

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/272170179>

Ultrasound RF Time Series for Classification of Breast Lesions

Article · February 2015

CITATION

1

READS

839

1 author:



Mehdi Moradi

IBM Research

137 PUBLICATIONS 3,170 CITATIONS

SEE PROFILE

Ultrasound RF Time Series for Classification of Breast Lesions

Nishant Uniyal, Hani Eskandari, Purang Abolmaesumi, Samira Sojoudi, Paula Gordon, Linda Warren, Robert N. Rohling, Septimiu E. Salcudean, and Mehdi Moradi*

Abstract—This work reports the use of ultrasound radio frequency (RF) time series analysis as a method for ultrasound-based classification of malignant breast lesions. The RF time series method is versatile and requires only a few seconds of raw ultrasound data with no need for additional instrumentation. Using the RF time series features, and a machine learning framework, we have generated malignancy maps, from the estimated cancer likelihood, for decision support in biopsy recommendation. These maps depict the likelihood of malignancy for regions of size 1 mm^2 within the suspicious lesions. We report an area under receiver operating characteristics curve of 0.86 (95% confidence interval [CI]: 0.84%–0.90%) using support vector machines and 0.81 (95% CI: 0.78–0.85) using Random Forests classification algorithms, on 22 subjects with leave-one-subject-out cross-validation. Changing the classification method yielded consistent results which indicates the robustness of this tissue typing method. The findings of this report suggest that ultrasound RF time series, along with the developed machine learning framework, can help in differentiating malignant from benign breast lesions, subsequently reducing the number of unnecessary biopsies after mammography screening.

Index Terms—Biopsy, computer aided diagnosis, support vector machines, time series analysis, ultrasonic imaging.

I. INTRODUCTION

ACCORDING to the statistics published by the American Cancer Society in 2013, it was estimated that breast cancer is one of the most common types of cancer in women accounting for 29% of all cancer cases [1]. The World Health Organization estimated that almost 1.38 million women worldwide are diagnosed with breast cancer annually. That accounts

for 23% of all cancer cases [2]. Early detection and better diagnosis methods play a significant role in reducing the number of fatalities due to breast cancer.

Based on mammography images, the American College of Radiology has developed a method called BI-RADS (Breast Imaging-Reporting and Data System) for deciding whether biopsy of an identified suspicious lesion is indicated. Biopsy is the gold standard for breast cancer diagnosis; however it is an expensive, discomforting, and invasive procedure. Almost 80% of the biopsies carried out based on BI-RADS score turn out to be benign [3]. Hence, there exists a need for reducing the number of unnecessary breast biopsies and augment the current diagnosis methods.

Mammography alone is limited in its ability to detect cancer. Studies report 73% sensitivity for all breast types and only 48% for dense breasts [4]. Almost 50% of women under the age of 50 have dense breasts [5] and could be potentially deemed undiagnosed after mammography screening [6], [7]. Ultrasound, along with mammography, has been proven to improve the sensitivity in palpable breast masses [4], [8], [9], but due to its low sensitivity towards nonpalpable and noncystic breast lesions, it is not used as a primary screening modality. Ultrasound is now widely used for supplementary screening in women with dense breasts, both by hand-held and automated modalities. In addition to palpable masses, ultrasound is also used to evaluate nonpalpable masses seen on mammography. However, its drawback lies in low contrast, high speckle B-mode images, and operator variability [10]. To address these limitations, several research groups are working towards ultrasound-based computer-aided diagnosis (CAD) for breast cancer diagnosis.

For example, Zheng *et al.* [11] used sonographic texture features along with self-organizing maps for classification of breast abnormalities. Chen *et al.* [12] also used neural networks and texture information extracted from ultrasound images to classify breast nodules. Giger *et al.* [13] calculated features related to lesion margin, shape, homogeneity (texture), and posterior acoustic attenuation patterns in ultrasound images and used linear discriminant analysis to classify breast lesions. Lefebvre *et al.* [14] used texture and morphometric parameters with linear discriminant analysis and leave-one-out cross-validation to classify breast lesions. Donohue *et al.* [15] described a breast lesion classification method using texture and generalized spectrum features along with linear and quadratic discriminant analysis. Shankar *et al.* investigated the Nakagami distribution [16] and non-Rayleigh statistics [17] of the backscattered envelope for classification of breast lesions. Chang *et al.* [18], [19] and Huang *et al.* [20] used the

Manuscript received June 07, 2014; revised August 28, 2014; accepted October 16, 2014. Date of current version January 30, 2015. This work was supported by the Natural Sciences and Engineering Research Council of Canada through the following programs: NSERC Discovery (PI: M. Moradi), NSERC Strategic (PI: Abolmaesumi), NSERC I2I (PIs: Salcudean-Rohling). *Asterisk indicates corresponding author.*

N. Uniyal, P. Abolmaesumi, S. Sojoudi, and S. E. Salcudean are with the Department of Electrical and Computer Engineering, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada.

H. Eskandari is with Analogic Ultrasound at Ultrasonix Medical ULC, Richmond, BC, V6V 2K9 Canada.

P. Gordon and L. Warren are with the Faculty of Medicine, Department of Radiology, University of British Columbia, Vancouver, BC, V6T 1Z4 Canada.

R. N. Rohling is with the Departments of Electrical and Computer Engineering and the Department of Mechanical Engineering at the University of British Columbia, Vancouver, BC, V6T 1Z4 Canada.

*M. Moradi is currently with the IBM Almaden Research Center, San Jose, CA 95120 USA (e-mail: moradi@ece.ubc.ca).

Color versions of one or more of the figures in this paper are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/TMI.2014.2365030

same texture analysis approach but employed support vector machines to classify breast lesions. Drukker *et al.* [21], [22] performed analysis of posterior acoustic shadowing for breast lesion detection. Joo *et al.* [23] employed an artificial neural network to detect breast lesions based on five morphological features representing the shape, edge characteristics, and darkness of a lesion. Chen *et al.* [24] described an approach based on fractal parameter computation from ultrasound images and the *k*-means classification algorithm. Alam *et al.* [25] reported the performance of logistic regression in classification of breast lesions based on quantitative acoustic parameters calculated using the spectrum analysis of ultrasound RF echo signals and morphometric features related to the lesion shape. In a more recent work, Gomez *et al.* [26] investigated the effectiveness of co-occurrence texture statistics calculated at different quantization levels, as a method for classification of breast images. Yang *et al.* [27] reported another method for ultrasound image diagnosis using gray-scale invariant features extracted via multi-resolution ranklet transform along with support vector machines. Tan *et al.* [28] developed a CAD system to detect cancer in 3-D breast ultrasound. A detailed review of the literature in the area of ultrasound-based CAD for breast cancer is reported in [29].

Ultrasound strain imaging [30], [31], Acoustic Radiation Force Impulse (ARFI) imaging [32], and supersonic shear-wave imaging [33]–[36] have also been used for breast lesion characterization. These methodologies require the use of excitation mechanisms or specialized equipment.

Recently, ultrasound RF time series analysis has been proposed as a novel tissue classification method by our research group. The RF time series method is based on a temporal analysis of beamformed RF signals. This type of analysis has proven successful in the classification of prostate cancer *ex vivo* and *in vivo*, and animal tissue [37]–[40]. The receiver operating characteristic (ROC) analysis showed that the area under the curve (AUC) for RF time series was recorded as 0.82 for *ex vivo* prostate [37], 0.76 for *in vivo* prostate [39], [41], and 0.92 in animal tissue classification [40]. This method does not require additional equipment and the region of interest selection is as simple as drawing a rectangular box inside the lesion. The task of size selection and placement of the rectangular box is much less subjective compared to manual contouring of the lesions. As such, it can be a valuable and practical addition to the previously reported ultrasound-based methods for breast lesion classification.

This is the first report of using the RF time series analysis in the context of lesion classification in breast images. Preliminary results have been previously presented in [42]. For this study, we use quantitative analysis of spectral and fractal parameters extracted from RF time series. We report the use of both support vector machines (SVM) and Random Forests classification methods with RF time series features and demonstrate accurate malignancy maps that can be used for decision support in biopsy recommendation.

II. METHODS

We describe the use of a machine learning framework to assign labels to each sample and also estimate the cancer likeli-

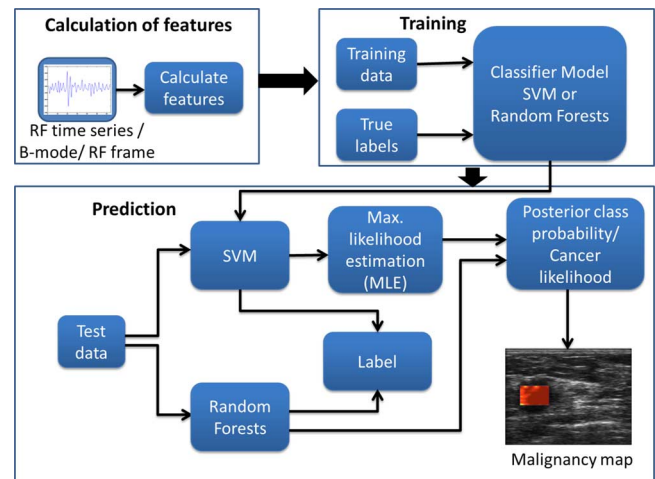


Fig. 1. Overall classification framework. This figure shows a graphical representation of the proposed approach.

hood (for SVM) and probability (in case of Random Forests). The estimated cancer likelihood is then used to generate a malignancy map. Fig. 1 shows a graphical representation of this approach.

A. Data Collection

An RF time series is formed by the sequence of RF echoes received from one location in the tissue over time. To acquire the RF time series data, the ultrasound probe and the tissue remain fixed and frames of RF signals are acquired. In this method of analysis, the tissue typing parameters are extracted from the temporal changes of the signal, as opposed to the classical method of spectral analysis on spatial segments of the RF signals [43].

The data collection for this work was performed during a study approved by the Clinical Research Ethics Board at the University of British Columbia. This study was designed to evaluate ultrasound-based tissue typing and vibro-elastography for breast lesion classification. The vibro-elastography technology was developed by Salcudean *et al.*, and the preliminary results were reported in [44]. Subjects referred to ultrasound-guided biopsy, based on mammography screening, were consented for collection of vibro-elastography and RF data during biopsy. The study was conducted between September 2012 and January 2013. Data was obtained on a SonixTouch ultrasound machine (Ultrasonix Medical Corp., Richmond, BC, Canada). The research platform provided by the manufacturer enabled acquisition of beamformed RF signals in real time.

For every subject, the sonographer first performed a preliminary ultrasound scan to find the suspected lesion. Once the lesion was located, the sonographer would hold her hand steady for 5 s while a computer program stored the RF data into the memory and consequently saved it in a file. Imaging was performed with an L14–5/38 ultrasound transducer at a center frequency of 10 MHz and a depth of 4 cm. Each RF line was sampled at 40 MHz and a total of 128 scan-lines were acquired for each RF frame. With these image settings, RF data at a frame-rate of 98 frames per second, for five seconds, was obtained. The data collection of each subject was followed by a

TABLE I
SUMMARY OF THE DATA

Subject	BI-RADS	Aprox. Lesion Size (mm ²)	# of 1 mm ² ROIs	Tissue Type	Biopsy Result
Subject 1	5	77	48	IDC	Malignant
Subject 2	5	139	80	IDC	Malignant
Subject 3	5	132	24	IDC	Malignant
Subject 4	4B	74	56	IDC	Malignant
Subject 5	4C	27	12	IDC	Malignant
Subject 6	4C	38	12	IDC	Malignant
Subject 7	4B	13	9	IDC	Malignant
Subject 8	4A	53	28	FA	Benign
Subject 9	3	135	44	FA	Benign
Subject 10	4	84	48	FA	Benign
Subject 11	4B	41	18	FA	Benign
Subject 12	4A	36	24	N/A	Benign
Subject 13	4B	72	28	FA	Benign
Subject 14	4A	65	28	FA	Benign
Subject 15	4B	53	10	CI	Benign
Subject 16	4A	46	32	FA	Benign
Subject 17	4A	30	16	FA	Benign
Subject 18	4A	136	12	FA	Benign
Subject 19	4A	110	15	Cyst	Benign
Subject 20	4A	322	160	FA	Benign
Subject 21	4A	210	120	FA	Benign
Subject 22	4A	60	39	NS	Benign

routine ultrasound exam and a core needle biopsy of the lesion under ultrasound guidance by the physician. In total, 35 subjects were imaged. However, in five subjects the radiologist decided that biopsy was unnecessary. Also, in eight subjects where biopsies were performed, the exact location of the biopsy sample could not be identified within the ultrasound image. As a result, we used the data corresponding to 22 subjects for whom the biopsy result was available and could be mapped to a specific lesion. Biopsy results for these subjects showed seven malignant lesions, all of the invasive ductal carcinoma type and 15 benign subjects, mostly of fibroadenoma type. Table I shows the BI-RADS score, approximate lesion size, number of 1 mm² ROIs, tissue type, and biopsy result for all the 22 subjects (IDC stands for invasive ductal carcinoma, FA stands for fibroadenoma, CI stands for chronic inflammation, and NS stands for nodular sclerosis). The tissue type is not available for subject 12 because the cytology result was unknown due to low cellularity. Note that the lesion size is an approximation.

To use the RF time series method, patient motion has to be minimized or compensated. In the current study, we used only the first 128 frames of the data, which accounts for 1.3 s of acquisition, to form the RF time series. Our motion estimation using the method described in [45] showed that when the time series is limited to 128 frames, the displacement caused by patient motion is on average smaller than the resolution of the RF samples. This means that we could, on average, assume that series of RF samples obtained at a certain image coordinate were, in fact, from the same physical location in the tissue.

B. Features

1) *Features Per ROI*: The biopsied lesions were divided into 1 mm² regions of interest (ROIs). The tissue typing features were extracted from these ROIs. In the RF domain, this ROI size was equivalent to 3 × 52 samples each forming a time series. In the interpolated B-mode image the ROI size of 1 mm² was equivalent to 26 × 26 pixels. The RF frame size was 128 × 2080 and the B-mode image resolution was 988 × 1040. Note that

TABLE II
ULTRASOUND RF TIME SERIES, B-MODE TEXTURE, AND ATTENUATION FEATURES CALCULATED FOR CLASSIFICATION OF MALIGNANT AND BENIGN BREAST LESIONS

#	Feature	Origin
1	RF time series first quadrant	RF time series
2	RF time series second quadrant	RF time series
3	RF time series third quadrant	RF time series
4	RF time series fourth quadrant	RF time series
5	Intercept of regression line fitted to norm. spect.	RF time series
6	Slope of regression line fitted to norm. spect.	RF time series
7	Higuchi fractal dimension	RF time series
8	Intercept of line fitted to calibrated spectrum	Single-frame RF
9	Slope of line fitted to calibrated spectrum	Single-frame RF
10	Texture: Mean	B-mode
11	Texture: Standard Deviation	B-mode
12	Texture: Skewness	B-mode
13	Texture: Kurtosis	B-mode
14-17	Correlation: $F_d(\theta) = \sum_{i,j} \frac{(i - \mu_i)(j - \mu_j)p(i,j)}{\sigma_i \sigma_j}$	B-mode
18-21	Energy: $F_d(\theta) = \sum_{i,j} p(i,j)^2$	B-mode
22-25	Contrast: $F_d(\theta) = \sum_{i,j} i - j ^2 p(i,j)$	B-mode
26-29	Homogeneity: $F_d(\theta) = \sum_{i,j} \frac{p(i,j)}{1 + i - j }$	B-mode

† Features 14-29 are calculated for each $d \in (1, 2, \dots, 10)$ resulting in 160 (16 × 10) feature values.

θ is the orientation angle for GLCM calculation ($\theta = 0^\circ, 45^\circ, 90^\circ, 135^\circ$).

d is the pixel distance for GLCM calculation.

$p(i, j)$ is the $(i, j)^{th}$ entry in the co-occurrence probability matrix.

$$\mu_i = \sum_{i,j} i \cdot p(i, j) \text{ and } \mu_j = \sum_{i,j} j \cdot p(i, j)$$

$$\sigma_i^2 = \sum_{i,j} (i - \mu_i)^2 \cdot p(i, j) \text{ and } \sigma_j^2 = \sum_{i,j} (j - \mu_j)^2 \cdot p(i, j)$$

from the 22 lesions in the dataset, a total of 863 ROIs were extracted. Among these, 241 were malignant. From each subject, we extracted multiple ROI and the number of ROIs depends on the size of the lesion. The lesion size and the number of 1 mm² ROIs can be found in Table I.

In addition to the RF time series features, we also calculated and used the values of texture features extracted from the first B-mode image and spectral RF features extracted from the first RF frame. A similar approach to B-mode texture profiling for tissue typing is reported in prior work [12]–[15], [46]–[48], [26], [27]. A complete list of the studied parameters can be found in Table II-B2.

2) *Features Per Lesion*: We also calculated the RF time series, single-frame RF spectral, and B-mode texture features for each lesion. This was done to make a fair comparison of the texture features with the RF time series. In this method, the area of interest was the entire lesion rather than several 1 mm² ROIs inside the lesion. For each of the 22 lesions, features were calculated for the biggest possible rectangular region inside the lesion boundary.

3) *RF Time Series Features*: A variety of methods have been used for tissue classification using RF time series. We use the method originally described in [49] which proposes summarizing the power spectrum of the RF time series in six parameters. The frequency spectrum was estimated by calculating the FFT-based periodogram of the Hamming windowed time series. This estimated spectrum was divided into four frequency bands and each averaged to deliver a feature. In other words, the first four features (Features 1–4) were the average of the frequency

spectrum in $[0, \pi/4)$, $[\pi/4, \pi/2)$, $[\pi/2, 3\pi/4)$, $[3\pi/4, \pi)$ frequency bands in the discrete frequency domain. Note that the sampling rate here is equivalent to the frame rate of the ultrasound machine. Two other spectral features were the intercept (Feature 5) and the slope (Feature 6) of a regression line fitted to the magnitude of the spectrum versus normalized frequency.

Feature 7 was the average fractal dimension (FD) of RF time series in a region of interest. In this context, fractal dimension is a measure of the nonlinear complexity of the signal. For calculation of the FD, we used the algorithm proposed by Higuchi [50] which decomposes the signal into different scales and evaluates the signal complexity. We used the Higuchi algorithm with 16 levels of decomposition for the time series of length 128. The use of this feature in ultrasound-based tissue typing was originally proposed in [51].

4) *Single-Frame RF Spectral Features*: The tissue attenuation coefficient and RF spectral parameters that are related to attenuation have been used for tissue typing. Similar to texture features, these features were also calculated using the first RF frame. One common method for evaluating the frequency-dependent attenuation of ultrasound in tissue is the spectral analysis of segments of echoed RF signals, after calibration to remove the effects of the imaging system. Researchers at the Lizzi Center for Biomedical Engineering at Riverside Research have shown that the parameters of a linear model fitted to the calibrated power spectrum of an RF segment contain valuable tissue classification information [52], [53]. To estimate these parameters, the averaged tissue power spectrum within an ROI had to be calibrated to account for the ultrasound machine transfer function. The calibration spectrum was acquired from the surface of a flat glass plate in a water bath at the transducer focal zone, with minimum amplifier gain and flat time gain control (TGC). Following [52], we subtracted the logarithm of the averaged spectrum of the glass from the logarithm of the spectrum of the tissue. For our data, we found that the useful part of the spectrum was at the range of 1–10 MHz and this is apparent in Fig. 2. Outside of this range, the power drops to less than half of the peak value. Therefore, we fitted the regression line to the data in this range. As can be seen in, Fig. 3 the calibrated spectrum shows a quasi-linear behavior. Features 8 and 9 are the intercept and slope of the linear model fitted to the normalized frequency spectrum.

5) *B-Mode Texture Features*: Features 10–173 are 164 texture parameters calculated using the B-mode image reconstructed from the first RF frame. The RF frame was converted into a B-mode image by taking the Hilbert transform and then resized to 988×1040 to match the physical dimension (3.8×4 cm) of the image. The first four parameters were simply the mean, standard deviation, skewness and kurtosis of the pixel intensities in an ROI and the other 160 parameters were the correlation, energy, contrast, and homogeneity calculated from the gray-level co-occurrence matrix (GLCM). The GLCM was created by calculating how often a pixel with a grayscale intensity i occurs adjacent to a pixel with the value j . For the purpose of this study, we calculated the co-occurrence matrices at a distance of ten pixel ($d = 1, 2, \dots, 10$), for directions of $\theta = 0^\circ$, $\theta = 45^\circ$, $\theta = 90^\circ$, and $\theta = 135^\circ$. For every one of the 40 combinations of d and θ values a GLCM was calculated.

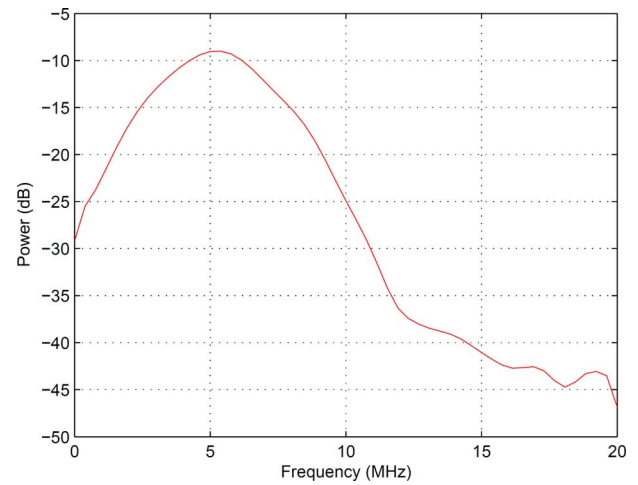


Fig. 2. Average power spectrum of the reflected RF signal for breast tissue.

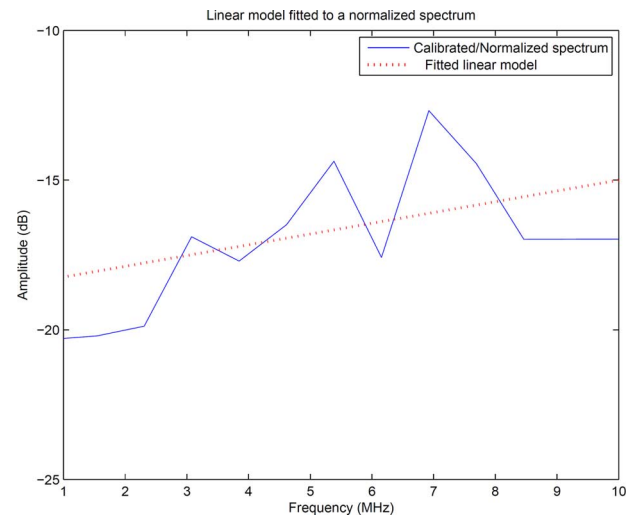


Fig. 3. Attenuation estimation by calculating the slope of the calibrated spectrum.

Four texture statistics are calculated from each GLCM and the definition of these features is provided in Table II-B2.

We also calculated the single-frame RF spectral features and B-mode texture features for ROI sizes of 4 mm^2 and 9 mm^2 , for comparison with ROIs of 1 mm^2 . These features were also calculated for the entire lesion.

C. Classification and Cancer Likelihood/Probability Estimation

1) *Support Vector Machine (SVM)*: The soft margin SVM algorithm implemented in the scikit-learn Python package and reported in [54] was used for tissue classification. The radial basis function (RBF) kernel was employed. The values of the RBF exponent (γ), which determines the range of support, and the parameter C , which controls the trade-off between the slack penalty and the margin in the soft margin SVM, were chosen based on a grid search through leave-one-subject-out cross-validation.

Cancer likelihood for each ROI, which is a number between 0 and 1, was estimated using Platt's algorithm [55] as follows. Assume that the SVM hyperplane obtained after training is

$\mathbf{w}^\top \phi(\mathbf{x}) + b$ where $\mathbf{x}$ is the feature vector and ϕ is the function that maps the feature vector to a higher dimensional space. Cancer likelihood (L_c) is computed by mapping the distance of each test sample to the decision boundary using a sigmoid function of form

$$L_c = (c | (\mathbf{w}^\top \phi(\mathbf{x}_i) + b)) = \frac{1}{1 + \exp(\alpha(\mathbf{w}^\top \phi(\mathbf{x}_i) + b) + \beta)} \quad (1)$$

Where c stands for cancer (class). Maximum likelihood estimation from the observations for which the true labels are known (training data) was used to calculate the values of the parameters α and β .

2) *Random Forests*: Random Forests implementation within the scikit-learn Python package was used [54]. Forest class a posteriori probabilities or malignancy probabilities $p(c|\mathbf{v})$, is calculated as the average of all tree a posteriori probabilities $p_t(c|\mathbf{v})$ and can be described as

$$p(c|\mathbf{v}) = \frac{1}{T} \sum_t p_t(c|\mathbf{v}) \quad (2)$$

where $\mathbf{v}$ is the input data, c stands for cancer (class), and T is the number of decision trees in the forest. The a posteriori class probability $p_t(c|\mathbf{v})$ is calculated at each leaf of the decision tree as the fraction of the number of samples with the majority label over the total number of samples. The label assigned to the test data is based on the forest class a posteriori probability [56]. The algorithm parameters were optimized by performing a grid search and the decision trees were constructed shallow to avoid over-fitting. The classification algorithm implemented in scikit-learn package closely follows the method described in [57] and [58]. More details about both, SVM and Random Forests formulations can be found in the Appendix.

3) *Feature and Model Selection*: To use the most informative features and reduce the dimensionality of the data, only a subset of the features were chosen for classification. It is important to note that the Random Forests algorithm provides a built-in method for ranking features in terms of their ability to separate samples from the two classes. This is not the case for SVM. Therefore, our approach to feature selection was different with the two classifier models.

In case of the SVM classifier, we used recursive feature elimination (RFE). The RFE algorithm, originally reported in [59], recursively eliminates the features with the smallest contribution to an estimated linear maximum margin model and is shown to outperform correlation-based methods in gene selection [59]. In this method, features are eliminated recursively based on their corresponding weight in a linear SVM classifier of the $y = \mathbf{w}^\top \mathbf{x} + b$. Initially, the model is trained on all the features and weights are calculated for each one of them. Then the feature with the smallest absolute weight value is eliminated. This process is then repeated recursively until the desired number of features is reached [59].

In Random Forests algorithm, feature importance is estimated as follows: each decision tree is created using a bootstrap sample of the original data. Hence, the left out data is labelled as out-of-bag (OOB) samples and can be used as “test data” to estimate feature importance. In this step, first all OOB samples

are propagated through the trained model and the accuracy is recorded. Then the values of one feature for OOB samples are randomly permuted and the OOB samples with this permuted feature value are propagated down the trees. The expectation is that the permutation results in decreased accuracy. The permutation of the “important” features will cause a larger reduction in classification accuracy. The difference between average ensemble accuracy with and without permutation provides a raw measure of feature importance. The features were ranked based on this importance measure. The choice of the number of features to select for classification was considered as one of the model parameters. All model parameters, which included the depth of trees, the number of trees, and the number of features were optimized through a grid search. Leave-one-subject-out cross-validation was used for this grid search.

D. Classification Experiments and Evaluation of Classification

The classification label was compared with the biopsy result, which was used as the ground truth. We assess the performance of the classifiers by calculating the AUC. The ROC curves were generated using the method described in [60] and the AUC was estimated using the trapezoidal rule. We also calculated the standard errors and p -values for the AUCs using the Hanley and McNeil method reported in [61].

In order to evaluate the performance of other groups of features as opposed to RF time series, we trained an SVM classifier using parameter selection that included only texture and attenuation parameters. All other aspects of the method were similar to the time series experiment.

We also performed a second set of classification experiments where the classified object was the whole lesion, as opposed to ROIs. This was primarily due to the fact that texture analysis is most accurate for large areas, such as the whole lesion. This experiment, therefore, affords a fair comparison with B-mode texture features. The disadvantage of this analysis is that the sample size is reduced to 22, effectively making the estimation of likelihoods based on [55] impractical. As a result, in this analysis, we have used the binary labels themselves as the decision threshold for plotting the ROC, resulting in a three-point ROC curve as opposed to a smooth ROC curve.

III. RESULTS

We report the results separately for the classification experiments with the ROI as the sample, and whole lesion as the sample. Note that cancer likelihood estimation was only possible when the ROIs were classified.

A. Classification of 1 mm² ROIs

1) *Feature Selection*: The feature selection algorithm showed that using three features resulted in the lowest cross-validation error in both SVM and Random Forests. Features 2, 3, and 7 were the best features selected for classification using SVM and features 2, 5, and 7 yielded the highest importance score for classification using Random Forests. It should be noted, that the fractal dimension feature extracted from the RF time series, feature 7, was the best overall. Note that for both classifiers the final set of features only included RF time series features.

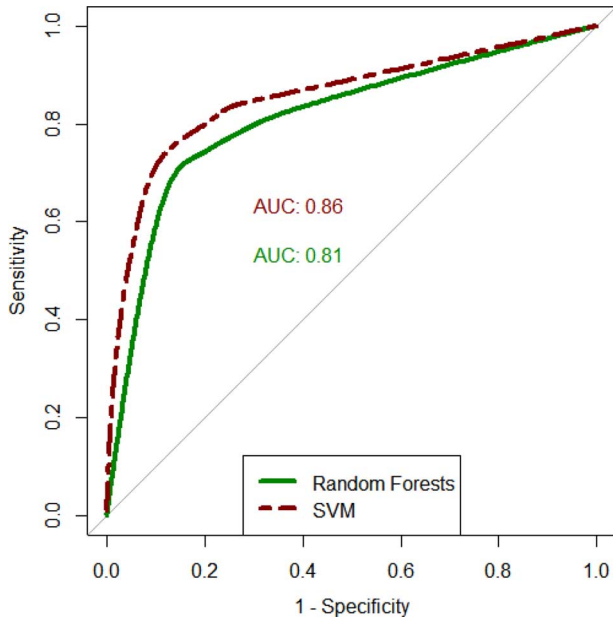


Fig. 4. ROC curve for classification of ROIs using SVM and Random Forests.

TABLE III
AREA UNDER THE ROC CURVE

Algorithm	AUC (95% CI)	Standard Error
SVM	0.86 (0.84–0.90)	0.016
Random Forests	0.81 (0.78–0.85)	0.018

2) *Classification Results*: As shown in Fig. 4, the AUC using SVM was 0.86 and using Random Forests was 0.81. Table III shows the confidence intervals for the AUC values.

Using the statistical tests proposed in [61], the standard errors were calculated as 0.016 and 0.018 for AUC using SVM and Random Forests, respectively. Using the same method, and performing the *z* test, the *p*-value was calculated to be 0.008, hence, the difference in performance of the two classifiers was statistically significant ($p < 0.05$) [61].

Cancer likelihood translated to colormaps and applied to the lesion area on the B-mode image can be found in Fig. 6. In 18 of the 22 subjects, the majority of the 1 mm² ROIs were correctly classified. This is also visually evident in Fig. 6. In subjects 3, 7, 10, and 16, the majority of ROIs were incorrectly classified. The corresponding likelihood colormaps are presented in Fig. 7. Further investigation of these subjects showed calcification inside the lesion for Subject 3 and poor lesion boundaries for Subjects 7 and 16. It should also be noted that Subject 7 was the smallest lesion in our dataset.

The effectiveness of the RF time series method is evident from the fact that only RF time series features were selected by our systematic ranking method. Classification using the best three single-frame RF spectral and B-mode texture features was also performed for comparison with time series features. The three selected features in this experiment were the intercept and slope from the attenuation group (Features 8 and 9) and mean intensity from the texture group (Feature 10). This combination resulted in an AUC of 0.68. However, these parameters provide

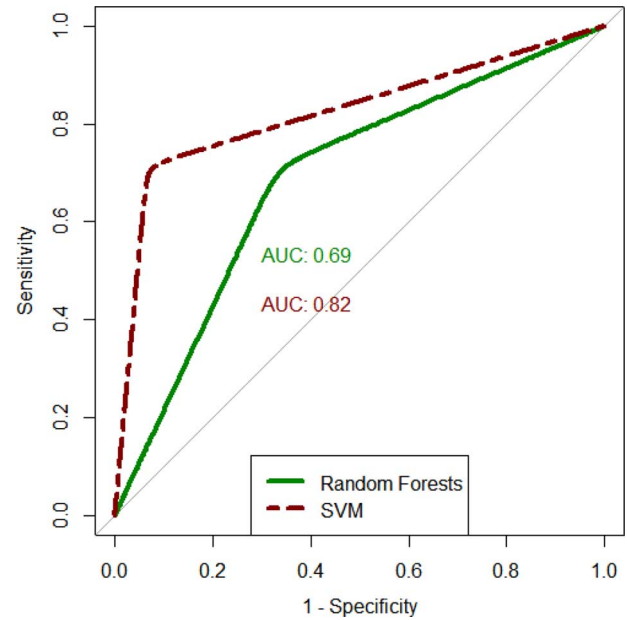


Fig. 5. ROC curve for classification of lesions using SVM and Random Forests.

a benefit. Specifically, the classifier trained on texture and attenuation parameters was successful in classifying both of the benign subjects—subjects 10 and 16—that were misclassified by the time series method. This was, however, at the cost of reduced sensitivity.

The AUC for ROI sizes of 4 and 9 mm² was calculated to be 0.69 and 0.70, respectively. We failed to find a statistically significant difference in AUCs based on ROI sizes ($p = 0.91$).

B. Classification of Lesions

1) *Feature Selection*: Both the feature selection methods, RFE and Random Forests consistently selected a time series feature, a single-frame RF spectral feature, and a B-mode texture feature as the best features for classification. These features were: Feature 7 (fractal dimension), Feature 9 (slope of the line fitted to the calibrated spectrum), and Feature 164 (B-mode texture feature calculated from the co-occurrence matrix). It is notable that when the texture features are extracted from the whole lesion, one of the B-mode texture features appears in the top three features.

2) *ROC Analysis*: Using only the best three features out of the 173 features, the AUC values were calculated to be 0.82 (95% CI: 0.63–1.00) using SVM and 0.69 (95% CI: 0.47–0.91) using Random Forests. It should be noted that the ROC curves were created using the classification labels, as opposed to the cancer likelihood estimation. The ROC curves can be found in Fig. 5. Note that these ROC curves are not comparable to the ROCs in Fig. 4 in which the cancer likelihood estimates are used as the decision variable.

The AUC value, using only the single-frame RF spectral features and the B-mode texture features was calculated to be 0.68. This was obtained using a SVM classifier.

IV. DISCUSSION AND CONCLUSION

There exists a need to reduce the number of false positives based on a BI-RADS score, hence, augmentation of the cur-

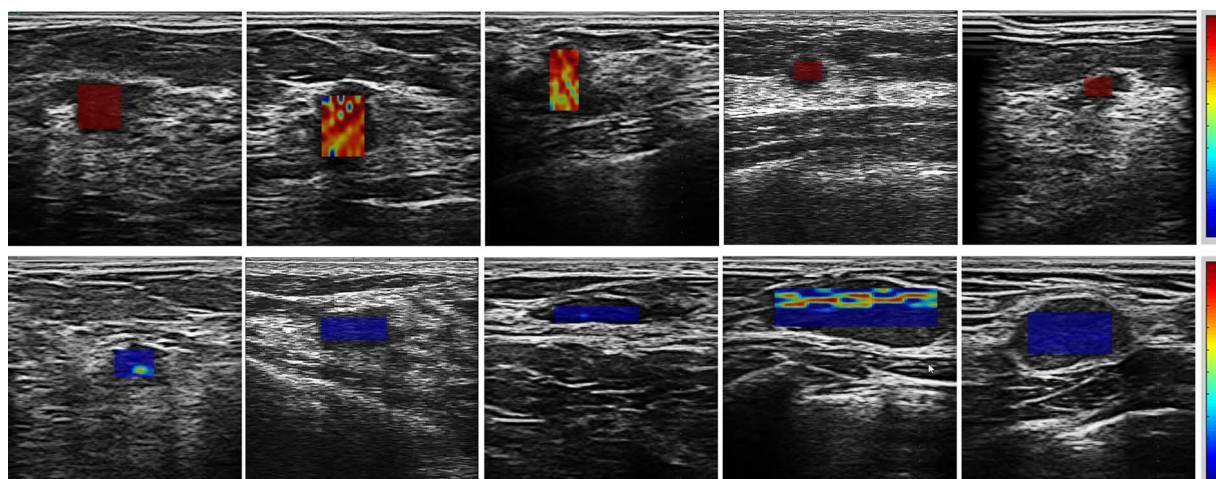


Fig. 6. Malignancy maps created by plotting the the cancer likelihood (L_c) of the ROIs overlaid on the B-mode image. Images on the top are from malignant subjects, and the ones on the bottom are benign subjects. For each of the above cases we were able to correctly estimate the cancer likelihood for the majority of the ROIs inside the lesion.

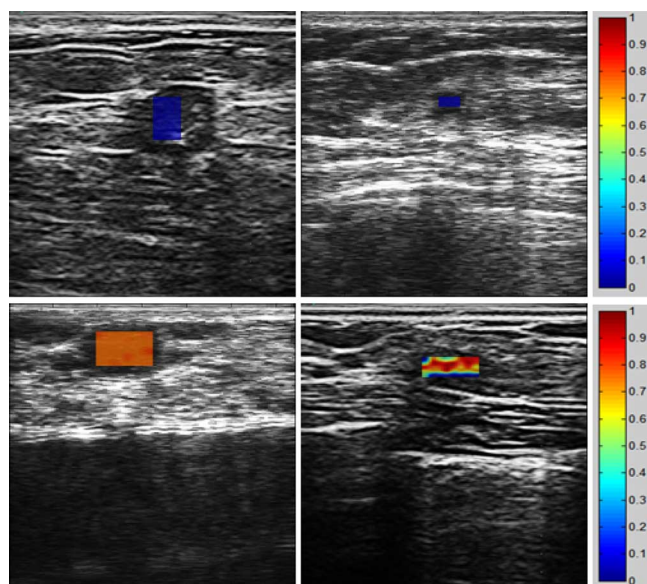


Fig. 7. Malignancy maps for the subjects for which the estimated cancer likelihood of the majority of the ROIs inside the lesion was incorrect. On the top are two malignant subjects: Subjects 3 (left) and 7 (right), and the ones on the bottom row are benign subjects: Subjects 10 (left) and 16 (right).

rent breast cancer diagnosis methods is a pressing need. In this work we presented an ultrasound-based method for classification of breast lesions. This method uses RF time series parameters within a machine learning framework. A set of 173 parameters extracted from RF time series analysis, single-frame RF spectral analysis, and B-mode texture were ranked for their importance with two different feature ranking approaches. In both feature-ranking algorithms, all the three best performing features were from the RF time series group. Using these three features, we obtained an AUC of 0.86 and 0.81 on SVM and Random Forests classifiers, respectively. The best performing subset of features that were not from RF time series resulted in an AUC of 0.68. We also performed classification using features

calculated from the entire lesion and found the AUC to be 0.82 using SVM and 0.69 using Random Forests.

The performance of B-mode texture features in breast cancer classification was poor in our study. However, it should be noted that the B-mode images used in this work were reconstructed from the RF signals offline. The B-mode images from the commercial ultrasound machines are filtered and optimized in terms of dynamic range and have a higher quality compared to the B-mode images used here. Therefore, the reported performance of the B-mode texture features could potentially be improved by using the B-mode images produced by the scanner. In addition, small (1 mm^2) ROI size could also be the cause of poor classification performance of the single-frame RF spectral features and the B-mode texture features. We increased the ROI size to 4 and 9 mm^2 and observed a slight improvement in the AUC. We did not go above 9 mm^2 because that would have resulted in loss of data from two subjects due to small lesion size. However, we performed a second kind of analysis where the entire lesion areas were considered as samples of unequal physical size for classification. Even in this experiment, RF time series and B-mode and single frame RF features resulted a higher AUC of 0.82 compared to B-mode and RF single frame features alone which resulted in $\text{AUC} = 0.68$.

Our dataset consists of mainly benign lesions and hence the majority of the training data belongs to the benign class. The decrease in classification performance using Random Forests could be due to this class imbalance problem [62], [63]. Another possible cause of the decreased AUC values calculated using the Random Forests algorithm is the small number of selected features. The power of Random Forests is in random sampling of features, which is most useful in high-dimensional feature vectors [58].

RF time series method is susceptible to patient motion. The time series used in this work were obtained over less than 2 s. We found that by asking the subjects to hold their breath, the amount of motion could be minimized. An ultrasound transducer stabilization mechanism could also reduce the operator-introduced motion and increase the length of motion-free time series, which

is linked to improved tissue classification in previous work [40]. Our measurement based on [45] showed that the amount of displacement in the RF data had the highest value in the misclassified malignant subjects. Also, our dataset includes mostly benign subjects which introduces a class imbalance problem. This affects the classification performance, specially the sensitivity. We need a larger dataset with improved representation of calcification and malignancy. As the size of the dataset increases, we anticipate improved performance.

RF time series method is a data-driven approach that has yielded consistently strong results in animal tissue experiments [40], prostate cancer *ex vivo* [37], prostate cancer *in vivo* [41], and in monitoring of tissue ablation [64]. While we have not discussed the potential reasons for improved performance obtained from the use of a sequence of frames, several hypotheses describing the physical phenomenon exist and have been explored elsewhere. It has been shown experimentally and in simulations, that the temperature increase in controlled irradiation of tissue with RF time series can partly explain the phenomenon [65]. Also, increased energy delivered to the tissue (through increased frame rate and transmit power) results in improved tissue classification accuracy in phantom and animal studies [40]. Given these observations, micro-vibrations of the tissue microstructure caused by acoustic radiation force is another likely phenomenon. The acoustic radiation force is related to both the acoustic energy and to the attenuation and scattering properties of the tissue, which are different for different pathological tissue types. Simulation studies are ongoing to further validate these hypotheses.

Accurate malignancy maps could eventually be used as a diagnosis method which would significantly reduce the negative biopsy outcome. Our proposed method can be part of the overall solution for multiparametric ultrasound analysis of breast cancer. We argue that RF time series can be a practical component of that approach without the need for additional equipment. This method is also independent of the contour shape and the potentially subjective process of segmentation.

APPENDIX

Support Vector Machines: SVM is a widely used maximum margin classifier. The soft margin SVM classifier is a hyperplane of the form $\mathbf{w}^\top \phi(\mathbf{x}_i) + b$, where ϕ is the function that maps the feature vector, $\mathbf{x}_i$, to a higher dimensional space and $\mathbf{w}$ and b are determined to minimize [66]

$$\frac{1}{2} \times \mathbf{w}^\top \mathbf{w} + C \sum_{i=1}^N \xi_i \quad (3)$$

subject to

$$y_i(\mathbf{w}^\top \phi(\mathbf{x}_i) + b) \geq 1 - \xi_i \quad (4)$$

where $C > 0$ is the regularization or penalty parameter that minimizes the error by controlling the trade-off between the slack variable penalty and the margin, and $\xi_i \geq 0$ are the slack variables that provide flexibility when fitting the data by permitting incorrect classification of noisy and difficult data points [66]. The function $\phi(\cdot)$ maps the data to a higher dimensional space. This new space is defined by its kernel function. It can

be shown that in the SVM formulation, one does not need an explicit expression for the kernel function and the SVM optimization and decision hyperplane are defined fully given the dot product format of the kernel $K(\mathbf{x}_i, \mathbf{x}_j) = \phi(\mathbf{x}_i)^\top \cdot \phi(\mathbf{x}_j)$.

The classification problem can also be written in the form of its *dual representation* in terms of the kernel function. In this representation, the class label is determined based on the sign of $y(\mathbf{x})$ in the following equation:

$$y(\mathbf{x}) = \sum_{i=1}^N a_i t_i K(\mathbf{x}, \mathbf{x}_i) + b \quad (5)$$

where $y(\mathbf{x})$ is the model prediction of input $\mathbf{x}$, t_i is the true label for the support vector $\mathbf{x}_i$, a_i is the Lagrange multiplier used to convert the maximum margin SVM optimization problem from 3 to its dual representation. The solution to the SVM training is a quadratic programming problem that solves for a_i 's. See [66] for details and [67] for implementation.

Random Forests: Random Forests is a robust classification algorithm employed for medical image processing applications [57], [58]. Random Forests algorithm is a collection of decision trees that exploits the bias variance trade-off. The underlying concept behind the algorithm is that by taking the prediction of each decision tree and averaging it across the ensemble/forest the high variance in the prediction of decision trees can be reduced. The Random Forests algorithm can be described as follows.

- 1) For 1 to T trees:
 - a) Draw a bootstrap (bagging) sample from the training data.
 - b) Recursively repeat the steps below at each terminal leaf, until the desired depth is reached, to construct a random tree from the bootstrapped data.
 - i) Randomly select a subset of n features from f features.
 - ii) Select the split point that results in maximum information gain along each n feature.
 - iii) Split the leaf into two child leaves.
- 2) Output the forest of T trees.

The objective function at each leaf of the tree is $I = I(S_j, \theta)$, where S_j is the j th node, θ is the leaf parameter or the split point, and I is the information gain. The information gain is calculated as

$$I(S, \theta) = H(S) - \sum_{i \in \{L, R\}} \frac{|S^i|}{|S|} H(S^i) \quad (6)$$

where, $H(S)$ is Shannon's entropy calculated as

$$H(S) = - \sum_{c \in C} p(c) \log(p(c)). \quad (7)$$

In the above equation C is the output/classes. From the n randomly selected features, the feature that results in the highest information gain is used at the root node and the best split point is also calculated with the objective of maximizing the information gain. During testing, the sample is pushed down each tree simultaneously until it reaches the corresponding leaf.

REFERENCES

- [1] R. Siegel, "Cancer statistics, 2013," *CA: Cancer J. Clinicians*, vol. 63, no. 1, pp. 11–30, 2013.
- [2] P. Boyle and B. Levin, International Agency for Research on Cancer World cancer report 2008 IARC, 2008.
- [3] S. Poplack, P. Carney, and J. Weiss, "Screening mammography: Costs and use of screening-related Services1," *Radiology*, vol. 234, no. 1, pp. 79–85, 2005.
- [4] T. Kolb, J. Lichy, and J. Newhouse, "Comparison of the performance of screening mammography, physical examination, and breast us and evaluation of factors that influence them: an analysis of 27,825," *Radiology*, vol. 225, no. 1, pp. 165–175, 2002.
- [5] P. Stomper and D. D'souza, "Analysis of parenchymal density on mammograms in 1353 women 25–79 years old," *AJR. Am. J. Roentgenol.*, vol. 167, no. 5, pp. 1261–1265, 1996.
- [6] W. A. Berg, L. Gutierrez, M. S. Nasseaiver, W. B. Carter, M. Bhargavan, R. S. Lewis, and O. B. Ioffe, "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer," *Radiology*, vol. 233, no. 3, pp. 830–849, 2004.
- [7] N. Boyd, H. Guo, and L. Martin, "Mammographic density and the risk and detection of breast cancer," *N. Eng. J. Med.*, vol. 356, no. 3, pp. 227–236, 2007.
- [8] W. Berg and J. Blume, "Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer," *JAMA*, vol. 299, no. 18, pp. 2151–2163, 2008.
- [9] K. M. Kelly, J. Dean, W. S. Comulada, and S.-J. Lee, "Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts," *Eur. Radiol.*, vol. 20, no. 3, pp. 734–742, Mar. 2010.
- [10] G. Rahbar, A. Sie, and G. Hansen, "Benign versus malignant solid breast masses: US Differentiation1," *Radiology*, vol. 213, no. 3, pp. 889–894, 1999.
- [11] Y. Zheng, J. F. Greenleaf, and J. J. Gisvold, "Reduction of breast biopsies with a modified self-organizing map," *IEEE Trans. Neural Netw.*, vol. 8, no. 6, pp. 1386–1396, Nov. 1997.
- [12] D. Chen, R. Chang, and Y. Huang, "Computer-aided diagnosis applied to US of solid breast nodules by using neural Networks1," *Radiology*, vol. 213, no. 2, pp. 407–412, 1999.
- [13] M. L. Giger, H. Al-Hallaq, Z. Huo, C. Moran, D. E. Wolverton, C. W. Chan, and W. Zhong, "Computerized analysis of lesions in US images of the breast," *Acad. Radiol.*, vol. 6, no. 11, pp. 665–674, Nov. 1999.
- [14] F. Lefebvre, M. Meunier, and F. Thibault, "Computerized ultrasound B-scan characterization of breast nodules," *Ultrasound Med. Biol.*, vol. 26, no. 9, pp. 1421–1428, 2000.
- [15] K. Donohue, L. Huang, and T. Burks, "Tissue classification with generalized spectrum parameters," *Ultrasound Med. Biol.*, vol. 27, no. 11, pp. 1505–1514, 2001.
- [16] P. M. Shankar, V. a. Dumane, J. M. Reid, V. Genis, F. Forsberg, C. W. Piccoli, and B. B. Goldberg, "Classification of ultrasonic B-mode images of breast masses using Nakagami distribution," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 48, no. 2, pp. 569–580, Mar. 2001.
- [17] P. M. Shankar, V. a. Dumane, C. W. Piccoli, J. M. Reid, F. Forsberg, and B. B. Goldberg, "Computer-aided classification of breast masses in ultrasonic B-scans using a multiparameter approach," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 50, no. 8, pp. 1002–1009, Aug. 2003.
- [18] R.-F. Chang, W.-J. Wu, W. K. Moon, Y.-H. Chou, and D.-R. Chen, "Support vector machines for diagnosis of breast tumors on US images," *Acad. Radiol.*, vol. 10, no. 2, pp. 189–197, Feb. 2003.
- [19] R.-F. Chang, W.-J. Wu, W. K. Moon, and D.-R. Chen, "Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis," *Ultrasound Med. Biol.*, vol. 29, no. 5, pp. 679–686, May 2003.
- [20] Y.-L. Huang and D.-R. Chen, "Support vector machines in sonography: Application to decision making in the diagnosis of breast cancer," *Clin. Imag.*, vol. 29, no. 3, pp. 179–184, 2005.
- [21] K. Drukker, M. L. Giger, and E. B. Mendelson, "Computerized analysis of shadowing on breast ultrasound for improved lesion detection," *Med. Phys.*, vol. 30, no. 7, p. 1833, 2003.
- [22] K. Drukker, M. L. Giger, and C. E. Metz, "Robustness of computerized lesion detection and classification scheme across different breast US platforms," *Radiology*, vol. 237, no. 3, pp. 834–840, 2005.
- [23] S. Joo, Y. S. Yang, W. K. Moon, H. C. Kim, and A. Abstract, "Computer-aided diagnosis of solid breast nodules: Use of an artificial neural network based on multiple sonographic features," *IEEE Trans. Med. Imag.*, vol. 23, no. 10, pp. 1292–1300, Oct. 2004.
- [24] D.-R. Chen, R.-F. Chang, C.-J. Chen, M.-F. Ho, S.-J. Kuo, S.-T. Chen, S.-J. Hung, and W. K. Moon, "Classification of breast ultrasound images using fractal feature," *Clin. Imag.*, vol. 29, no. 4, pp. 235–245, 2005.
- [25] S. Alam, E. Feleppa, and M. Rondeau, "Ultrasonic multi-feature analysis procedure for computer-aided diagnosis of solid breast lesions," *Ultrason. Imag.*, vol. 38, no. 1, pp. 17–38, 2011.
- [26] W. Gomez, W. C. a. Pereira, and F. C. Infantosi, "Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound," *IEEE Trans. Med. Imag.*, vol. 31, no. 10, pp. 1889–1899, Oct. 2012.
- [27] M.-C. Yang, W. K. Moon, Y.-C. Wang, M. S. Bae, C.-S. Huang, J. H. Chen, and R.-F. Chang, "Robust texture analysis using multi-resolution gray-scale invariant features for breast sonographic tumor diagnosis," *IEEE Trans. Med. Imag.*, vol. 32, no. 12, pp. 2262–2273, Dec. 2013.
- [28] T. Tan, B. Platel, R. Mus, L. Tabar, R. M. Mann, and N. Karssemeijer, "Computer-aided detection of cancer in automated 3-D breast ultrasound," *IEEE Trans. Med. Imag.*, vol. 32, no. 9, pp. 1698–1706, Sep. 2013.
- [29] H. Cheng, J. Shan, W. Ju, Y. Guo, and L. Zhang, "Automated breast cancer detection and classification using ultrasound images: A survey," *Pattern Recognit.*, vol. 43, no. 1, pp. 299–317, Jan. 2010.
- [30] B. S. Garra, E. I. Cespedes, J. Ophir, S. R. Spratt, R. A. Zuurbier, C. M. Magnant, and M. F. Pennanen, "Elastography of breast lesions: Initial clinical results," *Radiology*, vol. 202, no. 1, pp. 79–86, 1997.
- [31] E. Burnside, T. Hall, and A. Sommer, "Differentiating benign from malignant solid breast masses with US strain Imaging1," *Radiology*, vol. 245, no. 2, pp. 401–410, 2007.
- [32] J. R. Doherty, G. E. Trahey, K. R. Nightingale, and M. L. Palmeri, "Acoustic radiation force elasticity imaging in diagnostic ultrasound," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 60, no. 4, pp. 685–701, Apr. 2013.
- [33] J. Bercoff, M. Tanter, and M. Fink, "Supersonic shear imaging: A new technique for soft tissue elasticity mapping," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 51, no. 4, pp. 396–409, Apr. 2004.
- [34] M. Tanter *et al.*, "Quantitative assessment of breast lesion viscoelasticity: Initial clinical results using supersonic shear imaging," *Ultrasound Med. Biol.*, vol. 34, no. 9, pp. 1373–1386, Sep. 2008.
- [35] A. Athanasiou, A. Tardivon, and M. Tanter, "Breast lesions: Quantitative elastography with supersonic shear imaging-preliminary results," *Radiology*, vol. 256, no. 1, pp. 297–303, 2010.
- [36] W. Berg, D. Cosgrove, and C. Doré, "Shear-wave elastography improves the specificity of breast us: The BE1 multinational study of 939 masses," *Radiology*, vol. 262, no. 2, pp. 435–449, 2012.
- [37] M. Moradi, P. Abolmaesumi, D. R. Siemens, E. E. Sauerbrei, A. H. Boag, and P. Mousavi, "Augmenting detection of prostate cancer in transrectal ultrasound images using SVM and RF time series," *IEEE Trans. Biomed. Eng.*, vol. 56, no. 9, pp. 2214–2224, Sep. 2009.
- [38] F. Imani, M. Daoud, M. Moradi, P. Abolmaesumi, and P. Mousavi, "Tissue classification using depth-dependent ultrasound time series analysis: In-vitro animal study," in *Soc. Opt. Photon. Technol.*, 2011, pp. 79680F–79680F-7.
- [39] M. Moradi, S. S. Mahdavi, G. Nir, E. C. Jones, S. L. Goldenberg, and S. E. Salcudean, "Ultrasound RF time series for tissue typing: First in vivo clinical results," *Soc. Opt. Photon. Technol.*, pp. 86701I–86701I-8, Mar. 2013.
- [40] M. Moradi, P. Abolmaesumi, and P. Mousavi, "Tissue typing using ultrasound RF time series: Experiments with animal tissue samples," *Med. Phys.*, vol. 37, no. 8, pp. 4401–4413, 2010.
- [41] F. Imani, P. Abolmaesumi, and E. Gibson, "Ultrasound-based characterization of prostate cancer: an in vivo clinical feasibility study," in *Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, 2013, vol. 8150, LNCS, pp. 279–286.

- [42] N. Uniyal, H. Eskandari, P. Abolmaesumi, S. Sojoudi, P. Gordon, L. Warren, R. N. Rohling, S. E. Salcudean, and M. Moradi, "A new approach to ultrasonic detection of malignant breast tumors," in *Proc. IEEE Int. Ultrason. Symp.*, Jul. 2013, pp. 96–99.
- [43] E. J. Feleppa *et al.*, "Typing of prostate tissue by ultrasonic spectrum analysis," *IEEE Trans. Ultrason. Ferroelectr. Freq. Control*, vol. 43, no. 4, pp. 609–619, 1996.
- [44] H. Eskandari, S. E. Salcudean, R. Rohling, A. Baghani, S. Frew, P. B. Gordon, and L. Warren, "Identifying malignant and benign breast lesions using vibroelastography," in *Proc. IEEE Int. Ultrason. Symp.*, Jul. 2013, pp. 25–28.
- [45] R. Zahir-Azar and S. E. Salcudean, "Motion estimation in ultrasound images using time domain cross correlation with prior estimates," *IEEE Trans. Biomed. Eng.*, vol. 53, no. 10, pp. 1990–2000, Oct. 2006.
- [46] B. Garra, B. Krasner, and S. Horii, "Improving the distinction between benign and malignant breast lesions: The value of sonographic texture analysis," *Ultrason. Imag.*, vol. 15, no. 4, pp. 267–285, 1993.
- [47] K. Horsch, M. L. Giger, and C. J. Vyborny, "Computerized diagnosis of breast lesions on ultrasound," *Med. Phys.*, vol. 29, no. 2, p. 157, 2002.
- [48] D. Chen, R. Chang, and W. Kuo, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound Med. Biol.*, vol. 28, no. 10, pp. 1301–1310, 2002.
- [49] M. Moradi, P. Mousavi, D. R. Siemens, E. E. Sauerbrei, P. Isotalo, A. Boag, and P. Abolmaesumi, "Discrete Fourier analysis of ultrasound RF time series for detection of prostate cancer," in *Proc. 29th Int. Conf. IEEE Eng. Med. Biol. Soc.*, Jan. 2007, pp. 1339–1342.
- [50] T. Higuchi, "Approach to an irregular time series on the basis of the fractal theory," *Physica D: Nonlinear Phenomena*, vol. 30, no. 2, pp. 277–283, 1988.
- [51] M. Moradi, P. Abolmaesumi, P. Isotalo, D. Siemens, E. Sauerbrei, and P. Mousavi, "Detection of prostate cancer from RF ultrasound echo signals using fractal analysis," in *Proc. 28th Annu. Int. Conf. IEEE Eng. Med. Biol. Soc.*, 2006, pp. 2400–2403.
- [52] F. Lizzi and M. Ostromogilsky, "Relationship of ultrasonic spectral parameters to features of tissue microstructure," *IEEE Trans. Ultrason., Ferroelectr. Freq. Control*, vol. 34, no. 3, pp. 319–329, May 1987.
- [53] F. Lizzi, E. Feleppa, S. K. Alam, and C. Deng, "Ultrasonic spectrum analysis for tissue evaluation," *Pattern Recognit. Lett.*, vol. 24, no. 2003, pp. 637–658, 2003.
- [54] F. Pedregosa and G. Varoquaux, "Scikit-learn: Machine learning in Python," *J. Mach. Learn. Res.*, vol. 12, pp. 2825–2830, 2011.
- [55] J. Platt, "Probabilistic outputs for support vector machines and comparisons to regularized likelihood methods," *Adv. Large Margin Classifiers*, vol. 10, no. 3, pp. 61–74, 1999.
- [56] T. Hastie, R. Tibshirani, and J. J. H. Friedman, *The Elements of Statistical Learning*. New York: Springer, 2001, vol. 1.
- [57] L. Breiman, "Random forests," *Mach. Learn.*, vol. 45, no. 1, pp. 5–32, 2001.
- [58] A. Criminisi and J. Shotton, *Decision Forests for Computer Vision and Medical Image Analysis*. New York: Springer, 2013.
- [59] I. Guyon, J. Weston, S. Barnhill, and V. Vapnik, "Gene selection for cancer classification using support vector machines," *Mach. Learn.*, vol. 46, no. 1–3, pp. 389–422, Jan. 2002.
- [60] X. Robin *et al.*, "PROC: An open-source package for R and S+ to analyze and compare ROC curves," *BMC Bioinformat.*, vol. 12, no. 1, p. 77, Jan. 2011.
- [61] J. Hanley and B. McNeil, "A method of comparing the areas under receiver operating characteristic curves derived from the same cases," *Radiology*, vol. 148, no. 3, pp. 839–843, 1983.
- [62] N. Chawla, N. Japkowicz, and A. Kotcz, "Editorial: special issue on learning from imbalanced data sets," *ACM SIGKDD Explor. Newslett.*, vol. 6, no. 1, pp. 1–6, 2004.
- [63] H. He and E. Garcia, "Learning from imbalanced data," *IEEE Trans. Knowl. Data Eng.*, vol. 21, no. 9, pp. 1263–1284, Sep. 2009.
- [64] F. Imani *et al.*, "Ultrasound-guided characterization of interstitial ablated tissue using RF time series: feasibility study," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 6, pp. 1608–1618, Jun. 2013.
- [65] M. I. Daoud, P. Mousavi, F. Imani, R. Rohling, and P. Abolmaesumi, "Tissue classification using ultrasound-induced variations in acoustic backscattering features," *IEEE Trans. Biomed. Eng.*, vol. 60, no. 2, pp. 310–320, Feb. 2013.
- [66] C. M. Bishop, *Pattern Recognition and Machine Learning*. New York: Springer Science, 2006.
- [67] C. Chang and C. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–39, 2011.