

Received January 13, 2019, accepted January 30, 2019, date of publication February 15, 2019, date of current version March 4, 2019. *Digital Object Identifier 10.1109/ACCESS.2019.2898121*

A Novel Algorithm for Breast Lesion Detection Using Textons and Local Configuration Pattern Features With Ultrasound Imagery

U. RAJENDRA ACHARYA 1,2,3 , KRISTEN [M.](https://orcid.org/0000-0001-9719-4451) MEIBURGER 4 , JOEL EN WEI KOH 1 , EDWARD J. CIACCIO⁵, N. ARUNKUMAR^{®6}, MEE HOONG SEE⁷, NUR AISHAH MOHD TAIB⁷, ANUSHYA VIJAYANANTHAN⁸, KARTINI RAHMAT⁸, FARHANA FADZLI⁸, SOOK SAM LEONG⁹, CAROLINE JUDY WESTERHOUT⁸, AN[GE](https://orcid.org/0000-0002-1338-8820)LA CHANTRE-ASTAIZA¹⁰,

AND GUSTAVO RAMIREZ-GONZALEZ⁰¹¹

¹Department of Electronics and Computer Engineering, Ngee Ann Polytechnic, Singapore

²Department of Biomedical Engineering, School of Science and Technology, Singapore University of Social Sciences, Singapore

³Faculty of Health and Medical Sciences, School of Medicine, Taylor's University, Subang Jaya, Malaysia

¹⁰Department of Accounting, Economic and Administrative Sciences, University of Cauca, Popayan, Colombia ¹¹ Department of Electronic Engineering and Telecommunications, University of Cauca, Popayan, Colombia

Corresponding author: N. Arunkumar (arun.nura@gmail.com)

ABSTRACT Breast cancer is the most commonly occurring cancer in women worldwide. While mammography remains the gold standard in breast cancer screening, ultrasound is an important imaging modality for both screening and cancer diagnosis. This paper presents a novel method for the detection of breast lesions in ultrasound images using texton filter banks, local configuration pattern features, and classification, without employing any segmentation technique. The developed method was able to accurately detect and classify breast lesions and achieved an accuracy, sensitivity, specificity, and positive predictive

value of 96.1%, 96.5%, 95.3%, and 97.9%, respectively. The paradigm that we describe may, therefore, be useful as an effective tool to detect breast nodules during screening and in whole breast imaging, enabling clinicians to focus on images where a lesion is already known to be present. The developed method may also serve as a component for automatic breast nodule detection, and, when found, for the subsequent classification between lesion type benign versus malignant.

INDEX TERMS Breast, ultrasound, image, texton, local configuration pattern, malignant, benign, classifier.

I. INTRODUCTION

Breast cancer is a life-threatening cancer affecting women worldwide [1]. The frequency of new cases continues to rise, and the lifetime risk of being diagnosed with breast cancer for a woman in the United States is equal to 1-in-8 [2]. The early detection of breast cancer is extremely important for improved survival rate [3], and various imaging modalities are used in clinical practice for early detection and accurate assessment. Among imaging techniques, ultrasound (US)

The associate editor coordinating the review of this manuscript and approving it for publication was Victor Hugo Albuquerque.

has had a fundamental role in breast cancer detection for many years, thanks to its low cost, the fact that it does not use ionizing radiation, and its ability to assess various important aspects from multiple planes with high resolution. Specifically, ultrasound enables the assessment of orientation, morphology, lesion margin, and internal structure for both dense glandular structures and predominantly fatty breasts. Many features, such as surrounding tissue, margin contour, shape, posterior acoustic features, and lesion boundary, are important for correct lesion differentiation. The Breast Imaging Reporting and Data System (BI-RADS) of the American College of Radiology [4] is widely adopted for

⁴Department of Electronics and Telecommunications, Politecnico di Torino, Turin, Italy

⁵Department of Medicine, Columbia University, New York City, NY, USA

 6 Department of Electronics and Instrumentation, SASTRA University, Thanjavur, India

⁷Department of Surgery, University of Malaya, Kuala Lumpur, Malaysia

⁸Department of Biomedical Imaging, University of Malaya, Kuala Lumpur, Malaysia

⁹Department of Biomedical Imaging, University of Malaya Medical Centre, Kuala Lumpur, Malaysia

breast cancer screening. This classification system reduces the analysis variability between radiologists. The large study carried out by the American College of Radiology Imaging Network (ACRIN protocol 6666) has demonstrated how adding ultrasound imaging to the screening process alongside mammography enables identification of an additional 4.3 cancers per 1000 women screened [5]. Their work reported a decrease in specificity and an increase in sensitivity in detecting small breast cancers. However, they reported that performance of detection may increase marginally by using ultrasound images with mammogram images [5].

As a downside to ultrasound imaging, this technique is operator-dependent, and the imaging and diagnosis often present a high inter-operator variability. Due to these issues, design and implementation of computer aided detection (CAD) systems could be useful to provide a second opinion for detecting and classifying breast cancer [6].

There are numerous studies in the literature that present CAD systems for the differentiation between benign and malignant nodules in breast ultrasound (BUS) images [7]–[14]. However, all of these studies rely on a breast nodule being definitely present within the ultrasound image. As the use of ultrasound imaging in screening becomes more prevalent, it is important to focus attention not only on the differentiation between benign and malignant nodules, but also on automatic detection of the presence or absence of a breast mass. During the screening process, a large number of acquired ultrasound images would not present with a breast nodule. Therefore, the automatic detection of such nodules is of fundamental importance, and can be considered the first step in developing an automatic classification system that first detects whether a lesion is present or not, and then, if present, determines if it is benign or malignant.

Surprisingly, few published studies focus on automatic classification. Two studies, by Mogatadakala *et al.* [15] and Kutay *et al.* [16] showed satisfactory results in distinguishing tumorous vs non-tumorous regions ($AUC = 91\%$ and $AUC = 97\%$, respectively), but both of these studies were based on the raw RF signal, which is not typically given as output from clinical ultrasound devices. Other studies focus on whole breast ultrasound images, where 200-300 slices typically represent each lesion. Ikedo *et al.* [17] used 109 whole breast images (23 abnormal and 86 normal) and employed an edge direction technique, obtaining an 80.6% sensitivity with 3.8 false positives (FPs) per whole breast image. The same group further extended their database and with 260 whole breast images (208 normal and 52 with masses) using a similar but improved version of the previous algorithm [18] were able to obtain an 80% sensitivity in mass detection at 16.8 FPs per breast. Moon *et al.* [19] presented an algorithm using 148 whole breast images, where however each volume included a lesion, so that the problem was focused on determining which slices contained an actual lesion. Their technique was based on a 3D mean shift and fuzzy c-means algorithm, and they obtained an 89% sensitivity with 2 FPs per volume. Drukker *et al.* [20] presented a first study that used 757 traditional 2D ultrasound images of the breast, but in this study the authors focused on the classification of breast masses (complex cysts, benign nodules and malignant nodules), and only constructed 36 normal images from the acquired data. The algorithm rendered a sensitivity equal to 87% at 0.76 FPs, but showed 6 FPs on the 36 constructed healthy images. The same group continued their work [21], focusing more on the detection between normal and pathological images, using 757 training images and 1740 test images, which included 578 normal breast images. In their study, the authors detected potential lesions based upon expected lesion shape and margin characteristics, and then classified the candidate lesions by a Bayesian neural network developed from lesion features. The method provided an accuracy of 94% for the training set and 91% for the test set.

Herein, we present an automatic method for the first classification problem; hence the determination of the presence or absence of a breast mass within the ultrasound image. The developed technique is based on texton and local configuration pattern features and does not require any image segmentation or potential lesion detection.

FIGURE 1. Example breast images: (A) normal breast tissue; (B) benign nodule; (C) malignant nodule.

II. MATERIALS AND METHODS

A. IMAGE DATABASE

In this study, we used a database consisting of 448 B-mode ultrasound images of breast cancer of three different types: 147 images of a normal breast, 210 images of a benign nodule, and 91 images of a malignant nodule. Figure 1 shows examples of the US images acquired for each type. The images were collected from 282 patients (100 patients with benign nodules, 82 patients with malignant nodules, and 100 patients with normal breast) at the University Malaya Medical Centre. All images were acquired with a high resolution ultrasound system using a 5-12 MHz linear transducer (IU-22 Philips Medical System, Seattle, USA) and were subsequently exported for offline processing. The benignity/malignancy of the breast nodules were confirmed with biopsy.

B. AUTOMATIC BREAST NODULE DETECTION

Figure 2 shows a flow chart of the proposed algorithm, which is based on preprocessing, the calculation of texton features, the extraction of local binary pattern features, synthetic data sampling, feature reduction, and classification.

IEEE Access®

FIGURE 2. Flow chart of the proposed algorithm.

1) PREPROCESSING

The first step in the proposed breast nodule detection algorithm is image preprocessing to standardize the intensity distribution and to increase contrast. In order to do so, an adaptive histogram equalization algorithm [22] was applied to each image.

2) TWO-DIMENSIONAL TEXTON FORMATION

The spatial disposition of colors and regions in images along with the image pixel intensities together make up complex visual patterns that can be systematically characterized via a textural analysis [23]–[25]. Textons are the basic microstructures in images, and they can be used to construct detailed pixel-based textural features. In practice, textons are obtained via the convolution of a particular image with a filter set, that therefore characterize various pixel relationships in specific areas of the image [26], [27]. In this study, the filter banks that were used to find the response vectors for each texton were the Leung-Malik (LM) filters [28], the Schmid filters [29], and the Maximum Response (MR) filters [30], presented in detail following.

3) LEUNG-MALIK (LM) FILTER BANK

The LM filter banks present the interesting characteristic of being rotationally variant, meaning that the LM filter derivatives alter the image according to filter orientation. The filter bank is composed of 48 filters: the first two Gaussian derivatives for six different directions and three scales, four Gaussian filters, and eight Laplacian of Gaussian filters [28]. The database images are convolved with the 48 filters, thereby producing the filter responses, which correspond to the textural characteristics for each texton. A graphical representation of these filters can be appreciated in the study by Acharya *et al.* [31]. Figure 3 shows examples of the obtained images after applying the LM filter bank in all three classes (normal, benign and malignant). The filter banks help to highlight the hidden minute changes present in the images.

4) SCHMID FILTER BANK

In contrast to the LM filter bank, the Schmid filter bank contains rotationally invariant filters, which are therefore orientation invariant [32]. Using these filters, texton clustering occurs in a higher dimensional space, making them very effective in capturing subtle changes that may be present in the images.

The Schmid filter bank is made up of thirteen filters, which are convolved with the database images. A graphical representation of these filters can be appreciated in the study

FIGURE 3. Example of images obtained after the Leung-Malik (LM) filter banks are applied to each of the image types in dataset: normal breast, benign nodule, and malignant nodule.

by Acharya *et al.* [31]. Figure 4 shows examples of obtained images after applying the Schmid filter bank in all three classes (normal, benign and malignant).

FIGURE 4. Example of images obtained after the Schmid filter banks are applied to each of the image types in dataset: normal breast, benign nodule, and malignant nodule.

5) MAXIMUM RESPONSE (MR) FILTERS

Similar to the Schmid filter bank, the MR filters are rotationally invariant and provide the maximum filter response over various orientations and scales [30]. Practically speaking, 38 different filter banks are used: an orientation filter is utilized over six different orientations at three scales (6 orientations $* 3$ scales = 18, 18 $*2$ orientations = 36 filters) in the first and second derivative, followed by Gaussian filtering and then Laplacian of Gaussian filtering. Once these filters are applied, the largest response for all different orientations is taken as the maximal response output. A graphical representation of these filters can be appreciated in the study by Acharya *et al.* [31].

In this study, we employed two different types of MR filters: MR8 and MR4. The MR8 filter bank results in eight responses: the responses from the Gaussian and Laplacian of Gaussian filters and six from the first and second derivatives, at the three different scales. On the other hand, the MR4 filter bank will record four responses: two of which are from the Gaussian and Laplacian of Gaussian filters, and two responses result from the first and second derivatives.

Figures 5 and 6 show examples of the obtained images after applying the MR8 and MR4 filter banks in all three classes (normal, benign and malignant).

6) LOCAL BINARY PATTERN EXTRACTION

Once all texton filter banks have been applied to the images, it is then necessary to extract features from the obtained images. In this study, we extracted the local binary pattern (LBP), which is a straightforward and efficient texture descriptor used in previous studies [33]. This texture descriptor is rotationally invariant and can be defined as:

$$
LBP_{P,R}(x) = \begin{cases} \sum_{P=0}^{P-1} s (g_P - g_c), & U(x) \le 2\\ P+1, & otherwise \end{cases}
$$
 (1)

where

$$
s(x) = \begin{cases} 1, & x \ge 0 \\ 0, & x < 0 \end{cases}
$$

Furthermore, P is the number of points on the circumference of a circular neighborhood. *g^c* is the central pixel intensity,

FIGURE 5. Example of images obtained after the Maximum Response 8 (MR8) filter banks are applied to each of the image types in the dataset: normal breast, benign nodule, and malignant nodule.

FIGURE 6. Example of images obtained after the Maximum Response 4 (MR4) filter banks are applied to each of the image types in the dataset: normal breast, benign nodule, and malignant nodule.

by an imbalanced data distribution. Adaptive synthetic sampling (ADASYN) is a method that has been used in prior work to increase classification accuracy by balancing the data samples, thereby reducing bias [34], [35]. As can be appreciated in Section 2.1, the original database is partially imbalanced, so we employed ADASYN to overcome this issue and balance the dataset. After ADASYN synthetic sampling, the database then consisted of 192 images of normal breast, 210 images of a benign nodule, and 215 images of a malignant nodule (Table 1).

TABLE 1. Details of data used.

8) FEATURE REDUCTION

Feature reduction is important to reduce the dimensionality of an original features set, aiming to maintain a large separation between inter-class data samples. The texton extraction of the LBP from each filtered image produces 36 features. Therefore, in the case of the application of the LM filter bank (48 filters), each image is represented with a total of 1728 features (36 \times 4 = 1728). In the case of the Schmid filter bank on the other hand, each image is represented with 468 features $(36 \times 13$ filters). In the case of the MR8 and MR4 filter banks, 288 and 144 features represent each image, respectively $(36 \times 8 \text{ and } 36 \times 4 \text{, respectively}).$

In order to reduce dimensionality, a locality sensitive discriminant analysis (LSDA) approach was employed separately to the features calculated using each filter bank.

whereas g_p , $p = 0,1, \ldots, P - 1$, is the pixel intensity of the P points. The created neighborhood is assigned with $U(x)$, which is a uniformity measurement that counts the number of bit transitions (from 0 to 1 and vice versa) in the circular domain. From the LBP calculated in the entire circular domain (considering therefore different angles), the frequency of the recurring pattern can be calculated and subsequently mapped to a histogram. In this study we used $P = 8$, 16, 24 and $R = 1, 2, 3$.

7) ADASYN SYNTHETIC SAMPLING

A very important detail in classification and learning is to present a balanced dataset to ensure that no bias is introduced

TABLE 2. Top 10 LSDA features obtained with the Leung-Malik (LM) filter banks ranked using the f-value.

#Mean= α ; SD= β

Feature	Normal			Benign		Malignant		
	α	β	α	β	α	β	p-value	f-value
LSDA26	0.9821	0.0024	0.9833	0.0026	0.9825	0.0007	0.0000	19.7613
LSDA14	0.9530	0.0027	0.9542	0.0046	0.9537	0.0002	0.0011	6.8688
LSDA17	0.6385	0.0419	0.6230	0.0761	0.6314	0.0041	0.0086	4.7901
LSDA28	0.6795	0.0636	0.6675	0.0497	0.6793	0.0114	0.0105	4.5898
LSDA24	0.7705	0.0352	0.7781	0.0292	0.7763	0.0097	0.0129	4.3781
LSDA3	0.6928	0.0257	0.6855	0.0385	0.6904	0.0012	0.0192	3.9772
LSDA29	0.9804	0.0061	0.9799	0.0047	0.9809	0.0010	0.0572	2.8741
LSDA5	0.3460	0.0247	0.3549	0.0683	0.3477	0.0016	0.0774	2.5692
LSDA22	0.6212	0.0951	0.6229	0.0883	0.6349	0.0187	0.1239	2.0952
LSDA ₂₀	0.6973	0.0687	0.6970	0.0709	0.7060	0.0062	0.1743	1.7520

TABLE 4. Top 10 LSDA features obtained with the Maximum Response 8 (MR8) filter banks ranked using the f-value.

The LSDA method presents the positive characteristic of preserving both discriminant and local geometrical structure [36], and has been used effectively in the medical image

analysis field [37]. The study by Cai *et al.* [36] presents a detailed description of the functionality of the LSDA technique.

Feature	Normal			Benign		Malignant		
	α	β	α	β	α	β	p-value	f-value
LSDA3	0.5416	0.1837	0.4451	0.1422	0.4024	0.1297	0.0000	43.8649
LSDA5	0.7273	0.0525	0.7513	0.0518	0.7753	0.0518	0.0000	43.2639
LSDA10	0.7115	0.0820	0.7436	0.1134	0.6630	0.0855	0.0000	38.8432
LSDA14	0.3258	0.0874	0.3014	0.1047	0.3542	0.0947	0.0000	16.0500
LSDA21	0.2569	0.0792	0.2789	0.0857	0.2349	0.1041	0.0000	12.4490
LSDA1	0.5525	0.2730	0.4768	0.2904	0.4109	0.3040	0.0000	12.0917
LSDA ₂	0.6842	0.0768	0.7057	0.0642	0.7097	0.0521	0.0002	8.9288
LSDA4	0.4028	0.0150	0.4037	0.0476	0.3923	0.0197	0.0002	8.6539
LSDA24	0.9335	0.0197	0.9274	0.0209	0.9335	0.0195	0.0017	6.4680
LSDA12	0.8688	0.0312	0.8797	0.0312	0.8757	0.0357	0.0038	5.6260

TABLE 5. Top 10 LSDA features obtained with the Maximum Response 4 (MR4) filter banks ranked using the f-value.

FIGURE 7. Scatterplot of LSDA features using Leung-Malik (LM) filter banks.

A total of 30 features remained after the application of LSDA, which were then ranked based on their f-value. Hence, since we considered four different filter banks in this study (i.e., LM, Schmid, MR8, and MR4) we obtained four different sets of 30 LSDA coefficients.

9) CLASSIFICATION

After the feature reduction, we employed various classification methods to distinguish automatically between the ultrasound images that presented a nodule and those that were normal. Specifically, we used decision tree (DT), linear discriminant analysis (LDA), quadratic discriminant analysis (QDA), support vector machine (SVM), k-nearest neighbor (k-NN), and the probabilistic neural network (PNN). SVM classifiers can be used with different kernel functions, and in this study we implemented both polynomial functions (i.e., polynomials 1, 2 and 3) and a radial basis function (RBF) [38]. For a more detailed description of these classification techniques, please refer to [37].

III. RESULTS

A. FEATURE EXTRACTION RESULTS

As mentioned in paragraph II section B.5, each image is overrepresented with a high number of features, where the number

FIGURE 8. Scatterplot of LSDA features using Maximum Response 4 (MR4) filter banks.

FIGURE 9. Scatterplot of LSDA features using Maximum Response 8 (MR8) filter banks.

of features depends on which filter bank is employed. Hence, the LSDA feature reduction method was used to reduce the high number of features to a standard number of 30 LSDA coefficients. The LSDA coefficients were then ranked based on their significance calculated with the f-value.

Tables 2–5 show the results of the highest 10 ranked LSDA coefficients for each feature bank (i.e., LM, Schmid, MR8 and MR4, respectively). Figure 7, 8, 9 and 10 show scatterplots of the two more significant LSDA coefficients for all four filter banks, which illustrate how these coefficients can

FIGURE 10. Scatterplot of LSDA features using Schmid (S) filter banks.

TABLE 6. Classification results for breast nodule determination using the Leung-Malik (LM) filter banks.

Classifier	No. feat.	TP	TN	FP	FN	Acc. $\frac{6}{2}$	PPV $(\%)$	Sens. $(\%)$	Spec. $(\%)$
DT	3	390	161	31	35	89.30	92.64	91.76	83.85
LDA	3	242		191	183	39.38	55.89	56.94	0.52
QDA	3	231	11	181	194	39.22	56.07	54.35	5.73
SVM Poly 1	3	117	122	70	308	38.74	62.57	27.53	63.54
SVM Poly 2	3	142	162	30	283	49.27	82.56	33.41	84.38
SVM Poly 3	$\overline{2}$	240	91	101	185	53.65	70.38	56.47	47.40
k -NN	3	389	149	43	36	87.20	90.05	91.53	77.60
PNN	3	247	8	184	178	41.33	57.31	58.12	4.17
SVM RBF	3	277	147	45	148	68.72	86.02	65.18	76.56

DT: DECISION TREE; LDA: LINEAR DISCRIMINANT ANALYSIS; QDA: QUADRATIC DISCRIMINANT ANALYSIS; SVM: SUPPORT VECTOR MACHINE; K-NN: K-NEAREST NEIGHBORS; PNN: PROBABILISTIC NEURAL NETWORK; RBF: RADIAL BAS FUNCTION; FEAT: FEATURES; TP: TRUE POSITIVES; TN: TRUE NEGATIVES; FP: FALSE POSITIVES; FN: FALSE NEGATIVES; ACC: ACCURACY; PPV: POSITIVE PREDICTIVE VALUE; SENS: SENSITIVITY; SPEC: SPECIFICITY.

be used to clearly distinguish between images with normal breast versus one with a nodule, and therefore provide a high classification accuracy.

B. CLASSIFICATION RESULTS

Figures 7 to 10 show scatterplots of LSDA features for the LM, MR4, MR8 and S filters, respectively. It is evident from these figures that MR8 has a greater discrimination power as compared to the other types, and hence has yielded the highest performance. The ranked LSDA coefficients are then used to classify the images with the techniques mentioned previously: DT, LDA, QDA, SVM with polynomial and RBF kernel functions, k-NN, and PNN. The classifiers were independently employed using the features extracted with each filter bank, in order to achieve the best classification result using the lowest number of features, and to be able to discuss which filter bank provides the highest classification accuracy. Tables 6–9 report the classification results obtained in terms of the confusion matrix, accuracy, positive predictive value, sensitivity, and specificity, for all filter banks used. As can be noted, the probabilistic neural network using the MR8 filter bank presented the best classification results using 18 features, with an accuracy of 96.11%, a positive predictive value of 97.85%, and sensitivity and specificity values equal to 96.47% and 95.31%, respectively. These results show how

TABLE 7. Classification results for breast nodule determination using the Schmid filter banks.

TABLE 8. Classification results for breast nodule determination using the Maximum Response 8 (MR8) filter banks.

Classifier	No. feat.	TР	TN	FP	FN	Acc. $\frac{(0)}{0}$	PPV $(\%)$	Sens. (0/0)	Spec. $\frac{6}{2}$
DT	10	361	148	44	64	82.50	89.14	84.94	77.08
LDA	16	353	143	49	72	80.39	87.81	83.06	74.48
QDA	15	330	140	52	95	76.18	86.39	77.65	72.92
SVM Poly 1	15	312	158	34	113	76.18	90.17	73.41	82.29
SVM Poly 2	18	381	177	15	44	90.44	96.21	89.65	92.19
SVM Poly 3	17	394	179	13	31	92.87	96.81	92.71	93.23
k -NN	18	392	171	21	33	91.25	94.92	92.24	89.06
PNN	18	410	183	9	15	96.11	97.85	96.47	95.31
SVM RBF	18	417	171	21	8	95.30	95.21	98.12	89.06

TABLE 9. Classification results for breast nodule determination using the Maximum Response 4 (MR4) filter banks.

the developed technique was able to correctly distinguish the presence or absence of a breast nodule.

IV. DISCUSSION AND CONCLUSIONS

In this work, we used a dataset of 147 images of normal breast tissue and 301 images of benign and malignant nodules to develop and assess an automated system for the recognition of abnormalities in breast tissue, based on 2D B-mode ultrasound breast images. We applied the LM, Schmid, MR8 and MR4 texton filter banks to each image, and then extracted features from the obtained images using the local binary pattern. An LSDA feature reduction approach was then employed to reduce the number of features to 30 for each filter bank.

We then showed how the LSDA coefficients obtained from the MR8 filter banks were able to correctly detect the presence or absence of a breast nodule using various classifiers with a satisfactory level of accuracy.

As can be appreciated from Figure 5, it can be noted how the MR8 features of the benign and malignant nodules can present some features that can be extracted with the local binary pattern that are distinct from the normal MR8 features. In particular, the fifth column shows an example of the filter bank that helps obtain a high accuracy performance equal to 96%.

Breast cancer continues to be the most common malignant tumor that women can face, and while a mammogram

acquired with x-ray imaging is still the gold standard for diagnosis, ultrasound breast imaging is becoming an important complementary modality to increase diagnostic accuracy [39].

Many studies are present in the literature that focus on the classification between benign and malignant nodules, but there are few studies focusing on the first fundamental classification process: the differentiation between normal and abnormal 2D ultrasound breast images. The techniques that do provide this classification are still mainly based on the segmentation of a potential lesion, which can be a source of error in classification [15], [40]. Moreover, as stated in the Introduction, ultrasound imaging is becoming increasingly appealing for breast cancer screening, as it does not use ionizing radiation and it is an inexpensive imaging modality. As ultrasound imaging becomes more fundamental for breast cancer screening, so does the need for highly accurate and reliable CAD methods that can aid the practitioner in distinguishing between normal breast tissue and breast tissue with abnormality. This is increasingly important as the amount of data also increases, allowing experts to focus their attention on dubious and/or critical cases.

There are several advantages to the method that we propose in this work, which are:

• The method does not attempt any potential nodule segmentation technique, which can provide an increase in classification accuracy;

• The application of texton filter banks and the subsequent feature extraction can distinguish basic microstructures that characterize various pixel relationships in specific areas of the image;

• The system showed high accuracy, sensitivity and specificity. In particular, the best developed system demonstrated an accuracy equal to 96.31%, a sensitivity equal to 96.47% and a specificity equal to 95.31%.

• The developed method requires no user interaction and is therefore automated.

• The technique does not require any volume imaging, but is rather based on the analysis of single 2D B-mode ultrasound images of the breast.

Limitations and Future Directions

A limitation to this study is the fact that the database contained only 282 patients for evaluation, with an image database equal to 488 images. Therefore, it is necessary to test the developed technique on a larger database to confirm and further evaluate the method.

The technique presented here can prove to be a useful tool for automated detection of breast abnormality, the first step in developing methodology that focuses upon the differentiation between benign and malignant nodules. In the future, we plan to further extend the image database and develop a method for this subsequent classification. In particular, with the extension of the database, we plan to implement a convolutional neural network approach for the detection and classification of breast masses.

REFERENCES

- [1] L. A. Torre, F. Bray, R. L. Siegel, J. Ferlay, J. Lortet-Tieulent, and A. Jemal, ''Global cancer statistics, 2012,'' *CA, Cancer J. Clin.*, vol. 65, no. 2, pp. 87–108, 2015.
- [2] C. DeSantis, J. Ma, L. Bryan, and A. Jemal, ''Breast cancer statistics, 2013,'' *CA. Cancer J. Clin.*, vol. 64, no. 1, pp. 52–62, Jan. 2014.
- [3] W. Duncan and G. R. Kerr, ''The curability of breast cancer,'' *BMJ*, vol. 2, no. 6039, pp. 781–783, Oct. 1976.
- [4] American College of Radiology. *BI-RADS*. [Online]. Available: https://www.acr.org/Clinical-Resources/Reporting-and-Data-Systems/Bi-Rads
- [5] W. A. Berg, A. I. Bandos, E. B. Mendelson, D. Lehrer, R. A. Jong, and E. D. Pisano, ''Ultrasound as the primary screening test for breast cancer: Analysis from ACRIN 6666,'' *J. Nat. Cancer Inst.*, vol. 108, no. 4, p. djv367, Apr. 2016.
- [6] H. D. 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.
- [7] P. Raha, R. V. Menon, and I. Chakrabarti, ''Fully automated computer aided diagnosis system for classification of breast mass from ultrasound images,'' in *Proc. Int. Conf. Wireless Commun., Signal Process. Netw. (WiSPNET)*, 2017, pp. 48–51.
- [8] B. K. Singh, K. Verma, and A. S. Thoke, ''Fuzzy cluster based neural network classifier for classifying breast tumors in ultrasound images,'' *Expert Syst. Appl.*, vol. 66, pp. 114–123, Dec. 2016.
- [9] Q. Zhang, H. Chang, L. Liu, A. Li, and Q. Huang, ''A computer-aided system for classification of breast tumors in ultrasound images via biclustering learning,'' in *Machine Learning and Cybernetics*. Berlin, Germany: Springer, 2014, pp. 24–32.
- [10] 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.
- [11] A. V. Alvarenga, W. C. A. Pereira, A. F. C. Infantosi, and C. M. Azevedo, ''Classifying breast tumours on ultrasound images using a hybrid classifier and texture features,'' in *Proc. IEEE Int. Symp. Intell. Signal Process.*, Oct. 2007, pp. 1–6.
- [12] X. Shi, H. D. Cheng, L. Hu, W. Ju, and J. Tian, "Detection and classification of masses in breast ultrasound images,'' *Digit. Signal Process.*, vol. 20, no. 3, pp. 824–836, May 2010.
- [13] Y. Su, Y. Wang, J. Jiao, and Y. Guo, ''Automatic detection and classification of breast tumors in ultrasonic images using texture and morphological features,'' *Open Med. Inform. J.*, vol. 5, pp. 26–37, 2011.
- [14] W. Gomez, W. C. A. Pereira, and A. 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.
- [15] K. V. Mogatadakala, K. D. Donohue, C. W. Piccoli, and F. Forsberg, ''Detection of breast lesion regions in ultrasound images using wavelets and order statistics,'' *Med. Phys.*, vol. 33, no. 4, pp. 840–849, Mar. 2006.
- [16] M. A. Kutay, A. P. Petropulu, and C. W. Piccoli, "Breast tissue characterization based on modeling of ultrasonic echoes using the power-law shot noise model,'' *Pattern Recognit. Lett.*, vol. 24, nos. 4–5, pp. 741–756, Feb. 2003.
- [17] Y. Ikedo et al., "Development of a fully automatic scheme for detection of masses in whole breast ultrasound images,'' *Med. Phys.*, vol. 34, no. 11, pp. 4378–4388, Nov. 2007.
- [18] Y. Ikedo *et al.*, "Improving mass detection performance by use of 3D difference filter in a whole breast ultrasonography screening system,'' *Proc. SPIE*, vol. 6915, p. 691523, Mar. 2008.
- [19] W. K. Moon et al., "Tumor detection in automated breast ultrasound images using quantitative tissue clustering,'' *Med. Phys.*, vol. 41, no. 4, p. 042901, Mar. 2014.
- [20] K. Drukker, M. L. Giger, K. Horsch, M. A. Kupinski, C. J. Vyborny, and E. B. Mendelson, ''Computerized lesion detection on breast ultrasound,'' *Med. Phys.*, vol. 29, no. 7, pp. 1438–1446, Jun. 2002.
- [21] K. Drukker, D. C. Edwards, M. L. Giger, R. M. Nishikawa, and C. E. Metz, ''Computerized detection and 3-way classification of breast lesions on ultrasound images,'' *Proc. SPIE*, vol. 5370, p. 1034, May 2004.
- [22] S. M. Pizer *et al.*, "Adaptive histogram equalization and its variations," *Comput. Vis., Graph., Image Process.*, vol. 39, no. 3, pp. 355–368, 1987.
- [23] M. Tuceryan and A. K. Jain, "Basic methods in computer vision: Texture analysis,'' in *The Handbook of Pattern Recognition and Computer Vision*. 1998, pp. 207–248.
- [24] M. Varma and A. Zisserman, "A statistical approach to texture classification from single images,'' *Int. J. Comput. Vis.*, vol. 62, nos. 1–2, pp. 61–81, 2005.
- [25] S.-C. Zhu, C.-E. Guo, Y. Wang, and Z. Xu, ''What are textons?'' *Int. J. Comput. Vis.*, vol. 62, nos. 1–2, pp. 121–143, 2005.
- [26] F. Jurie and B. Triggs, ''Creating efficient codebooks for visual recognition,'' in *Proc. 10th IEEE Int. Conf. Comput. Vis. (ICCV)*, vol. 1, Oct. 2005, pp. 604–610.
- [27] U. R. Acharya et al., "An integrated index for identification of fatty liver disease using radon transform and discrete cosine transform features in ultrasound images,'' *Inf. Fusion*, vol. 31, pp. 43–53, Sep. 2016.
- [28] T. Leung and J. Malik, ''Representing and recognizing the visual appearance of materials using three-dimensional textons,'' *Int. J. Comput. Vis.*, vol. 43, no. 1, pp. 29–44, Feb. 2001.
- [29] L. Zhang, X. Ye, T. Lambrou, W. Duan, N. Allinson, and N. J. Dudley, ''A supervised texton based approach for automatic segmentation and measurement of the fetal head and femur in 2D ultrasound images,'' *Phys. Med. Biol.*, vol. 61, no. 3, pp. 1095–1115, Feb. 2016.
- [30] J.-M. Geusebroek, A. W. M. Smeulders, and J. van de Weijer, "Fast anisotropic Gauss filtering,'' *IEEE Trans. Image Process.*, vol. 12, no. 8, pp. 938–943, Aug. 2003.
- [31] U. R. Acharya, S. Bhat, J. E. W. Koh, S. V. Bhandary, and H. Adeli, ''A novel algorithm to detect glaucoma risk using texton and local configuration pattern features extracted from fundus images,'' *Comput. Biol. Med.*, vol. 88, pp. 72–83, Sep. 2017.
- [32] C. Schmid, "Constructing models for content-based image retrieval," in *Proc. IEEE Comput. Soc. Conf. Comput. Vis. Pattern Recognit. (CVPR)*, vol. 2, Dec. 2001, pp. II-39–II-45.
- [33] U. R. Acharya *et al.*, "Shear wave elastography for characterization of breast lesions: Shearlet transform and local binary pattern histogram techniques,'' *Comput. Biol. Med.*, vol. 91, pp. 13–20, Dec. 2017.
- [34] H. He, Y. Bai, E. A. Garcia, and S. Li, "ADASYN: Adaptive synthetic sampling approach for imbalanced learning,'' in *Proc. IEEE Int. Joint Conf. Neural Netw., IEEE World Congr. Comput. Intell.*, Jun. 2008, pp. 1322–1328.
- [35] F. Molinari, U. Raghavendra, A. Gudigar, K. M. Meiburger, and U. R. Acharya, ''An efficient data mining framework for the characterization of symptomatic and asymptomatic carotid plaque using bidimensional empirical mode decomposition technique,'' *Med. Biol. Eng. Comput.*, vol. 56, no. 9, pp. 1579–1593, 2018.
- [36] D. Cai, X. He, K. Zhou, J. Han, and H. Bao, "Locality sensitive discriminant analysis,'' in *Proc. 20th Int. Joint Conf. Artif. Intell. (IJCAI)*, 2007, pp. 1713–1726.
- [37] U. R. Acharya et al., "Automated characterization of fatty liver disease and cirrhosis using curvelet transform and entropy features extracted from ultrasound images,'' *Comput. Biol. Med.*, vol. 79, pp. 250–258, Jun. 2016.
- [38] R. O. Duda, P. E. Hart, and D. G. Stork, *Pattern classification*, 2nd ed. Hoboken, NJ, USA: Wiley, 2001.
- [39] M. H. Yap *et al.*, ''Breast ultrasound lesions recognition: End-to-end deep learning approaches,'' *J. Med. Imag.*, vol. 6, no. 1, p. 011007, Oct. 2018.
- [40] K. Drukker, C. A. Sennett, and M. L. Giger, ''Computerized detection of breast cancer on automated breast ultrasound imaging of women with dense breasts,'' *Med. Phys.*, vol. 41, no. 1, p. 012901, Dec. 2013.

U. RAJENDRA ACHARYA received the Ph.D. degree from the National Institute of Technology Karnataka, Surathkal, India, and the D.Eng. degree from Chiba University, Japan. He is currently a Senior Faculty Member with Ngee Ann Polytechnic, Singapore. He is also an Adjunct Professor with Taylor's University, Malaysia, an Adjunct Faculty with the Singapore Institute of Technology–University of Glasgow, Singapore, and an Associate Faculty with the Singapore

University of Social Sciences, Singapore. He has published more than 400 papers in refereed international SCI-IF journals (345), international conference proceedings (42), books (17) with more than 20 000 citations in Google Scholar (with *h*-index of 75), and ResearchGate RG score of 47.1. He holds three patents. His major academic interests include biomedical signal processing, biomedical imaging, data mining, visualization, and biophysics for better healthcare design, delivery, and therapy. He is ranked in the top 1% of the Highly Cited Researchers for the last three consecutive years (2016, 2017, and 2018) in computer science according to the Essential Science Indicators of Thomson. He is involved in various funded projects, with grants worth more than 2 million SGD. He has served on the Editorial Board of many journals. He has served as a Guest Editor for many journals. Please visit https://scholar.google.com.sg/citations?user=8FjY99sAAAAJ&hl=en for more details.

KRISTEN M. MEIBURGER received the Ph.D. degree. She is currently an Assistant Professor, with time contract, with the Department of Electronics and Telecommunications, Politecnico di Torino, Torino, Italy. She has published numerous research papers in international conferences and journals. Her research interests include medical image and signal processing and the development of computer-aided diagnosis systems for various imaging modalities, such as OCTA, ultrasounds,

and photoacoustics. Please visit https://scholar.google.it/citations?user= JrPiIPgAAAAJ&hl=en for more details.

JOEL EN WEI KOH received the B.Eng. (Biomedical) degree from SIM University and the M.Sc. degree from Nanyang Technological University, Singapore. He is currently a Researcher with Ngee Ann Polytechnic, Singapore. He has published more than 31 journal papers in the past five years. His principal interests include artificial intelligence, biomedical imaging, and machine learning. Please visit https://www.researchgate.net/ profile/Joel_En_Wei_Koh for more information.

EDWARD J. CIACCIO received the Ph.D. degree from Rutgers. He is currently a Senior Research Scientist with the Division of Cardiology, Department of Medicine, Columbia University, and also with the Celiac Disease Center, Columbia University, and an Honorary Principal Research Fellow with the Division of Cardiology, Department of Medicine, Imperial College London. He has published over 100 peer-reviewed articles on such topics as biomedical signal processing of heart

electrograms, and image processing of villous atrophy in celiac disease patients. In computational biology, he has developed biophysical models of activation wavefront propagation for ventricular tachycardia and for atrial fibrillation. His research received a Paper of the Year Award from the *Heart Rhythm Journal*, in 2008. He has been a Keynote Speaker of the International Conference on Biomedical Engineering and Biotechnology, from 2013 to 2015, and the Conference on Innovation in Medicine and Healthcare 2014. He was a Faculty Speaker of the 1st Annual International Symposium on Ventricular Arrhythmias 2006, Atrial Signals 2015, the 13th Annual Congress of the European Cardiac Arrhythmia Society 2017, and Boston Signals Summit 2018. He is the Editor-in-Chief of the journals *Computers in Biology and Medicine* and *Informatics in Medicine Unlocked*.

N. Arunkumar received the B.E. degree the M.E. degree in electronics and communication engineering, with a specialization in biomedical engineering. He has a strong academic teaching and research experience of more than 11 years with SASTRA University, India. He is appreciated for his innovative research-oriented teaching related practical life experiences to the principles of engineering. He is active in research.

ANUSHYA VIJAYANANTHAN received the bachelor's degree from the University of Delhi, in 1992, and the M.Rad. degree from the University of Malaya, in 2001. She subsequently did her training stints overseas in vascular ultrasound with the Thane Ultrasound Center, Mumbai, in 2004. She was a Fellow of interventional radiology with the Gartnavel General Hospital, Glasgow, U.K., in 2006. She is currently a Professor with the Department of Biomedical Imaging, University of

Malaya. She is a Coordinator of the Continuing Biomedical Imaging Education Centre, responsible for the organization of courses in the department. She is also involved in medical education in the Faculty of Medicine and is currently the Coordinator of undergraduate OSCE. She also has a keen interest in magnetic resonance guided focused ultrasound surgery, especially for gyne pathologies. Her special interests include gynecology and vascular ultrasound and interventional radiology.

MEE HOONG SEE received the M.D. degree from University Putra Malaysia, in 2003, and the M.S. degree in general surgery from University Malaya, in 2011. She was an Honorary Fellow of the Breast Oncoplastic Surgical Training with Siriraj Hospital, Mahidol University, Bangkok, Thailand, in 2013. She joined the University Malaya Medical Faculty, in 2012, as a Clinical Lecturer, subspecializing in breast oncoplastic surgery. She is currently an Oncoplastic Breast Surgeon with

the University Malaya Medical Centre (UMMC) and the University Malaya Specialist Centre. She is also a Lecturer and an Oncoplastic Breast Surgeon with UMMC. She is an Investigator and a Co-Investigator for a few clinical trials such as risk for breast cancer and intraoperative radiotherapy (TARGIT-B trial), improving the outcome for breast conserving surgery and oncology treatment for Her-2 positive breast cancer (Samsung Trial). She is a member of the College of Surgeon and Breast Chapter, Academy of Medicine Malaysia. Her fields of interest include breast oncoplastic surgery, minimally invasive breast surgery (endoscopic), immediate and delayed breast and nipple reconstruction for breast cancer. She is also a PI and has been collaborated with other researchers for local, national, and international studies and grants.

KARTINI RAHMAT received the MBBS (Mal), FRCR, M.Rad. (Mal) degrees. She joined as an Academic Staff, in 2003, following post-graduate training with University Malaya and has received further subspeciality training from the Royal Perth Hospital WA and U.K. for Breast Imaging and Neuroimaging Fellowships. She is currently a Professor and a Radiology Consultant with the University of Malaya, Kuala Lumpur. She is active in clinical research serving currently as a Unit

Coordinator of biomedical research imaging and as a Collaborator for studies in neuro-oncology, degenerative disease, epilepsy, and breast oncology. She has published and presented numerous scientific papers in peer-reviewed journals and conferences. Her clinical interests and research contributions include neuroradiology, breast cancer imaging/interventions, advanced MRI, and ultrasound applications.

NUR AISHAH MOHD TAIB received the MBBS degree from UM, in 1995, the MRCSEd degree, in 2000, the Masters of Surgery degree from UM, in 2001, the Grad dip Genet Counsell degree from CSU Australia, in 2010, and the Doctor of Medicine degree from UM, in 2012. She has been a Professor with the Department of Surgery, University Malaya, since 2012. She has been the Cancer Research Thrust Lead with the Faculty of Medicine, University of Malaya, since 2017. Her

h-index is 14 and Researchgate score is 42.39. She has authored or coauthored over 130 publications in peer-reviewed journals. She published one of the earliest publications on breast cancer outcomes in Malaysia. Her areas of interest include strong correlation with patient care and outcomes. She is a member of the International Surgical Society. She is a Life Member of the Malaysian Oncology Society, Malaysia. She is a Principal Investigator of the Malaysian Breast Cancer Cohort Study. She is a Trainer for Masters of Surgery in Malaysia and is on the Board of Examiners for the exit viva for the Breast and Endocrine Surgery Subspecialty Exams in Malaysia. She has been a TA Fellow and a past Honorary Treasurer and a third term Council Member of the College of Surgeons of Malaysia, since 2015. She was the Chairperson of the Breast Chapter of the College of Surgeons of Malaysia, from 2015 to 2018. She has been a Vice President of Together Against Cancer, an advocacy group, since 2018.

FARHANA FADZLI received the MBChB (Hons.) degree in U.K., the MRCP degree in U.K., the M.Rad. degree from UM, the M.Med. (Diagnostic Radiology) (Sin) degree, and the FRCR degree from the University of Leicester, U.K., in 2005. She was a Houseman in U.K. She joined University Malaya, Malaysia, in 2007, as a Trainee Lecturer in radiology. During her time as a Trainee, she completed her radiology professional exams. She has a special interest in nuclear

medicine and breast imaging. Her recent publications include works in clinical radiology and breast imaging. Please see https://scholar.google.com/ citations?user=JrV5Uk4AAAAJ&hl=en for more information.

SOOK SAM LEONG received the bachelor's degree in medical imaging from the University of Teesside, U.K., and the master's degree in medical science from the University of Malaya, Malaysia, where she is currently pursuing the Ph.D. degree. She is currently a Senior Radiographer with the University of Malaya Medical Centre, Malaysia. She has a keen interest in ultrasonography and involved in multiple research projects in this field.

CAROLINE JUDY WESTERHOUT did her undergraduate training with the University of Malaya and qualified as a Doctor, in 2002. She completed the Masters of Radiology training with the University of Malaya and was also admitted as a Fellow to The Royal College of Radiologists, U.K., in 2010. She is currently a Clinical Radiologist and a Senior Lecturer with the University of Malaya. Her areas of interest include pediatric and breast imaging.

ANGELA CHANTRE-ASTAIZA received the master's degree in marketing and the Ph.D. degree in business economics from Rey Juan Carlos University. She is currently a Business Administrator with the University of Cauca, Popayan, Colombia, where she is also a Professor. Her research interest includes data analysis applied to different fields.

GUSTAVO RAMIREZ-GONZALEZ received the B.S. degree in electronic and telecommunications engineering and the M.S. degree in telematics engineering from the University of Cauca, Colombia, in 2001, and the Ph.D. degree in telematics engineering from the Universidad Carlos III de Madrid, Spain, in 2010. He is currently a Professor and a Researcher with the Department of Telematics, University of Cauca. He has participated in national and international projects in Colombia

and Spain. He has published several research papers in reputed journals. His research interests include image processing, secure communication, machine learning, and the Internet of Things. He has served as a Guest Editor for several special issues at many journals.

 \sim \sim \sim