Computer-Aided Lesion Diagnosis in Automated 3-D Breast Ultrasound Using Coronal Spiculation

Tao Tan*, Bram Platel, Henkjan Huisman, Clara I. Sánchez, Roel Mus, and Nico Karssemeijer, *Member, IEEE*

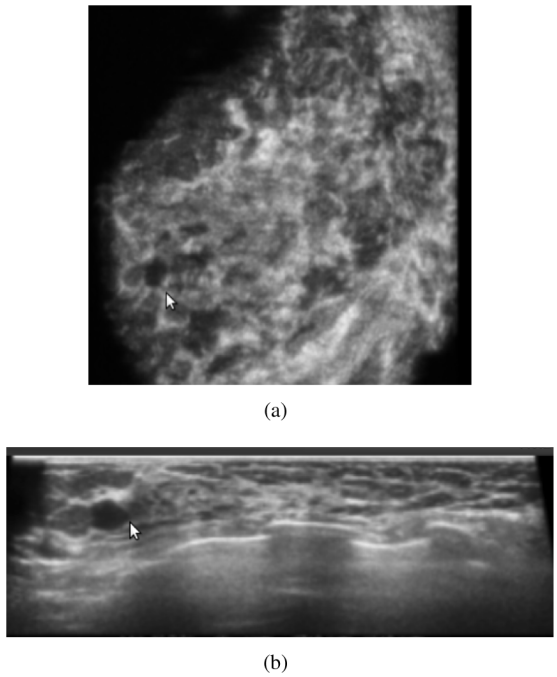
***Abstract*—A computer-aided diagnosis (CAD) system for the classification of lesions as malignant or benign in automated 3-D breast ultrasound (ABUS) images, is presented. Lesions are automatically segmented when a seed point is provided, using dynamic programming in combination with a spiral scanning technique. A novel aspect of ABUS imaging is the presence of spiculation patterns in coronal planes perpendicular to the transducer. Spiculation patterns are characteristic for malignant lesions. Therefore, we compute spiculation features and combine them with features related to echotexture, echogenicity, shape, posterior acoustic behavior and margins. Classification experiments were performed using a support vector machine classifier and evaluation was done with leave-one-patient-out cross-validation. Receiver operator characteristic (ROC) analysis was used to determine performance of the system on a dataset of 201 lesions. We found that spiculation was among the most discriminative features. Using all features, the area under the ROC curve (A_z) was 0.93, which was significantly higher than the performance without spiculation features ($A_z = 0.90, p = 0.02$). On a subset of 88 cases, classification performance of CAD ($A_z = 0.90$) was comparable to the average performance of 10 readers ($A_z = 0.87$).**

***Index Terms*—Automated 3-D breast ultrasound, computer-aided diagnosis (CAD), lesion segmentation, observer study, spiculation.**

I. Introduction

BREAST cancer is the most frequently diagnosed cancer and the leading cause of cancer death in females worldwide [1]. To reduce the mortality of breast cancer screening programs have been introduced. Because early detection reduces the risk of metastatic disease these programs are a key to successful treatment. Currently, mammography is the primary modality for breast cancer screening and diagnosis. As a complementary modality, targeted 2-D ultrasound (US) is widely used for diagnosing lesions, while most breast biopsies are performed with ultrasound guidance. For diagnosis, ultrasound has long been recognized as a valuable technique for characterization of cysts [2]–[4].

With new generations of the technology, the quality of US has strongly improved, increasing its potential for screening applications. In particular, there is an increasing interest to use US for screening in women with dense breasts, because of the poor performance of mammography in these women [5], [6]. Ultrasound has a high sensitivity for detecting invasive cancer in dense breasts. However, disadvantages of handheld ultrasound in screening are operator-dependency, a long examination time, and the inability to acquire and archive 3-D volumetric images of the breast. Automated ultrasound scanning approaches are developed to alleviate these problems. In a recent study, Kelly *et al.* [7] used an automated scanning technique operated by technicians, resulting in 2-D image sequences covering the entire breast which are presented to the radiologists as a movie. Detection sensitivity increased significantly when compared with mammography alone in dense breasts. Limitations of 2-D ultrasound can be overcome by automated 3-D breast ultrasound (ABUS). This new modality provides 3-D ultrasound images of large sections of the breast from the skin surface to the chest wall. The modality involves compression of the breast using a dedicated membrane and a wide transducer mounted in a scanning device. Fig. 1 shows a 3-D breast image generated by ABUS. Typically three or four views per breast are acquired that target different areas of the breast.



(a)

(b)

Fig. 1. Benign lesion in ABUS (indicated by the arrow). Two slices through the lesion are shown, a coronal slice (a) which is obtained from transforming the original stack of transversal images and a transversal slice (b) which is acquired by the probe at a particular location. The latter view is similar to a regular handheld ultrasound image.

Manuscript received November 07, 2011; revised January 10, 2012; accepted January 11, 2012. Date of publication January 16, 2012; date of current version May 02, 2012. This work was supported by the EU funded project HAMAM (IST-2007-224538) within the Seventh Framework Programme (FP7). *Asterisk indicates corresponding author.*

*T. Tan is with the Department of Radiology, Radboud University Nijmegen Medical Centre, 6525 GA Nijmegen, The Netherlands (e-mail: t.tan@rad.umcn.nl)

H. Huisman, C. I. Sanchez, R. Mus, and N. Karssemeijer are with the Department of Radiology, Radboud University Nijmegen Medical Centre, 6525 GA Nijmegen, The Netherlands.

B. Platel is with Fraunhofer MEVIS, 28359 Bremen, Germany.

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.2012.2184549

Computer-aided diagnosis (CAD) systems have been developed for different breast imaging modalities to support radiologists in detecting and diagnosing breast lesions more accurately and efficiently [8]–[11]. In screening mammography CAD is nowadays widely used. Development of CAD in 2-D breast ultrasound has received much attention in the literature [12]–[17], but applications have not yet been introduced in clinical practice. It is expected, however, that interest in breast ultrasound CAD will strongly increase when the technique is getting used in screening. Computerized features of breast lesions that have been reported are related to lesion shape, margin, posterior acoustic behavior, intensity distribution (echogenicity and echo texture).

Due to limited availability of 3-D acquisition devices, only few investigators have reported work in 3-D breast ultrasound. Wenkel *et al.* [18] showed that the use of ABUS achieved a high sensitivity in a selected patient group. Sahiner *et al.* [19] developed a CAD system based on extracted texture and morphological features of lesions in 3-D US volumetric images and it was shown that the use of this CAD system [9] significantly improved radiologists' accuracy in distinguishing malignant from benign breast masses. The data were obtained with an experimental 3-D breast ultrasound device, in which lesions were imaged after identification of their location using 2-D ultrasound and mammography and the images were acquired by manual translation of the transducer. Chen *et al.* [20] extracted the texture characteristics of lesions from 3-D US images obtained by an automated scanner using a transducer with a scan width of 40 mm. The 3-D US imaging in this study also required identification of lesion location using 2-D ultrasound. Ikedo *et al.* [21] developed an automatic scheme for detection of masses in whole breast ultrasound images in which 3-D characteristics were not used due to large interval between slices (2 mm). Chang *et al.* [22] developed a computer-aided detection system to automatically detect lesions from their whole breast images which were obtained by stitching three series of 2-D US images together. Moon *et al.* [23] applied the level set method to segment lesions and proposed a CAD system based on 3-D texture and morphology features using ABUS images, which yield an area under receiver operator characteristic (ROC) curve A_z of 0.95 on a dataset of 147 lesions. It should be noted that A_z values from different studies can not be directly compared because results depend strongly on the difficulty of the datasets.

In this paper, we present results of computerized analysis of lesions in ABUS data. A novel aspect of this modality is the ability to image high quality coronal views of the breast, representing slices in parallel to the skin surface. These views are strikingly different from regular B-mode ultrasound scans which radiologists are familiar with. In particular, spiculation patterns surrounding malignant lesions are often observed in coronal views, while spiculation cannot be observed well or is completely absent in 2-D handheld ultrasound. This additional value of coronal views of 3-D breast ultrasound was previously mentioned in [24] and [25]. In this paper we focus on automated classification of benign and malignant lesions, extending work previously reported [26]. In particular, we investigate features to quantify spiculation in 3-D ultrasound and to assess their value in classification of benign and malignant lesions. To our knowledge, this is the first paper in which the use of features extracted from coronal planes in breast ultrasound in a CAD system is investigated.

To assess the performance of the presented CAD system we performed experiments using a set of lesions with known pathology. To compare standalone performance of the system to that of radiologists an observer study was conducted in which a subset of the cases was presented to experienced breast imaging specialists.

II. Materials and Methods

A. Dataset

The breast ultrasound images used in this study are a representative sample of diagnostic cases obtained in routine clinical care from Radboud University Nijmegen Medical Centre (Nijmegen, The Netherlands), Falun Central Hospital (Falun, Sweden), and the Jules Bordet Institute (Brussels, Belgium). All data was anonymized. Institutional Review Board (IRB) approval was not required. All the centers used the automated breast ultrasound system developed by U-systems, Inc. (SomoVu, Sunnyvale, CA). Imaging involved frontal compression of the breast by a dedicated membrane, allowing a wide transducer with a width of 14.5 cm to move smoothly over the breast surface. The systems used were equiped with 8 Mhz transducers. In the ABUS scanning device the transducer is translated with a constant speed across the breast to obtain a 3-D volume of imaging data covering a large segment of the breast. Depending on the size of the breast, up to four views are acquired for each breast. Positioning and compression are to some extent standardized and include anteroposterior, lateral, medial or superior views. Each 3-D volumetric view was generated with an in-plane pixel size of 0.285 mm$\times$0.285 mm or 0.134 mm depending on the settings of the scan and a slice thickness of 0.6 mm. For processing, images were resampled to obtain 0.6 mm cubic voxels.

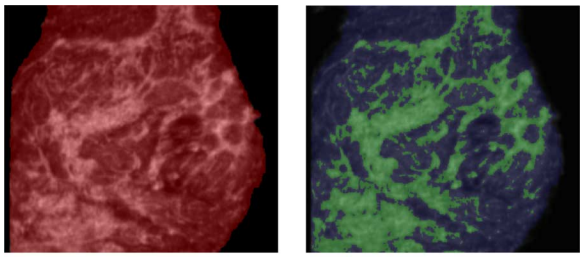
In this study, a complete dataset of ABUS studies from 158 patients with abnormalities was used. Among 610 views of those 158 patients, 179 views in which lesions are visible were used and annotated. Views in which a lesion is on the boundary of the image domain were not annotated and were excluded. In total, 201 lesions were annotated, including 85 malignant lesions and 116 benign lesions. Table I shows the distribution of lesions by histology types in the complete dataset and also the subset we used for an observer study (described in Subsection II-F). All malignant lesions were confirmed by biopsies and 38 benign lesions were confirmed by biopsies and the remaining benign lesions were confirmed by image interpretation alone or by combining information from different modalities.

B. Intensity Normalization

Images in our dataset were acquired over a longer period of time and in three institutes. Due to variations of scan parameter settings and use of different software versions the intensity levels of tissue types (fatty tissue, dense tissue, etc.) varied from patient to patient. Therefore, before lesion segmentation and feature computation, intensity normalization was performed on all images.

TABLE I
NUMBER OF LESIONS OF DIFFERENT HISTOLOGY TYPES IN THE COMPLETE DATASET AND OBSERVER STUDY (OS) DATASET

Histology	Complete dataset	OS dataset
Malignant lesions		
Infiltrative ductal carcinoma	60	38
Ductal carcinoma in situ	3	1
Infiltrative lobular carcinoma	7	6
Other	15	2
Total	85	47
Benign		
Cyst	47	9
Fibroadenoma	24	14
Fibrocystic change	24	11
Other	21	7
Total	116	41



(a) (b)

Fig. 2. (a) 2-D breast mask overlay (red) in a coronal plane. (b) Fatty tissue overlay (blue) and dense tissue overlay (green).

In ABUS imaging, US signal can only be acquired if a good contact between the probe and the skin exists. Therefore, only voxels in columns under the area with good contract hold valid information. The valid area on the skin is shown on coronal view in the resulted image. To be able to normalize intensities inside the breast in one image, a 2-D breast mask was generated in the coronal plane. The same 2-D mask is suitable for all coronal slices in the volume.

To this end, a coronal projection was obtained by averaging a series of 2-D coronal slices. The breast mask was obtained by applying Otsu's thresholding method [27] on the coronal projection. To avoid a negative effect of slices through the skin and thoracic volume on the normalization, only slices with depth between 10% and 50% of the total depth of the image volume were used in the projection.

After thresholding, in the projection, the boundary of the largest foreground connected component (breast) was refined by dilation and erosion operations. Holes inside the boundary caused by lesions or shadows were filled, obtaining the breast mask shown in Fig. 2(a).

Tissues in the breast, by rough approximation, can be categorized into fatty and dense tissue. We use this for normalization of the intensities in the 3-D volume. In the volume, within the breast mask, again using Otsu's thresholding, voxels with depth between 10% and 50% of the total image depth are categorized into fatty tissue and dense tissue [Fig. 2(b)]. Subsequently, the median intensities of fatty tissue and dense tissue are extracted I_f and I_d. The normalized intensity I for each voxel is computed as

$$I = \frac{(I'_d - I'_f)(I_o - I_d)}{I_d - I_f} + I'_d \qquad (1)$$

where I_o is the original voxel value and I'_d and I'_f are constants representing reference values for fatty tissue and dense tissue.

C. Lesion Segmentation

For classification of lesions a quantitative representation of lesion features has to be defined. As computation of such features generally requires an accurate delineation of the abnormality, segmentation is an important step in lesion classification algorithms. Various segmentation methods have been proposed to segment lesions in different modalities [14], [28]–[30]. In ultrasound, inhomogeneity of intensities inside lesions, background structure, partially undefined boundaries, and different posterior acoustic behavior of lesions, make lesion segmentation a difficult task. To obtain a reliable and computationally efficient lesion segmentation method we extended a spiral-scanning based dynamic programming technique, which was originally introduced by Wang *et al.* [29] for pulmonary nodule segmentation in CT.

The proposed method can be summarized as follows: First, using the center of the annotated lesion as a seed point, a volume of interest (VOI) is created by defining a sphere with a fixed radius (6 cm) at the lesion center. Second, by applying a spiral scanning technique [29] to the VOI, a 2-D resampled image is generated. Third, an optimal path in the resampled 2-D image which delineates the lesion boundary is determined by minimizing a cost function by dynamic programming [28], [29]. An example is shown in Fig. 3. To increase robustness, Wang's original approach is extended by using three spiral scans in orthogonal orientations instead of just one, resulting in three optimal paths. In this way the method is less dependent on the orientation of the spiral scanning. From these three paths edge point candidates are sampled in 3-D. The final lesion segmentation is obtained from these candidates by voxel classification as described in [29]. The classifier determines whether voxels are part of the lesion by taking into account the distance of the nearest edge candidate points to the lesion center.

An advantage of the proposed method is that it can be adapted to particular segmentation tasks by modification of the cost function determining the optimal path. While using the intensity gradient to detect edges locally, leakage of the segmentation can be prevented by means of a smoothness term in the cost function [29]. In addition, the cost function can be used to incorporate prior knowledge of the lesions to be segmented. In our cost function we used the expected radius (9 mm) of breast lesions [28]. This is important for lesions with ill defined boundaries and dark posterior shadowing. Fig. 4 shows the segmentation result of the same lesion shown in Fig. 3. In this example, the segmentation method manages to determine a good approximation of the surface of the lesion, despite the ill-defined boundaries,

D. Feature Computation

Given a segmentation V_l of a lesion, features can be computed to represent lesion properties as echotexture, echogenicity, shape, posterior acoustic behavior, and margins. These features are used for classification and should thus represent relevant aspects of breast lesions in ultrasound as good as possible. To define features, we have used descriptions in the radiological

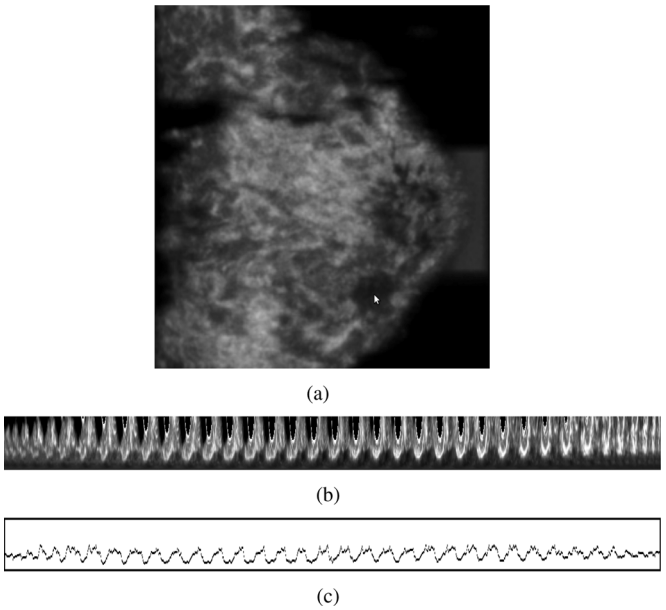


Fig. 3. (a) A 3-D ultrasound volume with the cursor indicating a malignant lesion in a coronal slice. Image (b) shows the resampled 2-D image obtained from one spiral scanning. Image (c) shows the optimal path delineating lesion boundary computed by the dynamic programming algorithm.

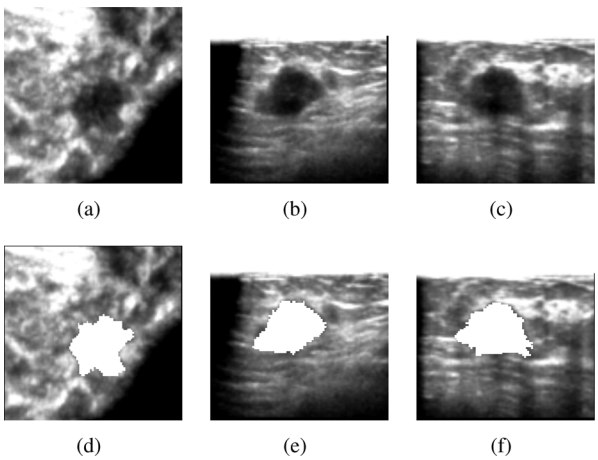


Fig. 4. Example of a segmentation result. Images (a), (b), and (c) show the coronal, transversal, and sagittal view of a malignant lesion, respectively. Images (d), (e), and (f) show the segmentation overlay (white) in coronal, transversal, and sagittal views, respectively.

literature as a guide [9], [20], [24], [25], [31]–[36]. Table II lists the 11 features investigated in this study and their abbreviations.

Spiculation patterns are common in mammography in association with invasive breast cancer, but are not observed well or are completely invisible in 2-D ultrasound. Interestingly, however, spiculation is often observed on coronal slices of 3-D ultrasound images [24], [25], [31] [Fig. 5(b)]. To extract features representing spiculation we use a method originally developed for mammography [37], [38]. In this method, in each voxel i a measure of linear concentration is computed from a 2-D line orientation map in the coronal plane through i. The line orientation map is computed by maximizing the directional second order Gaussian derivative at each voxel j over its orientation in the coronal plane. Indicating the line orientation at voxel j by

TABLE II
LESION FEATURES AND THEIR ABBREVIATIONS

Feature	Abbrevations
the mean of the spiculation values	SP_{mean}
the 90th percentile of the spiculation values	$SP_{90\%}$
the 75th percentile of the spiculation values	$SP_{75\%}$
variance of intensities	VI
entropy	H
average intensity	AI
margin contrast	MC
volumetric height-to-width ratio	$VHWR$
sphericity	SH
compactness	CP
posterior acoustic behavior	PAB

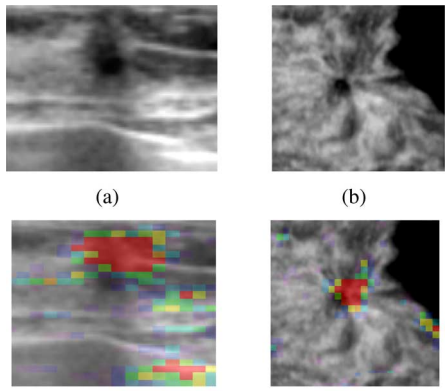


Fig. 5. Images (a) and (b) show a malignant lesion in transversal and coronal views, respectively. Spiculation can be observed well in the coronal view. Images (c) and (d) show the lesion with the spiculation feature map in color overlay. The spiculation value of the overlay increase with the color changing from blue to red.

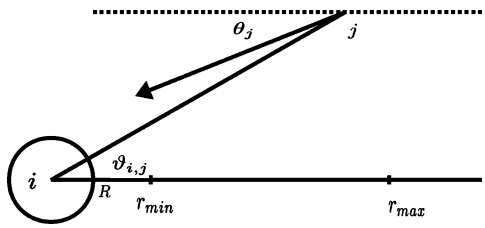


Fig. 6. Voxels j that are located at a distance $r_{r,j}$ between $r_{\min}$ and $r_{\max}$ and that are directed toward a disk of radius R centered at the test site i make contributions to spiculation calculation.

θ_j, the maximized magnitude of the directional derivative W_j has orientation $\theta_j + \pi/2$. To measure the concentration of line structures pointing to a central region of radius R at voxel i, a circular neighborhood is used as shown in Fig. 6. All voxels j with a distance r_{ij} ($r_{ij} \in [r_{\min}, r_{\max}]$) within the coronal planes to i are selected when the magnitude of W_j exceeds a small threshold. This selected set of voxels is denoted by N_i.

For any voxel j in N_i, a contribution x_j of voxel j to the concentration measure at site i is defined as

$$x_j = \begin{cases} 1 - p_j, & \text{if } h(\theta_j, \vartheta_{i,j}, r_{i,j}) = 1 \\ -p_j, & \text{else} \end{cases} \tag{2}$$

where p_j is the probability that voxel j would be pointing to the site i in a random pattern, defined by

$$p_j = \frac{2R}{\pi r_{i,j}} \tag{3}$$

and $h(\theta_j, \vartheta_{i,j}, r_{i,j})$ determines whether voxel j is oriented towards the center i, which is defined as

$$h(\theta_j, \vartheta_{i,j}, r_{i,j}) = \begin{cases} 1, & \text{if } |\vartheta_{i,j} - \theta_j| < \frac{R}{r_{i,j}} \\ 0, & \text{else} \end{cases} \tag{4}$$

where $\vartheta_{i,j}$ is the orientation of the line between voxel j and voxel i and θ_j is the line orientation at voxel j.

Using binomial statistics, the normalized spiculation measure [38] at i is defined by

$$f = \frac{\sum_{j \in N_i} x_j}{\sqrt{\sum_{j \in N_i} p_j (1 - p_j)}}. \tag{5}$$

To make this measure less dependent on the choice of the neighborhood size $(r_{\max})$, it is computed as a function of $r_{\max}$ (12 mm $< r_{\max} <$30 mm). The parameters $r_{\min}$ and R are both set to 3 mm. The maximum f_m of f over the range of neighborhood sizes $(r_{\max})$ was chosen as the spiculation value of voxel i.

The spiculation value f_m is computed for every voxel in the ABUS image, thus yielding a spiculation map in every coronal slice. An example is shown in Fig. 5. To extract lesion spiculation features from the voxel spiculation values, we computed the mean (SP_{mean}), the 90th percentile $(SP_{90\%})$, and 75th percentile $(SP_{75\%})$ of the spiculation values of voxels inside the lesion.

Echotexture and echogenicity of a lesion are often good indicators for the malignancy of a lesion [20], [32], [33]. Variance of intensities (VI), entropy (H) and the average intensity (AI) of voxels inside a lesion are computed to characterize these properties. Entropy (H) is defined as

$$H = -\sum_i p_i \cdot \log_2(p_i) \tag{6}$$

where p_i is the normalized histogram count of bin i computed from the intensities of voxels inside the segmented lesion.

There is usually a sharp demarcation between benign lesions and the surrounding tissue while malignant lesions often have indistinct margins [35]. We calculate a feature termed margin contrast (MC) by taking the difference of the average intensities in the inner and outer border of the lesion. The inner border consists of voxels inside the lesion whose distance to the lesion boundary is no more than 1.2 mm. The outer border consists of voxels outside the lesion within 1.2 mm of the boundary.

In breast ultrasound, it is well known that lesions that are wider than that they are tall, are more likely to be benign [34], [35]. This property is related to higher compressibility of benign lesions. To describe the wider-than-tall shape of lesions in ABUS, we introduce volumetric height-to-width ratio (VHWR) which is calculated as

$$\text{VHWR} = \frac{h}{w} \tag{7}$$

where h, the height, is defined by the difference of the maximal depth of the lesion $z_{\max}$ and the minimal depth $z_{\min}$. The width w is defined as the effective diameter of the coronal cross section of the lesion through the center, i.e., the diameter of the circle that has the same area as the cross section.

Due to different growth rates of breast lesions, malignant lesions tend to have an irregular shape and benign lesions, on the other hand, tend to have a spherical or oval shape [35]. Sphericity (SH) and compactness (CP) are used to characterize the irregularity of lesion shapes. Sphericity is defined as

$$SH = \frac{|V_l \cap S|}{|S|} \tag{8}$$

where S is a sphere at the lesion center which has equal volume as the segmented lesion. Compactness is defined as

$$\text{CP} = \frac{A^3}{|V_l|^2} \tag{9}$$

where A is the area of the surface of V_l.

The posterior acoustic behavior of a lesion is important for the classification as sound waves are attenuated differently in softer and harder tissues. Malignant lesions are harder than benign lesions on average [39], [40]. Therefore, in the posterior regions of lesions, benign solids tend to have no shadow, cysts usually have enhancement, and malignant lesions tend to have a dark shadow. The posterior acoustic behavior is a commonly analyzed feature in breast ultrasound [19], [36]. In this study, the posterior acoustic behavior feature (PAB) is computed as the difference between the average intensities of the region posterior to the lesion and the region surrounding the posterior region with the same depth range. The feature is defined as

$$\text{PAB} = \bar{I}_p - \bar{I}_s \tag{10}$$

where $\bar{I}_p$ and $\bar{I}_s$ are the average intensities of the posterior region V_p and the surrounding region V_s, respectively, which are defined as

$$V_p = \{(x, y, z) | (d(x, y) < 0) \wedge (z_{\max} < z < z_{\max} + l)\} \tag{11}$$

and

$$V_s = \{(x, y, z) | (0 < d(x, y) <= 2) \wedge (z_{\max} < z < z_{\max} + l)\} \tag{12}$$

where l is the predefined height of the posterior region, which is taken equal to the height h of the lesion in this study, $z_{\max}$ is the maximal depth of the lesion, and $d(x, y)$ is the signed distance map transformed from the binary 2-D projection in coronal view of the lesion. In the signed distance map, the absolute value of a pixel represents the shortest distance of that pixel to the region boundary. Pixels inside the border boundary are assigned negative values and pixels outside the boundary are assigned positive values.

E. Training and Evaluation

To automatically generate a score that indicates the malignancy of a lesion, a support vector machine (SVM) with a radial basis function (RBF) as the kernel was used as the classifier. In

order to evaluate the classifier performance and avoid possible bias, a nested leave-one-patient-out cross validation scheme was adopted. The inner cross-validation loop is used to perform the tuning of the SVM and kernel parameters using a grid search process while the outer loop is used to compute the estimate of the system performance. To investigate the benefits of incorporating coronal spiculation features, the experiments were performed using all features combined and using all features except spiculation features. The discriminative performance of the classification was evaluated by computing the area under ROC curve denoted as A_z. All ROC curves shown in this paper were fitted using the proper binormal model [41]. Performances of the CAD system with and without using spiculation features were compared using DBM MRMC 2.3 [42], [43]. Statistical analysis was performed using the fixed-reader with random-cases model. To investigate the effectiveness of each individual feature, we computed A_z values for each feature separately. To investigate how well the CAD system performs with solid masses using all features, we also performed ROC analysis on the 154 solid masses in our dataset, excluding the 47 cysts.

F. Observer Study

To determine how well the CAD system performs compared to human readers, an observer study was conducted. From the complete dataset, a total of 88 patients were included in this study, where for each patient one ABUS view with a lesion was selected. Cases for the study were manually selected to form a dataset with a roughly balanced number of malignant and benign lesions. During the selection, there is no preference on certain types of lesions except that most cases with cysts were left out because these are very easy to interpret. In the study set, 47 lesions were malignant and 41 were benign. The distribution of lesion types is shown in Table I. Ten readers (six radiologists and four residents) were invited to read the series of cases. Three radiologists and two resident had extensive experience with ABUS image interpretation. The remaining three radiologists and two residents had limited or no experience with ABUS but had extensive experience with interpreting 2-D ultrasound images. Readers rated the lesions on a continuous scale ranging from 0 (benign) to 100 (malignant). Before the study sessions, readers were presented a series of 26 training cases with 23 benign and 14 malignant lesions which were also randomly selected from the complete dataset of 201 lesions, to familiarize them with the 3-D viewing workstation. Performance of the readers was compared with that of standalone CAD using DBM MRMC 2.3 [42], [43], treating both readers and cases as random samples.

III. Results

Table III shows A_z values and the corresponding standard deviation obtained by the bootstrap method [44] using individual features on all 201 lesions. The A_z value of individual features varied from 0.51 to 0.84. MC and VHWR were the most discriminative features, both with an A_z value of 0.84. Three spiculation features (SP_{mean}, $SP_{75\%}$, and $SP_{90\%}$) led to A_z values of 0.83, 0.83, and 0.81. Among features characterizing the irregularity of lesion shapes, compactness (CP) had the best performance ($A_z = 0.82$). Posterior acoustic behavior feature (PAB) had an Az value 0.80. Scatter plots in feature space of MC and SP_{mean} and in feature space of MC and VHWR are shown in Figs. 7 and 8, respectively.

TABLE III
Performance of Feature(s)

Feature(s)	A_z	Std Dev
MC	0.84	0.03
$VHWR$	0.84	0.03
SP_{mean}	0.83	0.03
$SP_{75\%}$	0.83	0.03
CP	0.82	0.03
$SP_{90\%}$	0.81	0.03
PAB	0.80	0.03
H	0.72	0.04
VI	0.61	0.04
AI	0.59	0.04
SH	0.51	0.04
All features	0.93	0.02
All without spiculation features	0.90	0.02

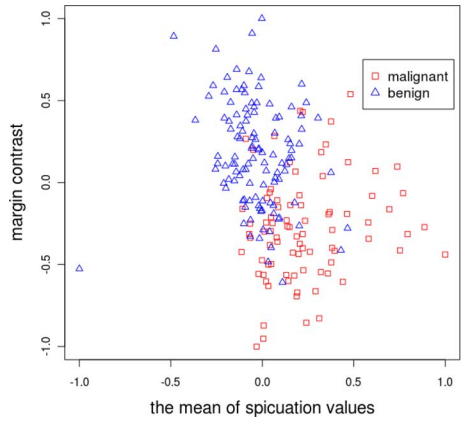


Fig. 7. Scatter plot in feature space of the mean of spicualtion values (SP_{mean}) and MC.

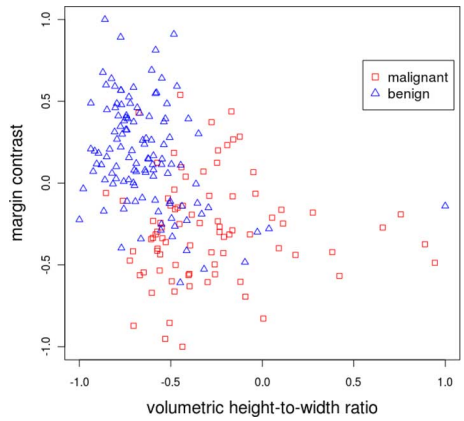


Fig. 8. Scatter plot in feature space of VHWR and MC.

The ROC curves of the CAD system with and without spiculation features on the dataset of all lesions are shown in Fig. 9. The CAD system achieved an A_z value of 0.93 using all the features and an A_z value of 0.90 without using spiculation features. The performance increase with spiculation was statistically significant ($p = 0.02$). On the complete dataset, it can be seen that the sensitivity of the CAD system using all features retained even high when the false positive (FP) fraction is reduced

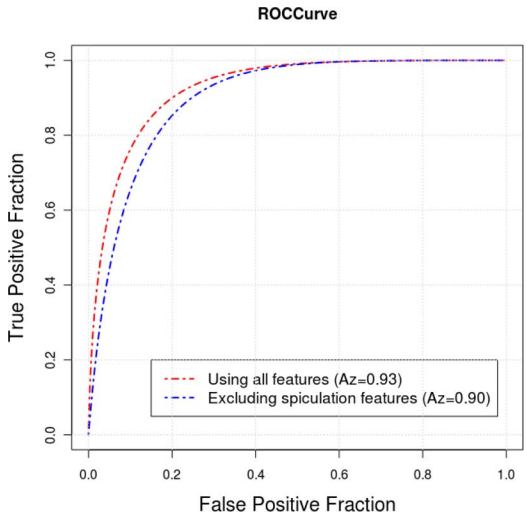


Fig. 9. Performances of the CAD system with and without spiculation features.

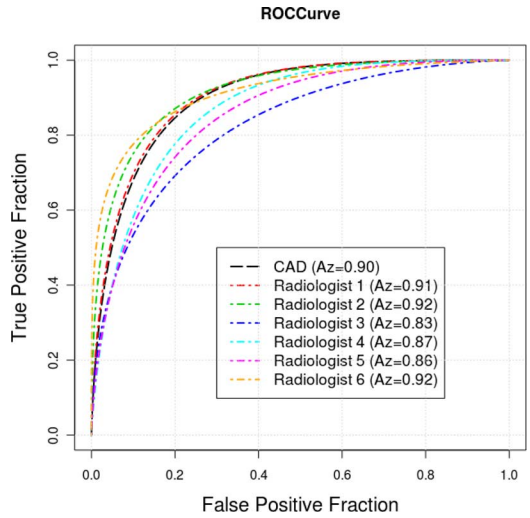


Fig. 10. Performances of the CAD system and the radiologisits.

TABLE IV
PERFORMANCE OF THE RADIOLOGISTS (RAD) AND THE RESIDENTS (RES) IN TERMS OF A_z VALUES

Reader	A_z	Std Dev
RAD1	0.91	0.03
RAD2	0.92	0.03
RAD3	0.83	0.04
RAD4	0.87	0.04
RAD5	0.86	0.04
RAD6	0.92	0.03
Mean RAD	0.88	0.03
RES1	0.87	0.04
RES2	0.77	0.05
RES3	0.86	0.04
RES4	0.89	0.03
Mean RES	0.85	0.04
Mean of All	0.87	0.03

from 1 to 0.4. At a FP rate of 0.4, the sensitivity is 0.98. On the dataset of solid lesions, the A_z value of the CAD system using all features was 0.90 ± 0.03.

Results of the reader performances from the observer study are shown in Table IV. The ROC curves of the CAD system and all radiologists are shown in Fig. 10. On the selected study cases our CAD system achieved an A_z value of 0.90 ± 0.04. The mean A_z values of the radiologists, the residents, and all readers were 0.88 ± 0.03, 0.85 ± 0.04, and 0.87 ± 0.03, respectively. The standard deviation for the mean Az values of the readers was obtained by treating readers and cases as random samples [42], [43]. The difference between the performance of CAD and all readers was not significant $(p = 0.26)$. Neither was the difference between the performance of CAD and the residents $(p = 0.14)$.

IV. CONCLUSION AND DISCUSSION

A computer aided diagnosis system was developed for the classification of malignant and benign lesions in ABUS. In the system, following automated lesion segmentation 11 features were computed to capture relevant lesion characteristics. Using a support vector machine as the classifier in combination with leave-one-patient-out cross-validation, an A_z value of 0.93 was obtained on a dataset of 201 lesions. The CAD system retained a sensitivity of 0.98 when the false positive rate was reduced from 1 to 0.4. This indicates that the CAD system has potential to improve diagnosis and that it also might be used to be more selective with biopsies in diagnostic procedures. It was found that the performance of the CAD system significantly improved when spiculation features were used $(p = 0.02)$. By conducting an observer study it was found that the CAD system had a performance comparable to that of experienced breast radiologists.

Our results highlight the diagnostic benefits that can be gained by using ABUS compared with regular 2-D ultrasound for breast lesion classification. Different to the work by Moon *et al.* [23], we focused on the contribution of spiculation features derived from ABUS. Spiculation features extracted from coronal planes of ABUS images turn out to be particularly effective, while this feature is difficult to observe in 2-D handheld ultrasound. The scatter plot (Fig. 7) shows malignant lesions tend to have higher spiculation values. In addition, VHWR was found to be one of the most discriminating features with an A_z value of 0.84. The effectiveness of this feature may result from the modest compression of the breast in ABUS. Differences in compressibility of benign and malignant lesions can result in lesion shape differences.

For some lesions the CAD system did not perform well. Fig. 11 shows a malignant lesion to which the CAD system assigned a low probability of being malignant. In the shown coronal plane, the segmentation failed to cover some of the lobes extended from the mass. Reasons might be that the intensity level of those lobes was different from that of the central region of the lesion. Also the smoothness constraint in the cost function may have prevented the lobes to be included in this case. With its oval shape and the absence of spiculation and posterior shadowing the CAD system rated this lesion as most likely benign, even though the inhomogeneity of intensity distribution inside the lesion was suggestive for malignancy. In our study, we found that variance of intensities (VI) and entropy (H) characterizing echotexture were not effective as single features. Reasons might be that those features are computed without considering spatial information of voxels inside the

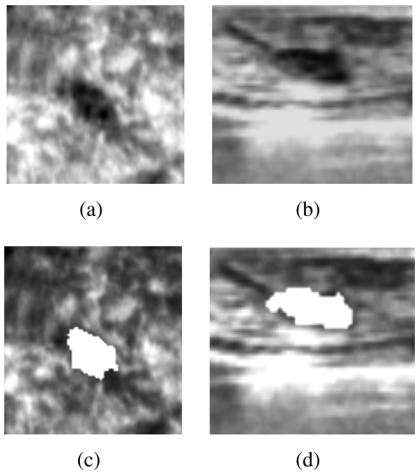


Fig. 11. An example of a malignant lesion which has a oval shape and has no shadowing in its posterior region. Images (a) and (b) show the lesion in coronal and transversal view, respectively. Images (c) and (d) show the white overlay of the lesion segmentation in coronal and transversal view, respectively.

lesion. The performance of the CAD system might be improved by incorporating more advanced texture features.

Lesion segmentation plays a crucial role in the CAD system since the computation of features related to lesion shape is largely dependent on the accuracy of segmentations. Due to ill-defined borders of breast lesions in ultrasound, in particular in the posterior regions, segmentation is often difficult. In this study we used a dynamic programming approach, because this is known to be an effective approach in situations where large sections of lesion boundaries are missing. In future work, it would be worthwhile to experiment with other segmentation approaches and to compare these to the method we used here.

Our study shows that MC is an effective feature in the classification task. This confirms the usefulness of computerized features extracted from lesion margins in breast ultrasound [12], [15], [36]. The computation of this feature depends on the choice of the width of inner border and outer border. Thus far we did not study the effect of border width parameters. Computing margin contrast at different scales (border widths) might be a feasible way to improve performance.

The Az value of the CAD system is higher than the average Az value of the observers. It should be noted, however, that most observers were lacking experience in reading ABUS images. Although they were familiar with the transversal sections displayed in the study, since these are similar to regular breast ultrasound, most readers did have little experience with interpreting lesions in coronal planes. In the study we failed to show a significant difference between the performance of CAD and the observers. This may be due to the fact that the power of the observer study was relatively low. We did use a fair number of lesions, but many were relatively easy to discriminate and therefore did not contribute much to the statistical power. Further, the CAD system is designed to assist radiologists. It is necessary to investigate whether the performance of observers improves with using CAD.

The CAD system was developed and validated with ABUS images from one manufacturer, while the technology is still under development. Newer systems use transducers with a higher frequency, resulting in a higher resolution. However, as long as some compression is used, the characteristics of lesions remain similar in 3-D breast ultrasound imaging. For example, benign lesions are more likely to have a wider-than-tall shape. Furthermore, the presence of spiculation patterns or architectural distortions associated with breast cancer were reported in 3-D breast ultrasound images acquired with different ultrasound systems [24], [25], [31]. Moreover, the intensity normalization we used makes feature computation less dependent on acquisition settings. Therefore, we expect that the system we developed is applicable to newer automated breast ultrasound systems, although retraining may be needed with data from these systems.

ACKNOWLEDGMENT

The authors would like to thank A. Grivegnée from Cancer Prevention and Screening Clinic, Jules Bordet Institute, Brussels, Belgium and L. Tabár from Department of Mammography, Falun Central Hospital, Sweden for providing the data.

REFERENCES

[1] A. Jemal, F. Bray, M. M. Center, J. Ferlay, E. Ward, and D. Forman, "Global cancer statistics," *CA Cancer J. Clinicians*, vol. 61, no. 2, pp. 69–90, 2011.

[2] S. V. Hilton, G. R. Leopold, L. K. Olson, and S. A. Willson, "Real-time breast sonography: Application in 300 consecutive patients.," *Am. J. Roentgenol.*, vol. 147, no. 3, pp. 479–486, 1986.

[3] D. B. Kopans, J. E. Meyer, and K. K. Lindfors, "Whole-breast US imaging: Four-year follow-up.," *Radiology*, vol. 157, no. 2, pp. 505–507, 1985.

[4] V. P. Jackson, "The role of us in breast imaging.," *Radiology*, vol. 177, no. 2, pp. 305–311, 1990.

[5] W. Berg, "Supplemental screening sonography in dense breasts.," *Radiologic Clin. N. Am.*, vol. 42, no. 5, pp. 845–851, 2004.

[6] T. M. Kolb, J. Lichy, and J. H. Newhouse, "Occult cancer in women with dense breasts: Detection with screening us—Diagnostic yield and tumor characteristics.," *Radiology*, vol. 207, no. 1, pp. 191–199, 1998.

[7] 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, 2010.

[8] H. P. Chan, B. Sahiner, M. A. Helvie, N. Petrick, M. A. Roubidoux, T. E. Wilson, D. D. Adler, C. Paramagul, J. S. Newman, and S. Sanjay-Gopal, "Improvement of radiologists' characterization of mammographic masses by using computer-aided diagnosis: An ROC study.," *Radiology*, vol. 212, no. 3, pp. 817–827, 1999.

[9] B. Sahiner, H.-P. Chan, M. A. Roubidoux, L. M. Hadjiiski, M. A. Helvie, C. Paramagul, J. Bailey, A. V. Nees, and C. Blane, "Malignant and benign breast masses on 3-D US volumetric images: Effect of computer-aided diagnosis on radiologist accuracy," *Radiology*, vol. 242, no. 3, pp. 716–724, 2007.

[10] L. A. Meinel, A. H. Stolpen, K. S. Berbaum, L. L. Fajardo, and J. M. Reinhardt, "Breast MRI lesion classification: Improved performance of human readers with a backpropagation neural network computer-aided diagnosis (CAD) system.," *J. Magn. Reson. Imag.*, vol. 25, no. 1, pp. 89–95, 2007.

[11] M. Samulski, R. Hupse, C. Boetes, R. Mus, G. den Heeten, and N. Karssemeijer, "Using computer aided detection in mammography as a decision support," *Eur. Radiol.*, vol. 20, no. 10, pp. 2323–2330, 2010.

[12] 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, 1999.

[13] K. G. Kim, J. H. Kim, and B. G. Min, "Classification of malignant and benign tumors using boundary characteristics in breast ultrasonograms.," *J. Digital Imag.*, vol. 15, pp. 224–227, 2002, Suppl 1.

[14] J.-Z. Cheng, Y.-H. Chou, C.-S. Huang, Y.-C. Chang, C.-M. Tiu, K.-W. Chen, and C.-M. Chen, "Computer-aided us diagnosis of breast lesions by using cell-based contour grouping.," *Radiology*, vol. 255, no. 3, pp. 746–754, 2010.

[15] A. V. Alvarenga, A. F. C. Infantosi, W. C. A. Pereira, and C. M. Azevedo, "Assessing the performance of morphological parameters in distinguishing breast tumors on ultrasound images," *Med. Eng. Phys.*, vol. 32, no. 1, pp. 49–56, 2010.

[16] H.-P. Chan, J. Wei, B. Sahiner, E. A. Rafferty, T. Wu, M. A. Roubidoux, R. H. Moore, D. B. Kopans, L. M. Hadjiiski, and M. A. Helvie, "Computer-aided detection system for breast masses on digital tomosynthesis mammograms: Preliminary experience," *Radiology*, vol. 237, no. 3, pp. 1075–1080, 2005.

[17] B. Liu, H. Cheng, J. Huang, J. Tian, X. Tang, and J. Liu, "Fully automatic and segmentation-robust classification of breast tumors based on local texture analysis of ultrasound images," *Pattern Recognit.*, vol. 43, no. 1, pp. 280–298, 2010.

[18] E. Wenkel, M. Heckmann, M. Heinrich, S. A. Schwab, M. Uder, R. Schulz-Wendtland, W. A. Bautz, and R. Janka, "Automated breast ultrasound: Lesion detection and bi-rads classification—A pilot study.," *RöFo*, vol. 180, no. 9, pp. 804–808, 2008.

[19] B. Sahiner, H.-P. Chan, M. A. Roubidoux, M. A. Helvie, L. M. Hadjiiski, A. Ramachandran, C. Paramagul, G. L. LeCarpentier, A. Nees, and C. Blane, "Computerized characterization of breast masses on three dimensional ultrasound volumes.," *Med. Phys.*, vol. 31, no. 4, pp. 744–754, 2004.

[20] D.-R. Chen, R.-F. Chang, W.-M. Chen, and W.-K. Moon, "Computer-aided diagnosis for 3-dimensional breast ultrasonography.," *Arch. Surg.*, vol. 138, no. 3, pp. 296–302, 2003.

[21] Y. Ikedo, D. Fukuoka, T. Hara, H. Fujita, E. Takada, T. Endo, and T. Morita, "Development of a fully automatic scheme for detection of masses in whole breast ultrasound images.," *Med. Phys.*, vol. 34, no. 11, pp. 4378–4388, 2007.

[22] R.-F. Chang, K.-C. Chang-Chien, E. Takada, C.-S. Huang, Y.-H. Chou, C.-M. Kuo, and J.-H. Chen, "Rapid image stitching and computer-aided detection for multipass automated breast ultrasound.," *Med. Phys.*, vol. 37, no. 5, pp. 2063–2073, 2010.

[23] W. K. Moon, Y.-W. Shen, C.-S. Huang, L.-R. Chiang, and R.-F. Chang, "Computer-aided diagnosis for the classification of breast masses in automated whole breast ultrasound images.," *Ultrasound Med. Biol.*, vol. 37, no. 4, pp. 539–548, 2011.

[24] D. Rotten, J. M. Levaillant, and L. Zerat, "Analysis of normal breast tissue and of solid breast masses using three-dimensional ultrasound mammography," *Ultrasound Obstetrics Gynecol.*, vol. 14, no. 2, pp. 114–124, 1999.

[25] D. O. Watermann, M. Fldi, A. Hanjalic-Beck, A. Hasenburg, A. Lghausen, H. Prmpeler, G. Gitsch, and E. Stickeler, "Three-dimensional ultrasound for the assessment of breast lesions," *Ultrasound Obstetrics Gynecol.*, vol. 25, no. 6, pp. 592–598, 2005.

[26] T. Tan, H. Huisman, B. Platel, A. Grivignee, R. Mus, and N. Karssemeijer, "Classification of breast lesions in automated 3-D breast ultrasound," in *Proc. SPIE Med. Imag.*, 2011, vol. 7963, p. 79630X.

[27] N. Otsu, "A threshold selection method from gray level histograms," *IEEE Trans. Syst., Man, Cybern.*, vol. 9, no. 1, pp. 62–66, Jan. 1979.

[28] S. Timp and N. Karssemeijer, "A new 2-D segmentation method based on dynamic programming applied to computer aided detection in mammography," *Med. Phys.*, vol. 31, no. 5, pp. 958–971, 2004.

[29] J. Wang, R. Engelmann, and Q. Li, "Segmentation of pulmonary nodules in three-dimensional CT images by use of a spiral-scanning technique," *Med. Phys.*, vol. 34, no. 12, pp. 4678–4689, 2007.

[30] M. Alemn-Flores, P. Alemn-Flores, L. Alvarez-Len, R. Fuentes-Pavn, and J. M. Santana-Montesdeoca, "Computer vision techniques for breast tumor ultrasound analysis.," *Breast J.*, vol. 14, no. 5, pp. 483–486, 2008.

[31] D. Kotsianos-Hermle, K. M. Hiltawsky, S. Wirth, T. Fischer, K. Friese, and M. Reiser, "Analysis of 107 breast lesions with automated 3-D ultrasound and comparison with mammography and manual ultrasound," *Eur. J. Radiol.*, vol. 71, no. 1, pp. 109–115, 2009.

[32] B. S. Garra, B. H. Krasner, S. C. Horii, S. Ascher, S. K. Mun, and R. K. Zeman, "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.

[33] C.-M. Chen, Y.-H. Chou, K.-C. Han, G.-S. Hung, C.-M. Tiu, H.-J. Chiou, and S.-Y. Chiou, "Breast lesions on sonograms: Computer-aided diagnosis with nearly setting-independent features and artificial neural networks.," *Radiology*, vol. 226, no. 2, pp. 504–514, 2003.

[34] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: Use of sonography to distinguish between benign and malignant lesions.," *Radiology*, vol. 196, no. 1, pp. 123–134, 1995.

[35] G. Rahbar, A. C. Sie, G. C. Hansen, J. S. Prince, M. L. Melany, H. E. Reynolds, V. P. Jackson, J. W. Sayre, and L. W. Bassett, "Benign versus malignant solid breast masses: US differentiation.," *Radiology*, vol. 213, no. 3, pp. 889–894, 1999.

[36] K. Horsch, M. L. Giger, L. A. Venta, and C. J. Vyborny, "Computerized diagnosis of breast lesions on ultrasound.," *Med. Phys.*, vol. 29, no. 2, pp. 157–164, 2002.

[37] N. Karssemeijer and G. M. te Brake, "Detection of stellate distortions in mammograms," *IEEE Trans. Med. Imag.*, vol. 15, no. 5, pp. 611–619, Oct. 1996.

[38] N. Karssemeijer, "Local orientation distribution as a function of spatial scale for detection of masses in mammograms," in *Information Processing in Medical Imaging*. New York: Springer, 1999, vol. 1613, pp. 280–293.

[39] K. Emerson, *Principles of Internal Medicine*. New York: McGraw Hill, 1974, pp. 582–587.

[40] E. E. Konofagou, "Quo vadis elasticity imaging?," *Ultrasonics*, vol. 42, no. 1–9, pp. 331–336, 2004.

[41] Metz and Pan, ""proper" binormal ROC curves: Theory and maximum-likelihood estimation," *J. Math. Psychol.*, vol. 43, no. 1, pp. 1–33, 1999.

[42] D. D. Dorfman, K. S. Berbaum, and C. E. Metz, "Receiver operating characteristic rating analysis: Generalization to the population of readers and patients with the jackknife method," *Invest. Radiol.*, vol. 27, pp. 723–731, 1992.

[43] S. L. Hillis, K. S. Berbaum, and C. E. Metz, "Recent developments in the Dorfman-Berbaum-Metz procedure for multireader ROC study analysis," *Acad. Radiol.*, vol. 15, no. 5, pp. 647–661, 2008.

[44] B. Efron, "Bootstrap methods: Another look at the jackknife," *Ann. Stat.*, vol. 7, no. 1, pp. 1–26, 1979.