scholarship from the Department of Mathematical Sciences. His research interests include hierarchical Bayesian methods, mixture model, spatial analysis, time series, and Bayesian shrinkage.

Jingxian Wu is currently an associate professor in the Department of Electrical Engineering, University of Arkansas, Fayetteville. His research interests mainly focus on statistical signal processing, large scale data analytics, and wireless communications. He served as an associate editor of the *IEEE Transactions on Vehicular Technology* from 2007 to 2011 and He is an editor of the *IEEE Transactions on Wireless Communications* from 2011 to 2016, and is now serving as an associate editor of the IEEE access.

Avishek Chakraborty is an assistant professor in the Department of Mathematical Sciences, University of Arkansas. He graduated from Duke University in 2010 with a PhD in statistical science. His primary research interest lies in spatial statistics, uncertainty quantification and Bayesian computation for high-dimensional data.

Narasimhan Rajaram is currently an assistant professor of biomedical engineering at the University of Arkansas, Fayetteville. His research is focused on the development of functional and molecular optical imaging strategies that can survey the tumor microenvironment and determine early response to therapy as well as long-term clinical outcome in terms of tumor recurrence and metastasis.

Keith Bailey received a DVM in veterinary medicine and PhD in pathobiology from the University of Missouri–Columbia in 1992 and 1996, respectively. He has been responsible for more than 100 animal

toxicity studies in accordance with good laboratory practices including mouse and rat lifespan bioassays and evaluation of various rodent cancers. He serves as a director and veterinary comparative pathologist in the Oklahoma Animal Disease Diagnostic Laboratory, Center for Veterinary Health Sciences, Oklahoma State University.

Magda El-Shenawee received her PhD in electrical engineering from the University of Nebraska, Lincoln, USA, in 1991. She currently serves as a professor of electrical engineering at the University of Arkansas, where she joined as an assistant professor in 2001. Her research interests include terahertz imaging and spectroscopy, breast cancer detection, computational inverse scattering algorithms, MEMS antennas, nanoantennas for energy enhancement of photovoltaic solar cells, and biopotentials modeling of breast tumor cancerous cells.