Ultrasound in Med. & Biol., Vol. 45, No. 8, pp. 2213–2225, 2019

0301-5629/$ - see front matter

https://doi.org/10.1016/j.ultrasmedbio.2019.03.018

● *Original Contribution*

TEXTURE ANALYSIS BASED ON AUTO-MUTUAL INFORMATION FOR CLASSIFYING BREAST LESIONS WITH ULTRASOUND

WILFRIDO GÓMEZ-FLORES,*
ARTURO RODRÍGUEZ-CRISTERNA,* and WAGNER COELHO DE ALBUQUERQUE PEREIRA†
* Center for Research and Advanced Studies of the National Polytechnic Institute, 87138 Ciudad Victoria, Tamaulipas, Mexico; and
† Biomedical Engineering Program/COPPE, Federal University of Rio de Janeiro, 21941-972 Rio de Janeiro, Brazil

(*Received* 18 *July* 2018; *revised* 22 *March* 2019; *in final from* 26 *March* 2019)

Abstract—Described here is a novel texture extraction method based on auto-mutual information (AMI) for classifying breast lesions. The objective is to extract discriminating information found in the non-linear relationship of textures in breast ultrasound (BUS) images. The AMI method performs three basic tasks: (i) it transforms the input image using the ranklet transform to handle intensity variations of BUS images acquired with distinct ultrasound scanners; (ii) it extracts the AMI-based texture features in the horizontal and vertical directions from each ranklet image; and (iii) it classifies the breast lesions into benign and malignant classes, in which a support-vector machine is used as the underlying classifier. The image data set is composed of 2050 BUS images consisting of 1347 benign and 703 malignant tumors. Additionally, nine commonly used texture extraction methods proposed in the literature for BUS analysis are compared with the AMI method. The bootstrap method, which considers 1000 bootstrap samples, is used to evaluate classification performance. The experimental results indicate that the proposed approach outperforms its counterparts in terms of area under the receiver operating characteristic curve, sensitivity, specificity and Matthews correlation coefficient, with values of 0.82, 0.80, 0.85 and 0.63, respectively. These results suggest that the AMI method is suitable for breast lesion classification systems. (E-mail: wgomez@tamps.cinvestav.mx)

Key Words: **Breast ultrasound, Texture analysis, Auto-mutual information, Lesion classification.**

INTRODUCTION

Breast cancer is the most frequently diagnosed cancer and the leading cause of cancer death among women worldwide (Ferlay et al. 2015). Early diagnosis is a crucial factor in breast cancer treatment, where medical images are an important source of diagnostic information. Over the past three decades, mammography has been the most cost-effective imaging technique for early diagnosis of breast cancer (American Cancer Society 2015). However, in patients with dense breast tissue, palpable nodules and normal or inconclusive mammograms, breast ultrasound (BUS) has emerged as the most important adjunct technique to mammography (Kelly et al. 2010). BUS images are particularly effective in distinguishing cystic from solid masses. They are also useful for determining the probable malignancy of tumors by using sonographic features such as shape, margin, echo patterns and posterior features (D'Orsi et al. 2013).

By observing the sonographic features of breast lesions, radiologists can interpret a BUS image and make clinical recommendations. Note that the diagnosis depends on the expertise and training of the radiologist, which can lead to discrepancies in image interpretation and, consequently, to different recommendations for clinical conduct (Calas et al. 2010; Timmers et al. 2012). Aiming to improve the performance of radiologists, computer-aided diagnosis (CAD) systems have become useful tools for determining the general histological class of tumors, that is, benign or malignant, as defined by the World Health Organization to categorize breast lesions (Azzopardi et al. 1982). Radiologists can use the CAD outcome as a second opinion for making a decision to: (i) perform a biopsy if the lesion is classified as malignant, or (ii) recommend a routine follow-up study if the lesion is classified as benign (Cristerna et al. 2018).

Address correspondence to: Dr. Wilfrido Gómez-Flores, CINVESTAV, Ciudad Victoria, Tamaulipas, Mexico. E-mail: wgomez@tamps.cinvestav.mx

In medical images, spatial variations of pixel intensities can reveal subtle pathologic features that can be quantitatively characterized by a texture analysis for the purpose of clinical diagnosis (Al-Kadi et al. 2015). In the case of ultrasound images, such a spatial variation of intensities is produced by speckle patterns that can be quantified in terms of "surface roughness" of the image to characterize tumor echogenicity, which is useful for distinguishing between benign and malignant breast lesions (Al-Kadi et al. 2015; Cheng et al. 2010).

Several texture analysis approaches have been used for breast lesion classification, including the auto-correlation and auto-covariance coefficients (Chang et al. 2003; Horsch et al. 2002; Huang et al. 2006), the block difference of inverse probabilities and the block variation of local correlation coefficients (Huang et al. 2006), the complexity curve (Alvarenga et al. 2007), brightness measures (Joo et al. 2004; Shen et al. 2007), wavelet coefficients (Chen et al. 2002), the Nakagami distribution (Shankar et al. 2003), fractal dimension (Chen et al. 2005), local binary patterns (Liu et al. 2014), Law's energy measures (Singh et al. 2017), the gray-level difference matrix (Huang et al. 2005), the gray-level run length matrix (GLRLM) (Prabusankarlal et al. 2015) and the gray-level co-occurrence matrix (GLCM) (Alvarenga et al. 2007; Gómez et al. 2012; Yang et al. 2013).

In general, texture analysis approaches process grayscale images to obtain texture features that should distinguish between benign and malignant textures regardless of the origin of the image. However, BUS images acquired under the same conditions differ greatly from each other, as ultrasound is highly dependent on the characteristics of the scanner; therefore, a different ultrasound scanner may result in a different performance in the texture analysis (Lo et al. 2015). To overcome this inconvenience, considerable attention has been devoted to intensity-invariant texture classification, where the input BUS image is transformed into a grayscale invariant representation for making texture descriptors robust to monotonic intensity transformations. In this context, both the ranklet transform (Lo et al. 2015; Yang et al. 2013) and the phase congruency (Cai et al. 2015) are multiresolution and orientation-selective approaches that have been used to perform robust texture analysis. Hence, from a single grayscale image, several texture channels are generated, from which typical texture features are computed, such as GLCM-based features and local binary patterns (Cai et al. 2015; Yang et al. 2013). Note that the higher the number of texture channels, the larger is the feature space. This effect is critical when the number of images is limited (which is typical in medical data) because the evaluations of the performance of the classifier can be optimistically biased (Foley 1972).

On the other hand, Farmer (2015) has shown that images with complex textures (such as speckle patterns) are generated by acquisition systems modeled as non-linear systems; thus, using non-linear estimators for analyzing a fundamentally non-linear process is appropriate. Mutual information (MI) is widely used to measure the non-linear relationship between two random variables, unlike the correlation, which measures only linear relationships. Thus, the auto-mutual information (AMI) function can be used to measure the relative MI between an input signal and a lagged version of itself over successively increasing lags (Farmer 2015).

Auto-mutual information has been widely used to analyze time series in medical data for applications such as detecting Alzheimer's disease (Coronel et al. 2017), characterizing the cortical areas of schizophrenic patients (Na et al. 2002), detecting pain responses during sedation (Melia et al. 2014), sleepiness characterization (Melia et al. 2015) and so on. However, AMI is practically unused in the context of image analysis, although it has the potential to be applied to texture description in the same way as the auto-correlation or the auto-covariance features previously used for breast lesion classification (Chang et al. 2003; Horsch et al. 2002).

In this article, a novel AMI-based texture analysis approach for breast lesion classification is proposed. Additionally, to make the texture description invariant to monotonic intensity transformations, the AMI is computed in the horizontal and vertical directions from the ranklets of the input BUS image. Because two texture features are computed from each ranklet image, the resultant feature space tends to be reduced in comparison to other feature spaces based on multiresolution and orientation-selective approaches. A support vector machine (SVM) with a Gaussian kernel is used for classifying breast lesions into benign and malignant classes. Moreover, nine texture analysis approaches commonly used in the literature for BUS classification are evaluated and compared with the proposed approach.

METHODS

The proposed approach is based on the AMI function to extract texture features for distinguishing between benign and malignant breast lesions in ultrasound images. Additionally, with the aim of performing an intensity-invariant texture analysis, the input BUS image is processed by the ranklet transform before feature extraction. The proposed method comprises three main steps: (i) transforming the input image using the ranklet decomposition; (ii) extracting the AMI-based features in the horizontal and vertical directions from each ranklet image; and (iii) classifying the breast lesion into either the benign or the malignant classes. Figure 1 is a block diagram of the feature

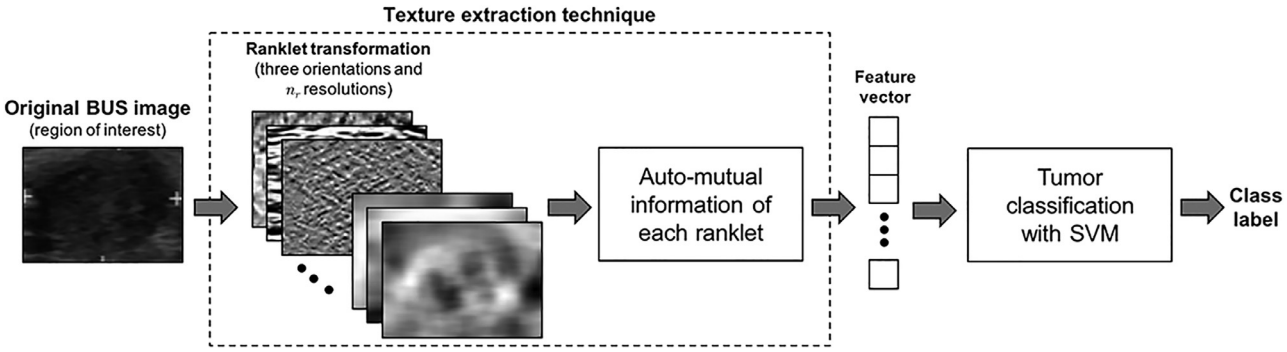

Fig. 1. Block diagram of the proposed approach. BUS = breast ultrasound.

extraction procedure for breast lesion classification followed in this study. This scheme represents a conventional approach to image classification; hence, the texture extraction technique module (*dashed rectangle*) can be substituted by other texture analysis methods to obtain a feature vector as the input to the SVM classifier.

Ranklet transform

The ranklet transform conducts a multiresolution, orientation-selective, non-parametric analysis to produce a series of transformed images referred to as *ranklet images*. It considers the relative rank of the pixels within a local region, instead of their grayscale values; hence, it can be defined as an operator invariant to linear and non-linear monotonic intensity transformations of the original image (Masotti and Campanini 2008).

The ranklet transform is applied to the input image in the spatial domain, where a single ranklet mask is defined in terms of both the resolution and the orientation parameters. Thus, the product of the number of resolutions and the number of orientations determines the total number of ranklet images.

Multiresolution analysis requires defining a set of resolution values that correspond to the linear sizes of the ranklet masks. Typically, the values of the resolutions are powers of 2. On the other hand, the orientation-selective analysis is performed by dividing the pixels within the ranklet mask into two subsets: treatment (T) and control (C). These subsets are defined differently for three orientations, namely, T_V and C_V for the vertical orientation, T_H and C_H for the horizontal orientation and T_D and C_D for the diagonal orientation (Masotti and Campanini 2008). Figure 2 illustrates the concepts of resolution and orientation for a ranklet mask with $r = 4$.

Finally, the non-parametric analysis is performed on the neighborhood covered by the ranklet mask. Here, the ranklet coefficients are calculated as (Yang et al. 2013)

$$R_j = \frac{\sum_{p \in T_j} \pi(p) - \frac{N}{4}\left(\frac{N}{2}+1\right)}{\frac{N^2}{4}} \tag{1}$$

where $j = \{V, H, D\}$ is the orientation; $\pi(p)$ denotes the pixel rank in the subset T_j; and $N = r^2$ is the number of pixels in the neighborhood. The ranklet coefficients are in the range [0,1], where values close to unity indicate that the pixels in T_j are brighter than the pixels in C_j, values close to zero indicate that the pixels in T_j are darker than the pixels in C_j and values close to 0.5 indicate that the pixels in T_j and C_j have similar intensities.

Because the computation of the MI is defined for discrete variables, the ranklet coefficients are quantized (Gómez et al. 2012):

$$R'_j = \lfloor q \cdot R_j + 1 \rfloor \tag{2}$$

Here, q is the desired number of quantization levels and $\lfloor \cdot \rfloor$ is the floor function; thus, the quantized ranklet coefficients are in the range [1, q]. Yang et al. (2013) propose quantizing the ranklet images to 256 levels before feature extraction, which is adopted herein.

Auto-mutual information

The MI of two discrete random variables X and Y measures the linear and non-linear dependencies between the two variables. MI is defined as (Cover and Thomas 1991)

$$I(X;Y) = \sum_{y \in Y} \sum_{x \in X} p(x,y) \log \frac{p(x,y)}{p(x)p(y)} \tag{3}$$

where $p(x,y)$ is the joint probability density function of X and Y, and $p(x)$ and $p(y)$ are the marginal probabilities of X and Y, respectively. Also, eqn (3) can be rewritten as (Cover and Thomas 1991)

$$I(X;Y) = H(X) - H(X|Y) = H(Y) - H(Y|X) \tag{4}$$

where $H(X)$ and $H(Y)$ are the entropies, and $H(X|Y)$ and $H(Y|X)$ are the conditional entropies. From eqn (4), MI is bounded by

$$0 \leq I(X;Y) \leq \min\{H(X), H(Y)\} \tag{5}$$

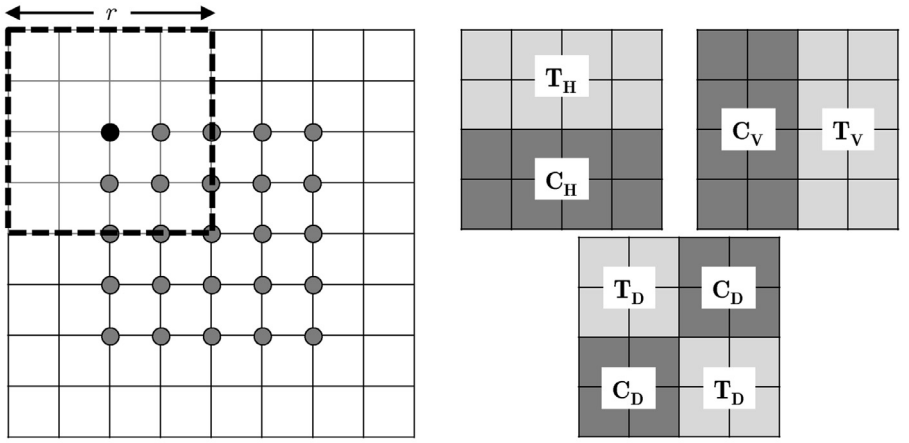

Fig. 2. (a) From an image of size 8 × 8, the ranklet mask (*dashed black square*) with resolution $r = 4$ generates 25 overlapping squares centered at gray points. (b) Three orientations of the ranklet mask in (a): horizontal (H), vertical (V) and diagonal (D), with their corresponding treatment (T) and control (C) subsets.

Note that in eqn (5), the MI is bounded above by the minimum of the entropies of X and Y, which could vary greatly; thus, this measure should be normalized to restrict its value to the range [0, 1]. The normalized MI can be calculated similarly to the Pearson correlation coefficient (Strehl and Ghosh 2002):

$$\tilde{I}(X;Y) = \frac{I(X;Y)}{\sqrt{H(X)H(Y)}}. \tag{6}$$

The AMI function proposed herein is inspired by the auto-correlation function proposed by Horsch et al. (2002) and uses the normalized version of MI in eqn (6). The AMI measures the degree of similarity between successive displacements of the input image, which can be performed in the horizontal and vertical directions. Let $f(x,y)$ be the input image with size $M \times N$. The AMI in the horizontal direction is defined as

$$A_x = \sum_{i=1}^{M-1} \tilde{I}(f(k_l, y); f(k_r, y)), \quad \forall y : 1 \le y \le N \tag{7}$$

where $k_l = 1, 2, \ldots, M-i$ and $k_r = i+1, i+2, \ldots, M$ denote the left and right displacements, respectively. Similarly, the AMI in the vertical direction is defined as

$$A_y = \sum_{i=1}^{N-1} \tilde{I}(f(x, k_u); f(x, k_d)), \quad \forall x : 1 \le x \le M \tag{8}$$

where $k_u = 1, 2, \ldots, N-i$ and $k_d = i+1, i+2, \ldots, N$ denote the up and down displacements, respectively.

In Figure 3 is an example of image displacements for computing the directional AMI in eqns (7) and (8). Note that vertical displacements create two subimages of the same size—the upper (*dark gray line*) and lower (*light gray line*) parts—from which the AMI is computed. At each displacement step, both subimages gradually become narrower and diminish their overlap until only the first and last rows remain. This same behavior is observed for horizontal displacements.

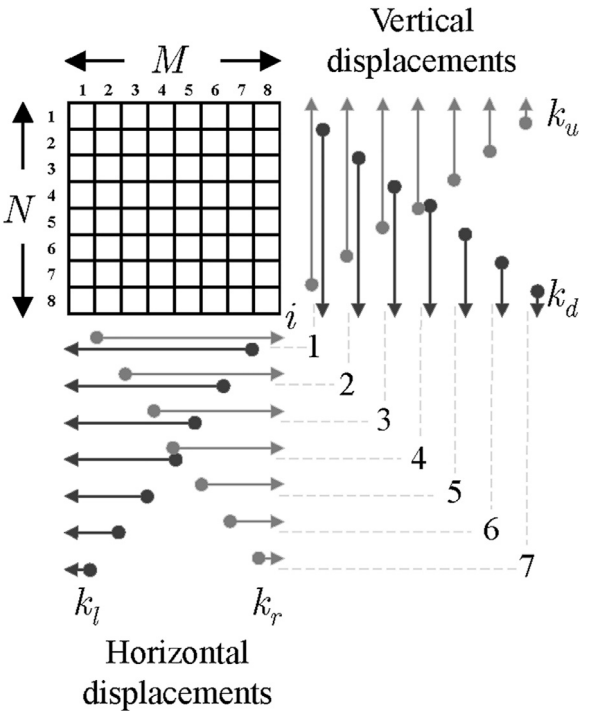

Fig. 3. Image displacements for computing the directional auto-mutual information. Dark gray and light gray lines mark the image displacements left/right (horizontal) and up/down (vertical), respectively. For an image of size 8 × 8, seven displacements are performed in the vertical and horizontal directions. BUS = breast ultrasound.

AMI-based texture features

Two steps are performed for calculating AMI-based texture features invariant to monotonic intensity transformations: (i) the input BUS image is converted to a series of ranklet images; and (ii) the directional AMI is computed on each ranklet image. Thus, given n_r resolutions and three orientations (horizontal, vertical and diagonal) for the ranklet transform and two directions (horizontal and vertical) for the AMI function, $n_r \times 6$ texture features are computed.

In the proposed approach, a radiologist manually outlines the breast lesion on the original image, from which a minimum rectangular region of interest (ROI) containing the lesion is cropped to compute ranklet images. Manual outlining is a simple and effective method for obtaining the breast lesion under analysis, although the segmentation generated by a computerized method can also be used to crop such an ROI. Hereafter, this texture extraction approach is referred to as the AMI method.

Algorithm 1 is the pseudocode of the AMI method. In addition, Figure 4 contains a BUS image and the ranklet images obtained from the ROI.

The running time depends on factors such as the programming language and computing platform. The AMI method was implemented in MATLAB 2014a (The MathWorks, Natick, MA, USA) because of its simplicity in manipulating matrices and control statements. However, the most time-consuming procedures are the computation of the ranklet transform and the MI; therefore, C language (Microsoft Visual Studio IDE, Redmond, WA, USA) was used to speed up both procedures, and they were compiled with MATLAB's mex function. The implementation of the AMI method is available from the authors on request. The computing platform employed a computer with a 3.60-GHz processor and 32 GB of RAM.

Algorithm 1. Pseudocode of the AMI method.

Input: input ROI image: $f(x, y)$; number of ranklet resolutions: n_r
Output: vector of AMI-based texture features: $\mathbf{x}$
Set ranklet resolutions: $\mathbf{r} \leftarrow \{2^1, \ldots, 2^{n_r}\}$
Set ranklet orientations: $\mathbf{o} \leftarrow \{\text{H, V, D}\}$
Initialize number of quantization levels: $q \leftarrow 256$
Initialize vector of texture features: $\mathbf{x} \leftarrow \emptyset$
For each $r \in \mathbf{r}$ **do**
  For each $o \in \mathbf{o}$ **do**
    Compute the ranklet image from $f(x, y)$; $R_{r,o}(x, y)$ (1)
    Quantize $R_{r,o}(x, y)$ with q levels: $R'_{r,o}(x, y)$ (2)
    Compute AMI in the horizontal direction from $R'_{r,o}(x, y)$: A_x (7)
    Compute AMI in the vertical direction from $R'_{r,o}(x, y)$: A_y (8)
    Concatenate the texture features: $\mathbf{x} \leftarrow \{\mathbf{x} \cup A_x \cup A_y\}$
  end
end
return x

Texture classification

The last step of the proposed approach is texture classification, where the AMI-based texture features are the inputs to an SVM classifier for distinguishing between benign and malignant tumors. This classifier is widely used for breast lesion classification because of its high generalization performance (Shan et al. 2016). Herein, the support vector machine library (LIBSVM) by Chang and Lin (2011) was employed to build an SVM using a Gaussian kernel.

The classification performance of the SVM depends on the choice of two hyperparameters—the penalty parameter C and the Gaussian kernel bandwidth γ. These hyperparameters are adjusted by using a procedure involving stratified k-fold cross-validation and grid search methods (Chang and Lin 2011; Hastie et al. 2009).

On the other hand, in practice, having more benign than malignant cases is expected (Guray and Sahin 2006); hence, supervised learning from a data set with an unbalanced class distribution could degrade the classification performance on the minority class. In this study, the class imbalance is considered during SVM training, in which a multiplicative coefficient adjusts the penalty parameter C for each class, which can be calculated by the relation (Ben-Hur and Weston 2010)

$$Cn = 2C_+n_+ = 2C_-n_- \tag{9}$$

where n_+ and n_- are the numbers of observations in the positive and negative classes, respectively, and $n = n_+ + n_-$. Thus, for the positive class, the C parameter is adjusted as $C_+ = Cn/2n_+$ and, similarly, the negative class as $C_- = Cn/2n_-$. With this strategy, errors in the minority class are more penalized than errors in the majority class.

Experiments

Breast ultrasound data set. The data set consisted of 2050 BUS images acquired during breast diagnostic procedures at the National Cancer Institute (INCa) in Rio de Janeiro, Brazil. The INCa Research Ethics Committee approved this study (Protocol 38/2001). Patients were informed of the purpose of the study before consenting to participate. The images were collected from three ultrasound scanners: LOGIQ 6p (GE Healthcare, Chicago, IL, USA), LOGIQ 5 (GE Healthcare) and Sonoline Sienna (Siemens Healthineers, Erlangen, Bavaria, Germany), with linear transducer arrays having frequencies between 7.5 and 12 MHz. The images were captured directly from the 8-bit video signal (*i.e.,* 256 gray levels) and saved in TIFF format. All the images were obtained from patients with posterior indication of biopsy, from which 1347 images revealed benign lesions, and 703 images malignant tumors. Figure 5 illustrates the histological types of

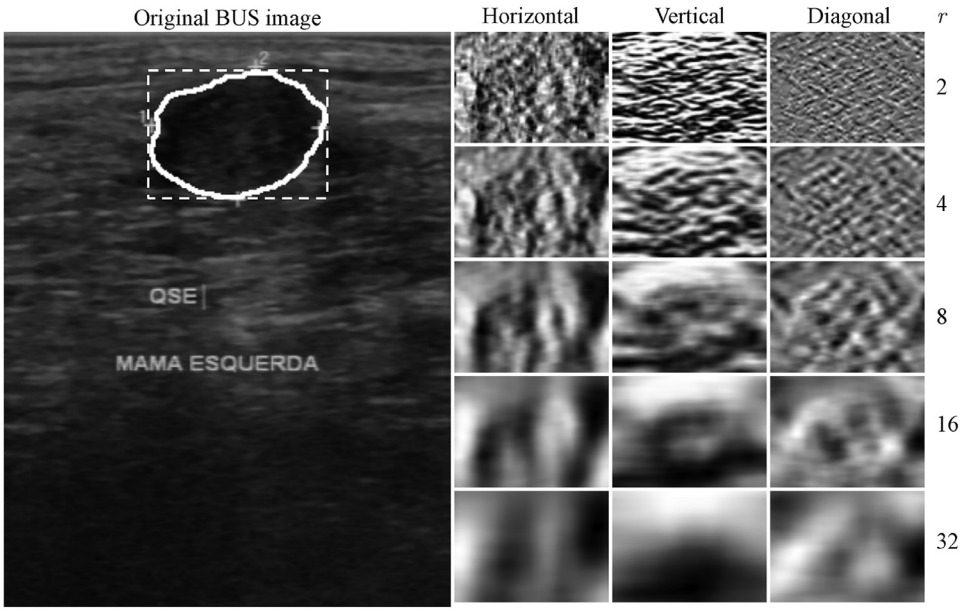

Fig. 4. Ranklet images obtained from the ROI (*white dashed square*) in the original BUS image. The lesion contour defined by the radiologist is delineated by a *white line*. From each ranklet image ROI, two directional AMI-based texture features are computed. BUS = breast ultrasound; ROI = region of interest.

lesions contained in the BUS data set, which reflects the diversity of textures that will be further characterized.

Each breast lesion in the data set was manually outlined by a senior radiologist with more than 15 years of experience in BUS interpretation. These outlines are useful to extract ROIs, from which the texture features were computed.

Classification performance assessment. By using the manual outline of the lesion, a minimum rectangular ROI is cropped from every image in the BUS data set, on which Algorithm 1 is run to create a vector of AMI-based texture features. For the experiments, five ranklet resolutions were evaluated.

From the resulting feature space, the SVM optimal hyperparameters were found by means of k-fold cross-validation (with $k = 10$) and a grid search, where the values searched were $C = \{2^{-5}, 2^{-4}, \ldots, 2^{15}\}$ and $\gamma = \{2^{-15}, 2^{-14}, \ldots, 2^{3}\}$ (Chang and Lin 2011).

The classification performance was assessed by the bootstrap method with 1000 independent bootstrap samples (Chatterjee and Chatterjee 1983). Note that for each bootstrap sample, the optimal pair (C, γ) was used to train the corresponding SVM. Before the classification, the feature space was rescaled to the range $[-1, 1]$ by softmax normalization to reduce the influence of extreme feature values (Priddy and Keller 2005).

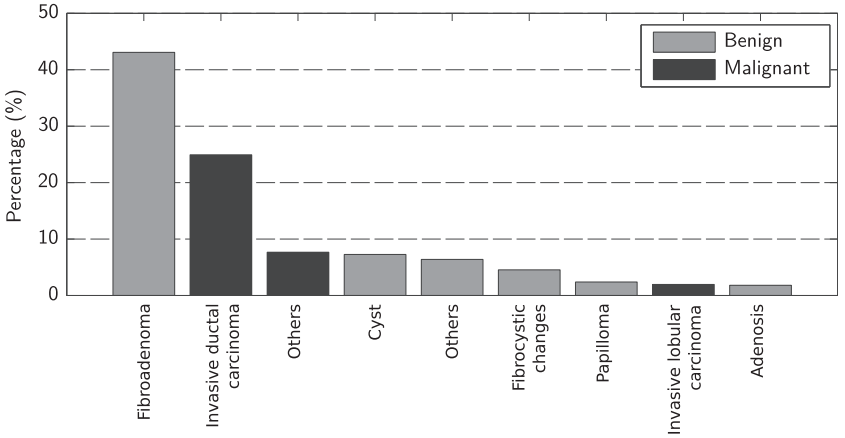

Fig. 5. Histological distribution of breast lesions in the image data set.

The outcome of the SVM is a score in the range $[-\infty, \infty]$ that represents the signed distance from the observation to the decision hyperplane. Next, the following classification rule was applied: a positive score indicates that the observation belongs to the malignant class; otherwise, the observation belongs to the benign class. From the classified observations, the following performance indices were calculated: area under the receiver operating characteristic curve (AUC), sensitivity, specificity, positive predictive value, negative predictive value and the Matthews correlation coefficient (MCC) (Baldi et al. 2000; Sokolova and Lapalme 2009).

In addition, the .632+ bootstrap estimator was used to obtain a low-variance measure of the prediction error, that is, the expected discrepancies between the classifier outputs and the correct responses (Efron and Tibshirani 1997).

Other texture analysis methods evaluated. Nine texture analysis methods proposed in the literature were considered for performance comparison with the proposed approach: auto-correlation coefficient (ACOR) (Horsch et al. 2002), auto-covariance coefficient (Chen et al. 2002), fractal features (Chen et al. 2005), Law's energy measures (Kriti et al. 2016), local binary pattern variance (Matsumoto et al. 2012), GLCM (Gómez et al. 2012), GLRLM (Prabusankarlal et al. 2015), ranklet transform-based co-occurrence matrix (RCM) (Yang et al. 2013) and phased-congruency based binary patterns (PCBP) (Cai et al. 2015). For each technique, the minimum rectangular ROI containing the lesion was used to compute the texture features. Details of these texture analysis techniques are summarized in Table 1. Furthermore, for the sake of objective comparison, the SVM hyperparameters were adjusted by k-fold cross-validation and a grid search procedure for each texture extraction method.

On the other hand, the proposed method was compared against the classification of an expert radiologist, who had classified 1731 of the 2050 images of the data set. The classification of the 1731 images (1190 benign and 541 malignant tumors) was performed in terms of BI-RADS (Breast Imaging-Reporting and Data System) categories. Only categories 2 (benign), 3 (probably benign), 4 (suspicious) and 5 (highly suggestive of malignancy) were used because they are oriented to images with lesions. Next, image categories were grouped according to their implied clinical conduct: categories 2 and 3 were grouped as *test negative* and 4 and 5 as *test positive*. Hence, a benign lesion classified as 2 or 3 was a true negative and a malignant lesion classified as 4 or 5 was a true positive, whereas a benign lesion classified as 4 or 5 was a false positive and a malignant lesion classified as 2 or 3 was a false negative (Calas et al. 2010). In this particular case, it is worth mentioning that the proposed method was evaluated with the same subset of images observed by the radiologist.

RESULTS

Classification performance of the AMI method

The AMI method depends on defining the number of resolutions n_r to compute $n_r \times 6$ texture features; thus, the effect of distinct values of n_r on classification performance was determined. Moreover, because the BUS images in the data set were acquired from distinct scanners, the classification performance using texture features extracted from ranklet images was compared with that using features extracted directly from the gray-level

Table 1. Features computed from texture analysis methods in the literature

Method	Texture feature
ACOR (2)*	Auto-correlation in the vertical and horizontal directions
ACOV (24)	Auto-covariance coefficients with displacements $\Delta m = \Delta n = 5$
FF (8)	Fractal dimension and fractal Brownian motion features with spatial distances $\Delta r_{min} = 1$ and $\Delta r_{max} = 8$
LEM (70)	Mean, standard deviation, skew, kurtosis and energy from 15 rotationally invariant texture energy measures
LBPV (10)	LBP variance histogram with $p = 8$ and $R = 1$
GLRLM (11)	Short-run emphasis; long-run emphasis; gray-level non-uniformity; run percentage; run length non-uniformity; low gray-level run emphasis; high gray-level run emphasis; short-run low gray-level emphasis; short-run high gray-level emphasis; long-run low gray-level emphasis; long-run high gray-level emphasis in the 0°, 45°, 90° and 135° directions; and quantization of 32 levels
GLCM (48)	Energy; entropy; correlation; contrast, homogeneity; Haralick's correlation; inverse difference normalized; cluster shade; cluster prominence; auto-correlation; dissimilarity; maximum probability features derived in the 0°, 45°, 90° and 135° directions with distances of 1, 2, 4 and 8 pixels; and quantization of 256 levels
RCM (72)	Same 12 GLCM features derived in the 0°, 45°, 90° and 135° directions; directions and 1-pixel distance, with ranklet resolutions $r = \{4,8\}$; and quantization of 256 levels
PCBP (80)	Local binary pattern variance measures with 1-pixel radius derived from phase congruency images with six scales and eight orientations

ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; FF = fractal features; GLCM = gray-level co-occurrence matrix; GLDM = gray-level difference matrix; GLRLM = gray-level run length matrix; LBPV = local binary pattern variance; LEM = Law's energy measures; PCBP = phased-congruency based binary patterns; RCM = ranklet transform-based co-occurrence matrix (RCM).

* The number of texture features is in parentheses.

image. Note that the latter option generates a 2-D texture feature vector. The goal is to demonstrate that intensity-invariant features are adequate to cope with images obtained using different ultrasound devices.

Figure 6 illustrates the classification performance of the AMI method regarding the AUC and the .632+ bootstrap estimator for different values of n_r, including the results of texture features extracted from the grayscale images. Note that extracting the texture features directly from the grayscale images has a negative impact on the performance of the classification.

On the other hand, the AUC values of the AMI method gradually increase as the number of resolutions increases, reaching stability after four resolutions. Significant differences between the results from using different ranklet resolutions have been found by statistical analysis. For the AUC values in Figure 6a, the Shapiro–Wilk test ($\alpha = 0.05$) indicated that all the groups are normally distributed; therefore, a one-way analysis of variance (ANOVA, $\alpha = 0.05$) was performed to compare the means of the groups. This test revealed that at least two groups are significantly different ($p < 0.0001$). Hence, Scheffe's post hoc test ($\alpha = 0.05$) was used to perform multiple comparisons of the means, where four and five resolutions were not statistically significantly different ($p = 0.435$), whereas the remaining pairwise comparisons presented statistical differences ($p < 0.0001$). These results indicate that using four resolutions is sufficient to obtain an adequate classification performance; thus, in the following experiments, the AMI method uses 12 ranklet images to compute 24 texture features.

Comparison with other texture analysis methods

Figure 7 illustrates the distributions of the measured performance indices for the texture extraction methods compared, where the AMI method clearly exhibits a tendency to outperform its counterparts. Table 2 summarizes the means and standard deviations from 1000 bootstrap replications. Note that the AMI method obtained the best classification results for all the indices considered. Furthermore, the RCM method obtained results similar to those of the AMI method for sensitivity. Nevertheless, a statistical analysis was conducted to find significant differences between the results obtained by the methods studied.

The Shapiro–Wilk test ($\alpha = 0.05$) revealed that all the groups come from normal distributions; thus, the ANOVA test ($\alpha = 0.05$) was applied to compare the means of the evaluated methods. For all the measured indices, the ANOVA test indicated that at least two compared groups were statistically different ($p < 0.0001$); therefore, Scheffe's post hoc test ($\alpha = 0.05$) was performed to determine whether the observed differences

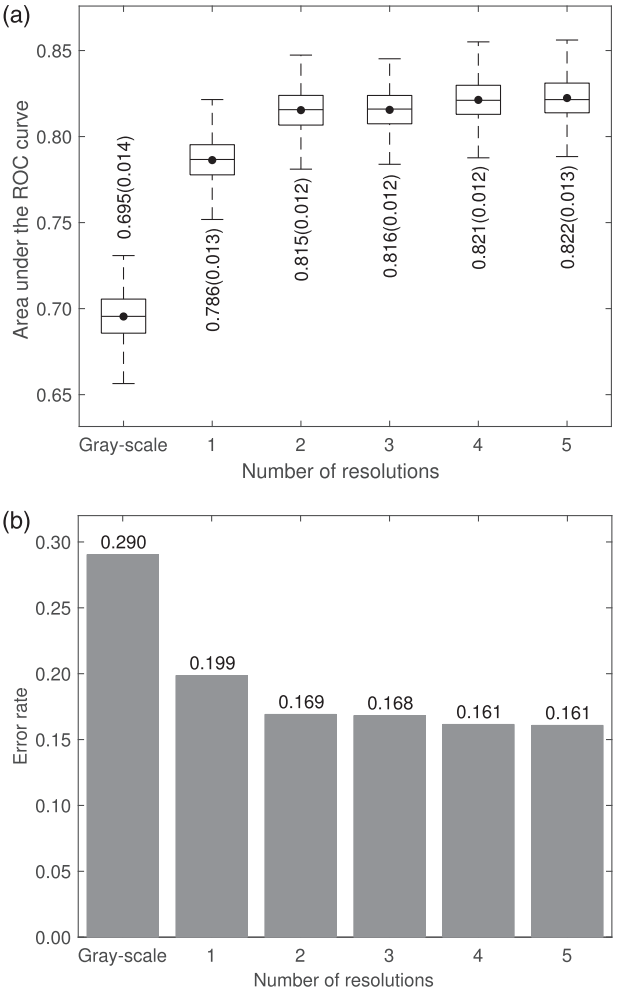

Fig. 6. Classification performance of the AMI method for different numbers of resolutions in terms of (a) AUC index, where "mean (standard deviation)" values are shown, and (b) error rate from the .632+ bootstrap estimator. In both graphs, 1000 independent bootstrap replications are used. *Black points* in (a) represent the mean values. AMI = auto-mutual information; AUC = area under the receiver operating characteristic curve.

were statistically significant. The results of these multiple comparisons of means are summarized in Table 3. The symbol '+' indicates that method A is statistically better than method B. The symbol '–' indicates that method A is statistically worse than method B. In both cases, all the p values were less than 0.001 for all the measured indices. The symbol '=' indicates that methods A and B are not statistically different. With respect to all the measured indices, note that the AMI method is statistically better than the other methods evaluated, except for the RCM method, for which the sensitivity is statistically similar ($p = 0.1006$). In this way, the AMI method is competitive with the RCM method in correctly identifying malignant tumors, although the correct identification of benign lesions is significantly improved; thus, the

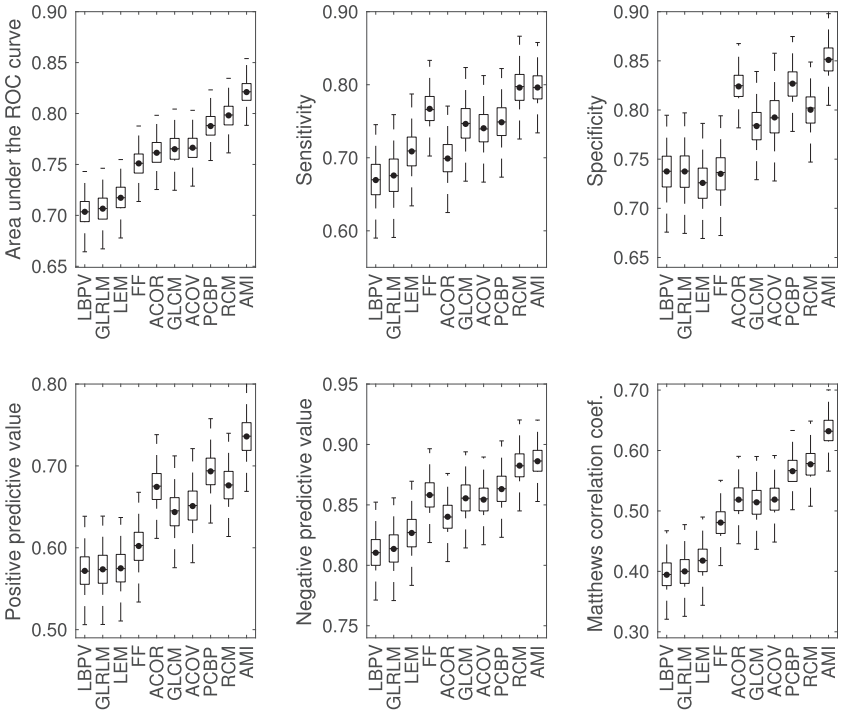

Fig. 7. Boxplots of the classification performance indices calculated from 1000 independent bootstrap samples. *Black points* represent the mean values. ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; FF = fractal features; GLCM = gray-level co-occurrence matrix; GLDM = gray-level difference matrix; GLRLM = gray-level run length matrix; LBPV = local binary pattern variance; LEM = Law's energy measures; PCBP = phased-congruency based binary patterns; RCM = ranklet transform-based co-occurrence matrix.

general classification quality, given by the AUC and MCC indices, is statistically significantly different from that of the other methods.

On the other hand, Figure 8 presents the .632+ bootstrap estimated errors reached by each texture extraction method, where the AMI method attained the lowest error rate, followed by the PCBP and RCM methods. Note that these three methods extract texture features invariant to monotonic intensity transformations of the BUS image. This characteristic is important for coping with distinct ultrasound scanners and configurations, which is presented in the BUS data set used in this study.

Additionally, to demonstrate that significant discriminant information can be found for the non-linear relationships of successive image displacements, the ACOR method was computed from ranklet images (with $n_r = 4$) in the vertical and horizontal directions to have the same conditions as the AMI method; both methods generated 24 texture features. The performance results of this comparison are summarized in Table 4. Note that the AMI

Table 2. Classification performance results of the texture analysis methods with the bootstrap method*

	Performance index					
Method	AUC	SEN	SPE	PPV	NPV	MCC
LBPV	0.70 (0.02)†	0.67 (0.03)	0.74 (0.02)	0.57 (0.02)	0.81 (0.02)	0.3 9(0.03)
GLRLM	0.71 (0.02)	0.68 (0.03)	0.74 (0.02)	0.57 (0.03)	0.81 (0.02)	0.40 (0.03)
LEM	0.72 (0.02)	0.71 (0.03)	0.73 (0.02)	0.57 (0.02)	0.83 (0.02)	0.42 (0.03)
FF	0.75 (0.01)	0.77 (0.03)	0.74 (0.02)	0.60 (0.02)	0.86 (0.01)	0.48 (0.03)
ACOR	0.76 (0.01)	0.70 (0.03)	0.82 (0.02)	0.67 (0.02)	0.84 (0.01)	0.52 (0.03)
GLCM	0.77 (0.02)	0.75 (0.03)	0.78 (0.02)	0.64 (0.03)	0.86 (0.02)	0.51 (0.03)
ACOV	0.77 (0.01)	0.74 (0.03)	0.79 (0.02)	0.65 (0.03)	0.85 (0.01)	0.52 (0.03)
PCBP	0.79 (0.01)	0.75 (0.03)	0.83 (0.02)	0.69 (0.02)	0.86 (0.01)	0.57 (0.03)
RCM	0.80 (0.01)	**0.80 (0.03)**	0.80 (0.02)	0.68 (0.02)	0.88 (0.01)	0.58 (0.03)
AMI	**0.82 (0.01)**	**0.80 (0.02)**	**0.85 (0.02)**	**0.74 (0.03)**	**0.89 (0.01)**	**0.63 (0.03)**

* The best classification results are in boldface.
† Mean (standard deviation).

Table 3. Pairwise comparisons performed with Scheffe's *post hoc* test

Method A	Method B	Performance index					
		AUC	SEN	SPE	PPV	NPV	MCC
AMI	RCM	+	=	+	+	+	+
AMI	PCBP	+	+	+	+	+	+
AMI	ACOV	+	+	+	+	+	+
AMI	GLCM	+	+	+	+	+	+
AMI	ACOR	+	+	+	+	+	+
AMI	FF	+	+	+	+	+	+
AMI	LEM	+	+	+	+	+	+
AMI	GLRLM	+	+	+	+	+	+
AMI	LBPV	+	+	+	+	+	+
RCM	PCBP	+	+	–	–	+	+
RCM	ACOV	+	+	+	+	+	+
RCM	GLCM	+	+	+	+	+	+
RCM	ACOR	+	+	–	=	+	+
RCM	FF	+	+	+	+	+	+
RCM	LEM	+	+	+	+	+	+
RCM	GLRLM	+	+	+	+	+	+
RCM	LBPV	+	+	+	+	+	+
PCBP	ACOV	+	+	+	+	+	+
PCBP	GLCM	+	=	+	+	+	+
PCBP	ACOR	+	+	=	+	+	+
PCBP	FF	+	–	+	+	+	+
PCBP	LEM	+	+	+	+	+	+
PCBP	GLRLM	+	+	+	+	+	+
PCBP	LBPV	+	+	+	+	+	+
ACOV	GLCM	=	–	+	+	=	=
ACOV	ACOR	+	+	–	–	+	=
ACOV	FF	+	–	+	+	–	+
ACOV	LEM	+	+	+	+	+	+
ACOV	GLRLM	+	+	+	+	+	+
ACOV	LBPV	+	+	+	+	+	+
GLCM	ACOR	+	+	–	–	+	=
GLCM	FF	+	–	+	+	=	+
GLCM	LEM	+	+	+	+	+	+
GLCM	GLRLM	+	+	+	+	+	+
GLCM	LBPV	+	+	+	+	+	+
ACOR	FF	+	–	+	+	–	+
ACOR	LEM	+	–	+	+	+	+
ACOR	GLRLM	+	+	+	+	+	+
ACOR	LBPV	+	+	+	+	+	+
FF	LEM	+	+	+	+	+	+
FF	GLRLM	+	+	=	+	+	+
FF	LBPV	+	+	=	+	+	+
LEM	GLRLM	+	+	–	=	+	+
LEM	LBPV	+	+	–	=	+	+
GLRLM	LBPV	+	+	=	=	+	+

ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; AUC = area under the receiver operating characteristic curve; FF = fractal features; GLCM = gray-level co-occurrence matrix; GLDM = gray-level difference matrix; GLRLM = gray-level run length matrix; LBPV = local binary pattern variance; LEM = Law's energy measures; MCC = Matthews correlation coefficient; NPV = negative predictive value; PCBP = phased-congruency based binary patterns; PPV = positive predictive value; RCM = ranklet transform-based co-occurrence matrix; SEN = sensitivity; SPE = specificity.

method outperformed the ACOR method for all indices evaluated, where Student's *t*-test ($\alpha = 0.05$) revealed that these two methods are statistically different ($p < 0.001$).

Lastly, Figure 9 illustrates the comparison of classification results between the AMI method and an expert radiologist. It is notable that the radiologist attained better sensitivity than the proposed method and, consequently, a better overall classification rate in terms of AUC and MCC, although the AMI method had higher specificity.

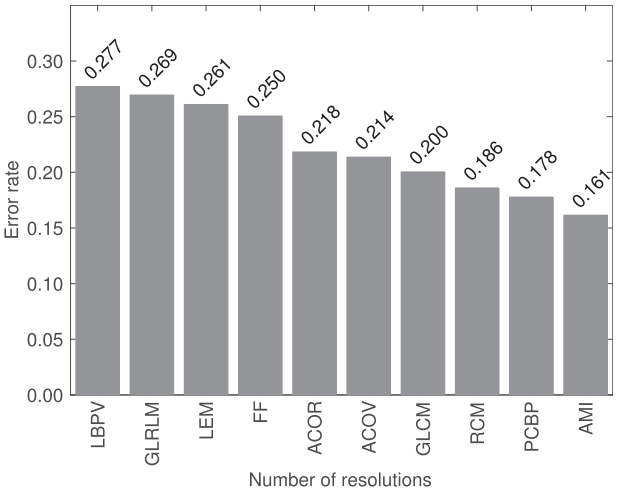

Fig. 8. The .632+ bootstrap error estimation of each texture extraction method. ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; FF = fractal features; GLCM = gray-level co-occurrence matrix; GLDM = gray-level difference matrix; GLRLM = gray-level run length matrix; LBPV = local binary pattern variance; LEM = Law's energy measures; PCBP = phased-congruency based binary patterns; RCM = ranklet transform-based co-occurrence matrix.

Computation time

When a feature extraction method is designed, both the transformation invariance and the discrimination power are the main goals, although obtaining results with a reasonable computation time should also be considered.

The time needed by Algorithm 1 varies with the size of the ROI; that is, the larger the area of the ROI, the longer is the computation time. Figure 10 illustrates the running time as a function of the area of the ROI

Table 4. Classification performance results for AMI and ACOR methods based on ranklet transform with the bootstrap method

Index	AMI	ACOR
AUC	0.82 (0.01)*	0.79 (0.02)
SEN	0.79 (0.01)	0.76 (0.02)
SPE	0.84 (0.01)	0.82 (0.02)
PPV	0.73 (0.02)	0.69 (0.02)
NPV	0.88 (0.01)	0.87 (0.01)
MCC	0.63 (0.02)	0.57 (0.02)

ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; AUC = area under the receiver operating characteristic curve; MCC = the Matthews correlation coefficient; NPV = negative predictive value; PPV = positive predictive value; SEN = sensitivity; SPE = specificity.

* Mean (standard deviation).

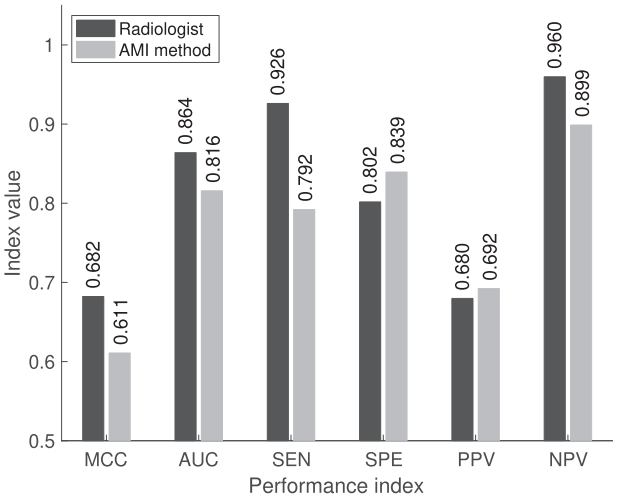

Fig. 9. Classification performance indices of an expert radiologist and the AMI method. AMI = auto-mutual information; AUC = area under the receiver operating characteristic curve; MCC = the Matthews correlation coefficient; NPV = negative predictive value; PPV = positive predictive value; SEN = sensitivity; SPE = specificity.

(*i.e.,* the number of pixels), where a linear tendency is observed. The mean size of the ROIs in the BUS data set was 95 × 147 pixels, and the mean running time was approximately 2 s. In addition, note that computing 12 ranklet images required less time than computing 24 AMI-based texture features, with mean times of 0.22 and 1.74 s, respectively. The largest ROI in the data set was 269 × 455 pixels and was processed in 12.5−1.4 s to compute the 12 ranklet images and 11.1 s to calculate the 24 AMI-based texture features.

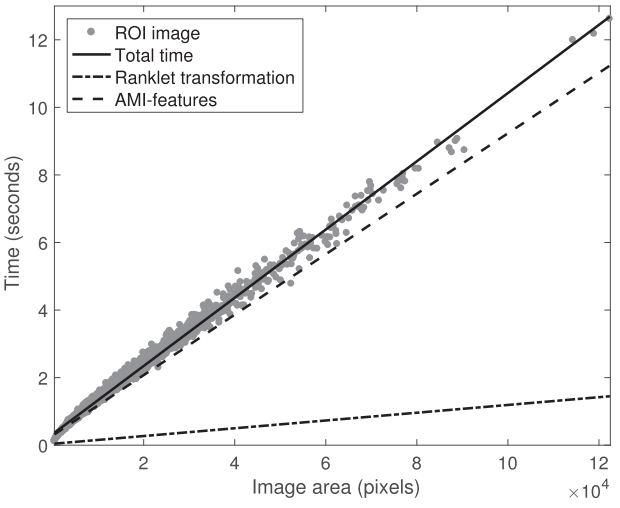

Fig. 10. Running time of the AMI method as a function of the area of the ROI. AMI = auto-mutual information; ROI = region of interest

Table 5. Computation time statistics of the texture analysis methods

	Computational time (s)			
Method	Mean	SD	Max	Min
ACOV	0.002	0.001	0.010	0.001
LBPV	0.006	0.005	0.040	0.001
FF	0.014	0.023	1.058	0.007
ACOR	0.016	0.015	0.154	0.000
GLRLM	0.025	0.017	0.301	0.004
LEM	0.035	0.019	0.171	0.014
PCBP	0.047	0.039	0.323	0.010
GLCM	1.100	0.011	1.166	1.070
RCM	1.605	0.016	1.709	1.568
AMI	1.978	1.522	12.617	0.114

ACOR = auto-correlation coefficient; ACOV = auto-covariance coefficient; AMI = auto-mutual information; FF = fractal features; GLCM = gray-level co-occurrence matrix; GLDM = gray-level difference matrix; GLRLM = gray-level run length matrix; LBPV = local binary pattern variance; LEM = Law's energy measures; PCBP = phased-congruency based binary patterns; RCM = ranklet transform-based co-occurrence matrix; SD = standard deviation.

In addition, for comparison, Table 5 summarizes the computation time reached by the texture extraction methods described in Table 1. Note that the AMI method required more computation time than its counterparts, although it is still reasonable for obtaining adequate classification performance.

DISCUSSION

In the literature, several approaches use the auto-correlation (or auto-covariance) measure to characterize the echogenicity of breast tumors, where a linear relationship between successive image displacements is assumed. However, discriminant information can be found from the non-linear relationships of the image; thus, auto-correlation suffers limitations concerning the characterization of non-linear dependencies of image textures. Hence, AMI was used to measure the non-linear relationship between successive BUS image displacements. The experimental results clearly indicate that AMI significantly outperformed the auto-correlation measure under the same experimental conditions, which suggests that taking into account the non-linear relationship of BUS textures provides significant discriminative information for distinguishing benign from malignant lesions.

On the other hand, to cope with the variability of the intensity of BUS images acquired from distinct ultrasound scanners, the results confirm that using texture features robust to intensity transformations is useful. In this way, the lesion classification performance was improved as compared with texture features directly computed from gray-level images. Hence, using an intensity-invariant representation before feature extraction is suitable for reducing the error rate, which has

been previously reported by several studies (Cai et al. 2015; Lo et al. 2015; Masotti and Campanini 2008; Yang et al. 2013).

Furthermore, the classification performance of the AMI method, as well as the ACOR approach, notably improved when texture features were computed from ranklet images instead of grayscale images. In the case of the AMI method, the classification performance in terms of the AUC index increased from 0.69−0.82, whereas the performance of the ACOR method increased from 0.71−0.79. In addition, the PCBP and RCM methods outperformed the texture methods that extracted features directly from the gray-level images. Therefore, robustness to monotonic intensity transformations reduces the misclassification of breast lesions.

A common drawback of texture description is the "curse of dimensionality." For instance, concerning the GLCM-based methods, the number of features is given by the number of distances × the number of orientations × the number of second-order statistics (*e.g.*, correlation, entropy, homogeneity). Generally, to reduce the dimensionality, the texture statistics of the same distance are averaged over all the orientations, although the number of features could remain large. Additionally, if some intensity-invariant transformation is performed, the number of features increases as a function of the number of resolutions (or scales) and orientations. This behavior is illustrated in the RCM and PCBP methods, which calculate 72 and 80 texture features, respectively. A feature selection technique can be used to reduce the dimensionality, although this procedure would introduce a higher computational cost at the training stage. In the case of the AMI method, the number of features is remarkably reduced right from the start: 24 texture features are calculated considering four resolutions of the ranklet transform. Moreover, with this reduced feature set, the AMI method surpassed both the RCM and PCBP methods in terms of classification performance.

Another typical issue found in medical data is the class imbalance; in this sense, benign lesions of the breast are far more frequent than malignant ones (Guray and Sahin 2006). In this study, the class imbalance was addressed by a class-weighting scheme used during the SVM training. This scheme imposed an additional cost on the model, in which errors in the minority class are more penalized than errors in the majority class. However, it is recommended that the number of malignant cases for obtaining a classification model be increased with greater diversity of cases.

The experimental results indicate that the AMI-based texture features present good discriminant capability to distinguish between benign and malignant breast tumors, such that all the measured classification performance indices were improved in comparison to nine texture extraction methods proposed in the literature. Moreover, the proposed approach has the potential to avoid recommending unnecessary biopsies in benign lesions. This observation can be inferred from the experimental comparison between the AMI method and the classification performed by an expert radiologist, in which the former attained a specificity of 0.84, whereas the latter attained 0.80. On the other hand, the sensitivity of the radiologist was higher (0.93) than that of the proposed approach (0.80), which may be due to the radiologist's concern over erroneously classifying a truly malignant tumor as benign, that is, BI-RADS categories 2 and 3.

Aiming to increase the classification rate of a CAD system, the combination of AMI-based texture features with morphological features and other clinical data (*e.g.,* patient age) is a feasible approach. In addition, applying a feature selection technique is recommended for obtaining the most significant features to distinguish between benign and malignant tumors (Alvarenga et al. 2012; Daoud et al. 2016; Flores et al. 2015).

CONCLUSIONS

We have described a novel texture description method based on the AMI measure for classifying breast lesions from ultrasound images. In this method, 24 AMI-based texture features that are robust to intensity transformations are calculated. This method is adequate for use with BUS images acquired with different ultrasound scanners. In addition, nine methods for texture description proposed in the literature were outperformed in terms of lesion classification. The AMI method has high performance because the discriminative information found in the non-linear relationships of textures from intensity-invariant images is extracted. In addition, the computation time of the proposed method is reasonable, so that it is suitable for implementation in breast lesion classification systems.

Acknowledgments—This research was supported by a Fondo SEP-Cinvestav 2018 grant (No. 145) and a National Council of Science and Technology (CONACyT, Mexico) research scholar grant (No. (362179).

REFERENCES

Al-Kadi OS, Chung DY, Carlisle RC, Coussios CC, Noble JA. Quantification of ultrasonic texture intra-heterogeneity via volumetric stochastic modeling for tissue characterization. Med Image Anal 2015;21:59–71.

Alvarenga AV, Pereira WCA, Infantosi AFC, Azevedo CM. Complexity curve and grey level co-occurrence matrix in the texture evaluation of breast tumor on ultrasound images. Med Phys 2007;34:379–387.

Alvarenga AV, Infantosi AFC, Pereira WCA, Azevedo CM. Assessing the combined performance of texture and morphological parameters in distinguishing breast tumors in ultrasound images. Med Phys 2012;39:7350–7358.

American Cancer Society. Breast cancer facts & figures 2015−2016. Atlanta, GA: Author; 2015.

Azzopardi JG, Chepick OF, Hartmann WH, Jafarey NA, Llombart-Bosch A, Ozzello L, Rilke F, Sasano N, Sobin LH, Sommers SC, Stalsberg H, Sugar J, Williams AO. The World Health Organization histological typing of breast tumors—Second edition. Am J Clin Pathol 1982;78:806–816.

Baldi P, Brunak S, Chauvin Y, Andersen CAF, Nielsen H. Assessing the accuracy of prediction algorithms for classification: an overview. Bioinformatics 2000;16:412–424.

Ben-Hur A, Weston J. A user's guide to support vector machines. New York: Humana Press; 2010. p. 223–239.

Cai L, Wang X, Wang Y, Guo Y, Yu J, Wang Y. Robust phase-based texture descriptor for classification of breast ultrasound images. BioMed Eng OnLine 2015;14:26.

Calas MJ, Almeida RM, Gutfilen B, Pereira WCA. Intraobserver interpretation of breast ultrasonography following the bi-rads classification. Eur J Radiol 2010;74:525–528.

Chang RF, Wu WJ, Moon WK, Chen DR. Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis. Ultrasound Med Biol 2003;29:679–686.

Chang CC, Lin CJ. LIBSVM: A library for support vector machines. ACM Trans Intell Syst Technol 2011;2(27):1–27.

Chatterjee S, Chatterjee S. Estimation of misclassification probabilities by bootstrap methods. Commun Stat Simul Comput 1983;12:645–656.

Chen DR, Chang RF, Kuo WJ, Chen MC, Huang YL. Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks. Ultrasound Med Biol 2002;28:1301–1310.

Chen DR, Chang RF, Chen CJ, Ho MF, Kuo SJ, Chen ST, Hung SJ, Moon WK. Classification of breast ultrasound images using fractal feature. Clin Imaging 2005;29:235–245.

Cheng HD, Shan J, Ju W, Guo Y, Zhang L. Automated breast cancer detection and classification using ultrasound images: A survey. Pattern Recog 2010;43:299–317.

Coronel C, Garn H, Waser M, Deistler M, Benke T, Dal-Bianco P, Ransmayr G, Seiler S, Grossegger D, Schmidt R. Quantitative EEG markers of entropy and auto mutual information in relation to MMSE scores of probable Alzheimer's disease patients. Entropy 2017;19:130.

Cover TM, Thomas JA. Elements of information theory. New York: Wiley; 1991.

Cristerna AR, Flores WG, de Albuquerque Pereira WC. A computer-aided diagnosis system for breast ultrasound based on weighted BI-RADS classes. Computer Methods Programs Biomed 2018;153:33–40.

Daoud MI, Bdair TM, Al-Najar M, Alazrai R. A fusion-based approach for breast ultrasound image classification using multiple-ROI texture and morphological analyses. Comput Math Methods Med 2016; 6740956.

D'Orsi C, Sickles E, Mendelson E, Morris E. Breast Imaging Reporting and Data System: ACR BI-RADS breast imaging atlas. 5th edition Reston, VA: American College of Radiology; 2013.

Efron B, Tibshirani R. Improvements on cross-validation: The .632+ bootstrap method. J Am Stat Assoc 1997;92:548–560.

Farmer KM. Application of chaos and fractals to computer vision. Flint, MI: Bentham Science; 2015.

Ferlay J, Soerjomataram I, Dikshit R, Eser S, Mathers C, Rebelo M, Parkin D, Forman D, Bray F. Cancer incidence and mortality worldwide: Sources, methods and major patterns in GLOBOCAN 2012. Int J Cancer 2015;136:E359–E386.

Flores WG, de Albuquerque Pereira WC, Infantosi AFC. Improving classification performance of breast lesions on ultrasonography. Pattern Recog 2015;48:1125–1136.

Foley D. Considerations of sample and feature size. IEEE Trans Inf Theory 1972;18:618–626.

Gómez W, Pereira W, Infantosi A. Analysis of co-occurrence texture statistics as a function of gray-level quantization for classifying breast ultrasound. IEEE Trans Med Imaging 2012;31:1889–1899.

Guray M, Sahin AA. Benign breast diseases: Classification, diagnosis, and management. Oncologist 2006;115:435–449.

Liu H, Tan T, Jvan Zelst J, Mann R, Karssemeijer N, Platel B. Incorporating texture features in a computer-aided breast lesion diagnosis system for automated three-dimensional breast ultrasound. J Med Imaging (Bellingham) 2014;1 024501.

Hastie T, Tibshirani R, Friedman J. The elements of statistical learning: Data mining, inference, and prediction. New York: Springer-Verlag; 2009.

Horsch K, Giger ML, Venta LA, Vyborny CJ. Computerized diagnosis of breast lesions on ultrasound. Med Phys 2002;29:157–164.

Huang YL, Lin SH, Chen DR. Computer-aided diagnosis applied to 3-D US of solid breast nodules by using principal component analysis and image retrieval. 2005 IEEE Engineering in Medicine and Biology 27th Annual Conference. Piscataway, NJ: IEEE; 2005. p. 1802–1805.

Huang YL, Wang KL, Chen DR. Diagnosis of breast tumors with ultrasonic texture analysis using support vector machines. Neural Comput Appl 2006;15:164–169.

Joo S, Yang YS, Moon WK, Kim HC. Computer-aided diagnosis of solid breast nodules: Use of an artificial neural network based on multiple sonographic features. IEEE Trans Med Imaging 2004;23:1292–1300.

Kelly KM, Dean J, Comulada WS, Lee SJ. Breast cancer detection using automated whole breast ultrasound and mammography in radiographically dense breasts. Eur Radiol 2010;20:734–742.

Kriti, Virmani J, Dey N, Kumar V. PCA-PNN and PCA-SVM based CAD systems for breast density classification. Berlin/New York: Springer; 2016. p. 159–180.

Lo CM, Moon WK, Huang CS, Chen JH, Yang MC, Chang RF. Intensity-invariant texture analysis for classification of BI-RADS category 3 breast masses. Ultrasound Med Biol 2015;41:2039–2048.

Masotti M, Campanini R. Texture classification using invariant ranklet features. Pattern Recog Lett 2008;9:1980–1986.

Matsumoto MMS, Sehgal CM, Udupa JK. Local binary pattern texture-based classification of solid masses in ultrasound breast images. Proc SPIE 2012;8320.

Melia U, Vallverdú M, Jospin M, Jensen EW, Valencia JF, Clariá F, Gambus PL, Caminal P. Auto-mutual information function for predicting pain responses in EEG signals during sedation. Berlin/New York: Springer; 2014. p. 623–626.

Melia U, Guaita M, Vallverdú M, Embid C, Vilaseca I, Salamero M, Santamaria J. Mutual information measures applied to EEG signals for sleepiness characterization. Med Eng Phys 2015;37:297–308.

Na SH, Jin SH, Kim SY, Ham BJ. EEG in schizophrenic patients: Mutual information analysis. Clin Neurophysiol 2002;113:1954–1960.

Prabusankarlal KM, Thirumoorthy P, Manavalan R. Assessment of combined textural and morphological features for diagnosis of breast masses in ultrasound. Human-centric Comput Inf Sci 2015;5:12.

Priddy KL, Keller PE. Artificial neural networks: An introduction. Washington, DC: SPIE; 2005.

Shan J, Alam SK, Garra B, Zhang Y, Ahmed T. Computer-aided diagnosis for breast ultrasound using computerized BI-RADS features and machine learning methods. Ultrasound in Med Biol 2016;42:98–988.

Shankar PM, Dumane VA, George T, Piccoli CW, Reid JM, Forsberg F, Goldberg BB. Classification of breast masses in ultrasonic B scans using Nakagami and K distributions. Phys Med Biol 2003;48:2229–2240.

Shen WC, Chang RF, Moon WK, Chou YH, Huang CS. Breast ultrasound computer-aided diagnosis using BI-RADS features. Acad Radiol 2007;14:928–939.

Singh BK, Verma K, Thoke A, Suri JS. Risk stratification of 2D ultrasound-based breast lesions using hybrid feature selection in machine learning paradigm. Measurement 2017;105:146–157.

Sokolova M, Lapalme G. A systematic analysis of performance measures for classification tasks. Inf Process Manage 2009;45:427–437.

Strehl A, Ghosh J. Cluster ensembles—A knowledge reuse framework for combining multiple partitions. J Mach Learn Res 2002;3:583–617.

Timmers JMH, van Doorne-Nagtegaal HJ, Verbeek ALM, den Heeten GJ, Broeders MJM. A dedicated BI-RADS training programme: Effect on the inter-observer variation among screening radiologists. Eur J Radiol 2012;81:2184–2188.

Yang MC, Moon WK, Wang YCF, Bae MS, Huang CS, Chen JH, Chang RF. Robust texture analysis using multi-resolution gray-scale invariant features for breast sonographic tumor diagnosis. IEEE Trans Med Imaging 2013;32:2262–2273.