

Robust Texture Analysis Using Multi-resolution Gray-scale Invariant Features for Breast Sonographic Tumor Diagnosis

Min-Chun Yang, Woo Kyung Moon^{*}, Yu-Chiang Frank Wang, Min Sun Bae, Chiun-Sheng Huang, Jeon-Hor Chen
^{*}and Ruey-Feng Chang^{*}, Senior Member, IEEE

Abstract—Computer-aided diagnosis (CAD) systems in gray-scale breast ultrasound images have the potential to reduce unnecessary biopsy of breast masses. The purpose of our study is to develop a robust CAD system based on the texture analysis. First, gray-scale *invariant* features are extracted from ultrasound images via multi-resolution ranklet transform. Thus, one can apply linear support vector machines (SVMs) on the resulting gray-level co-occurrence matrix (GLCM)-based texture features for discriminating the benign and malignant masses. To verify the effectiveness and robustness of the proposed texture analysis, breast ultrasound images obtained from *three* different platforms are evaluated based on cross-platform training/testing and leave-one-out cross-validation (LOO-CV) schemes. We compare our proposed features with those extracted by wavelet transform in terms of receiver operating characteristic (ROC) analysis. The AUC values derived from the area under the curve for the *three* databases via ranklet transform are 0.918 (95% confidence interval [CI], 0.848 to 0.961), 0.943 (95% CI, 0.906 to 0.968) and 0.934 (95% CI, 0.883 to 0.961), respectively, while those via

wavelet transform are 0.847 (95% CI, 0.762 to 0.910), 0.922 (95% CI, 0.878 to 0.958) and 0.867 (95% CI, 0.798 to 0.914), respectively. Experiments with cross-platform training/testing scheme between each database reveal that the diagnostic performance of our texture analysis using ranklet transform is less sensitive to the sonographic ultrasound platforms. Also, we adopt several co-occurrence statistics in terms of quantization levels and orientations (i.e., descriptor settings) for computing the co-occurrence matrices with 0.632+ bootstrap estimators to verify the use of the proposed texture analysis. These experiments suggest that the texture analysis using multi-resolution gray-scale *invariant* features via ranklet transform is useful for designing a robust CAD system.

Index Terms—Breast sonography, computer-aided tumor diagnosis, texture analysis, gray-scale invariant features, multi-resolution approach, 0.632+ bootstrap estimators.

I. INTRODUCTION

BREAST ultrasound (BUS) imaging is a useful tool for early detection of breast cancer as well as diagnosing the breast lesions [1, 2]. Stavros *et al.* [3] have reported the sensitivity for tumor diagnosis can reach 98.4% by interpreting BUS. Berg *et al.* [4] demonstrated supplemental screening ultrasound can depict small, node-negative breast tumors not obviously seen on mammography. In addition, the studies in [5, 6] have reported BUS screening can yield an incremental detection rate from 2.8 to 4.6 cancers per 1,000 women with dense breasts and negative mammograms.

Nonetheless, interpretation of BUS images for robust diagnosis requires experienced radiologists, and the diagnosis is often subjective. Hence, recent studies advocate to investigate and develop computer-aided diagnosis (CAD) [7-9] systems for addressing this issue. Using extracted textural or morphological features [7, 10] from the located tumor regions (i.e., regions of interest (ROIs)); the CAD system can automatically interpret and classify the breast tumors into malignant and benign ones. Several studies have shown the potential of BUS CAD to reduce the unnecessary biopsy [11, 12].

Texture patterns in BUS have been deemed a useful characteristic for distinguishing benign and malignant tumors [13, 14]. Recently, several studies aim to extract useful texture features for tumor diagnosis based on the gray-level co-occurrence matrix (GLCM) [15, 16]. Furthermore, Chen *et al.* [17] dedicate to investigate useful texture features extracted from wavelet transformed BUS images. Tsiparas *et al.* [18]

Manuscript received February 9, 2013; revised June 15, 2013; accepted July 18, 2013. The authors thank the National Science Council (NSC 101-2221-E-002-068-MY3), Ministry of Economic Affairs (102-EC-17-A-19-S1-164), and Ministry of Education (AE-00-00-06) of the Republic of China for the financial support. This work was also supported by Industrial Strategic technology development program (10042581) funded by the Ministry of Knowledge Economy (MKE, Korea) and by the Basic Science Research Program through the National Research Foundation of Korea (NRF) funded by the Ministry of Education, Science and Technology (2012R1A1A2009414).

Copyright (c) 2013 IEEE. Personal use of this material is permitted. However, permission to use this material for any other purposes must be obtained from the IEEE by sending a request to pubs-permissions@ieee.org.

M.-C. Yang is with Department of Computer Science and Information Engineering National Taiwan University, Taipei 10617, Taiwan.

W. K. Moon is with Department of Radiology, Seoul National University Hospital, Seoul 110-744, Korea (e-mail: moonwk@snu.ac.kr)

Y.-C. F. Wang is with Research Center for Information Technology Innovation, Academia Sinica, Taipei 105, Taiwan.

M. S. Bae is with Department of Radiology, Seoul National University Hospital, Seoul 110-744, Korea.

C.-S. Huang is with Department of Surgery, National Taiwan University Hospital and National Taiwan University College of Medicine, Taipei 106, Taiwan.

J.-H. Chen is with Tu & Yuen Center for Functional Onco-Imaging, Department of Radiological Sciences, University of California, Irvine, California 92697, United States and Department of Radiology, E-Da Hospital and I-Shou University, Kaohsiung 82445, Taiwan.

R.-F. Chang is with Department of Computer Science and Information Engineering and the Graduate Institute of Biomedical Electronics and Bioinformatics National Taiwan University, Taipei 10617, Taiwan (e-mail: rfchang@csie.ntu.edu.tw).

advance to extract the textures using wavelet-based transform with multi-resolution approach and SVM classifier [13, 19] is adopted for discriminating the atherosclerotic tissue from B-mode ultrasound. Gómes *et al.* [20] devote to analyze the gray-level co-occurrence statistics with six quantization levels and select the effective texture descriptors (e.g. quantization level, orientation and distance) for BUS tumor diagnosis.

Nevertheless, a challenge task arises for conventional texture analyses [15-17, 20, 21] while the adjustable parameters of the ultrasonic device [22] will introduce incoherent texture feature extraction from original BUS images or wavelet transformed BUS images and lead to non-robust diagnostic capability for tumor diagnosis. In this study, we refer the non-robust diagnostic capability to *variant* texture analysis. Since the *invariant* texture analysis is highly demanded for developing clinical applications, more recent studies are directed to extract the gray-scale *invariant* texture features for pattern recognition. The local binary patterns (LBPs) [23] are firstly investigated to deal with gray-scale and rotation *invariant* texture classification. Masumoto *et al.* [24] aim to extract the textures based on local binary patterns for classifying the solid masses in BUS images. As firstly proposed to use the *ranklets* (i.e., ranklet transformed images) for robust face recognition [25], several researches pay more attention to ranklet transform which is adopted for robust texture classification and mass classification in the mammograms [26, 27].

The ranklet transform is an image processing technique characterized by a multi-resolution and orientation-selective approach similar to that of the wavelet transform. Yet, differently from the latter, it deals with the rank of the pixels rather than their gray-scale intensity values. Herein, we develop a robust CAD system based on the gray-scale *invariant* features via ranklet transform. GLCM textures are extracted from multi-resolution ranklet transformed BUS images, which allow standard linear support vector machines (SVMs) [13, 19] for performing BUS tumor diagnosis. While performing on *three* different BUS platforms, we carry out the experiments concerning leave-one-out (LOO-CV) scheme and cross-platform training/testing schemes for BUS texture analyses to verify the robustness and stability of the proposed method. The experiments reveal the texture analyses using ranklet transform are less sensitive to different ultrasonic devices and properly adopted for designing a robust CAD system for tumor diagnosis.

II. MATERIALS

In this study, BUS images of three databases are used for evaluation. Database A includes 116 subjects (78 benign and 38 malignant cases) which were obtained (from August 2003 to January 2004) with Acuson Sequoia (Acuson Siemens, California, USA) equipment, using an 8-15 MHz linear-array 52-mm ultrasound probe. The ages of the subjects were from 17 to 82 years (mean age 44.48 ± 11.22 years). Database B includes 193 subjects (133 benign and 60 malignant cases) which were obtained (from April 2003 to February 2004) with GE LOGIQ 7 (GE Medical Systems, Milwaukee, Wisc., USA) equipment, using a 4.5-14 MHz linear-array 40-mm ultrasound probe. The

ages of the subjects were from 21 to 71 years (mean age 46.77 ± 9.07 years). Database C includes 161 subjects (104 benign and 57 malignant cases) which were obtained (from August 2004 to March 2005) with Voluson 730 expert (GE Medical systems, Kretz Ultrasound, Zipf, Austria) equipment, using a 4.0-10.5 MHz linear-array 38-mm ultrasound probe. The ages of the subjects were from 20 to 85 years (mean age 42.22 ± 8.34 years). The image pixel resolution of the three databases are 0.12 mm/pixel for database A, 0.10 mm/pixel for database B, and 0.11 mm/pixel for database C, respectively. Note that all the BUS images are captured with whole ultrasonic screen (640×480 pixels) and stored in 8-bit pixel depth (i.e., 256 gray scales) with DICOM format. Detailed tumor characteristics of the three databases are listed in Table I. All breast tumors were histopathologically proven by means of BUS-guided core needle biopsy or fine-needle aspiration cytology (FNAC). The institutional review board approved this retrospective study and informed consent was obtained from each patient prior to performing the biopsy.

For effectively retrieving the representative image frames to characterize the breast tumors, the images with largest diameters of the tumors are selected and captured by the radiologists while operating the whole breast examination. On the other hand, in order to extract the tumor regions (i.e., ROIs), the tumor boundaries were manually demarcated by two experienced radiologists who have clinical experience of 6 and 13 years in BUS screening respectively. Note that the sizes of the ROIs vary from 724 (65×28) to 55,186 (340×207) pixels for Database A. For Database B, the maximum and minimum ROI sizes are 788 (34×25) and 43,400 (325×197) pixels. As for Database C, the sizes are between 543 (36×18) and 27,781 (213×193) pixels.

III. METHODS

This study focuses on robust texture analysis for tumor diagnosis using multi-resolution gray-scale *invariant* features via ranklet transform [26, 28, 29]. Automatic texture analysis involves three major procedures [26, 29] as drawn in Fig. 1. First, we decompose each BUS image \mathbf{I} from the test database (as shown in Fig. 1 (a)) into *ranklets* (i.e., a series of images represented in the ranklet domain) in terms of multi-resolution and orientation-selective properties of the ranklet transform [26] as depicted in Fig. 1 (b). Afterwards, the gray-scale *invariant* texture features based on GLCM [21, 30, 31] can be calculated from each transformed (ranklet) image with the regions of interest (ROIs) as depicted in Fig. 1 (c). In order to retrieve multi-resolution features for texture analysis [23], the texture features derived from different resolutions (scales) and orientations of original BUS image can be combined into one compact texture features as the image texture representation for each breast tumor. As a result, the SVM [19, 41] classifier is adopted to distinguish a benign tumor from a malignant one as drawn in Fig. 1 (d). Several experiments are conducted using *three* different sonographic BUS platforms (i.e., Databases A, B and C) to evaluate the robustness and effectiveness of the proposed texture analysis approach for tumor diagnosis via ranklet transform.

A. Robust Gray-Scale Invariant Ranklet Transform

The ranklet transform involves non-parametric analysis, orientation-selective and multi-resolution properties as analog to Haar wavelets [32], which can be used as a rank descriptor of the pixels within a local region [26]. Since the ranklet transform considers the relative rank of the pixels [27] instead of corresponding gray values, it can be defined as an *invariant* operator to any monotonic changes of any arbitrary observed gray pixels. For example, given an arbitrary matrix S and an additive matrix P (all entries have the same value), the rank descriptive matrix RS with ranklet transform π of S , $S+P$ or $c \cdot S$ (c is a positive scalar) can be described as follows:

$$S = \begin{pmatrix} 12 & 28 & 25 & 78 \\ 72 & 55 & 44 & 13 \\ 46 & 11 & 22 & 89 \\ 14 & 58 & 33 & 10 \end{pmatrix}, P = \begin{pmatrix} 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \\ 3 & 3 & 3 & 3 \end{pmatrix}, RS = \pi(S) = \pi(S \pm P) = \pi(c \cdot S) = \begin{pmatrix} 3 & 8 & 7 & 15 \\ 14 & 12 & 10 & 4 \\ 11 & 2 & 6 & 16 \\ 5 & 13 & 9 & 1 \end{pmatrix} \quad (1)$$

In (1), the rank value of each entry of the matrix RS is started with 1 and ended up with the number of the matrix size. Note that the rank values in RS remain unchanged, since the transform operation (e.g., c or P) converts entries of S monotonically. Thus, this study aims to extract the gray-scale *invariant* features via ranklet transform for tumor diagnosis, which would be less sensitive to the database collection with operator-dependent issues. We refer *ranklets* to a series of transformed images from an input BUS image based on the multi-resolution and orientation-selective property of the ranklet transform and can be derived using ranklet decomposition [26, 28] as drawn in Fig. 1 (b).

More specifically, for a clipping square crop, the multi-resolution property of the ranklet decomposition can be described by giving an arbitrary point p and even resolutions r of the clipping crop as drawn in Fig. 2 (a). In the clipping crop, we separate the gray pixels into two clusters of pixels with the same size (i.e., subset A and B). The orientations of the two separated subsets A and B can be defined as horizontal (H), vertical (V) and diagonal (D) directions as drawn in Fig. 2 (b). To determine the ranklet coefficient with an arbitrary point p in ranklet image with specified resolution r and orientation t , we firstly rank the gray values of the pixels within the observed crop (i.e., all observed gray points in $A \cup B$) via ranklet transform to obtain the rank descriptive matrix ($R_{AB} = \pi(A \cup B)$). Meanwhile, we perform the sorting operation for subsets A and B in terms of the rank numbers of R_{AB} to obtain the sorted rank descriptive vectors SA and SB , respectively (i.e., the rank numbers existed in sub-region A or B are rearranged in ascending order to form the vectors, SA and SB). Hence, the ranklet coefficient $RC_r^t(p)$ can be defined as follows:

$$RC_r^t(p) = \left| \left(\sum_{i=1}^{N/2} SA_r^t(p, i) - SB_r^t(p, i) \right) / \left(\frac{N}{4} \right)^2 \right|, \quad (2)$$

$$t \in \{H, V, D\}, r = \{2, 4, 6, \dots\}$$

where $SA_r^t(p, i)$ and $SB_r^t(p, i)$ are the i -th sorted rank number in the respective subset while N ($N = r^2$) is the total pixels of the clipping crop. We note that the dynamic range of the ranklet

coefficient is $[0, +1]$ and represents the contrast strength between two sub-region $A_r^t(p)$ and $B_r^t(p)$. In other words, if the ranklet coefficient $RC_r^t(p)$ is approaching $+1$ indicating the gray pixels with higher gray-scales in original BUS image are centralized at one subset (i.e., $A_r^t(p)$ or $B_r^t(p)$). That is, the ranklet coefficient can be regarded as a directional gray-scale *invariant* gradient descriptor of a local region. The higher ranklet coefficients (gradients) might represent edge or corner structures and the lower ones might represent some specific texture patterns instead. Herein, we derive the ranklet coefficient for each target point within the ROI of a BUS image. Thus, the ranklet image is composed of the ranklet coefficients derived for this ROI, and this image will be applied for texture analysis.

B. Multi-Resolution Gray-scale Invariant GLCM Texture Extraction

Conventional texture analyses measures the local texture information as the histograms to characterize the histological textures for B-mode BUS diagnosis [33]. Nevertheless, regarding the related position between the pixels within a local region might offer more geometry information for feature representation [23]. GLCM textures [21, 30, 31] are proposed to calculate the texture feature depends on the spatial dependence of the gray values, which have shown the success in the image classification [30] and tumor diagnosis [15, 20].

The GLCM characterizes all the joint frequencies between the quantized level p and q by given a distance d and rotation angle θ between two arbitrary pixels (k, l) and (m, n) within the region of interest R . Herein, we can define the multi-resolution ranklet GLCMs as [20, 30]

$$RCM_{r,\theta}^t(p, q) = \# \left\{ (k, l), (m, n) \mid \begin{array}{l} m - k = d \cos \theta, n - l = d \sin \theta, \\ I(k, l) = p, I(m, n) = q \end{array} \right\}, \quad (3)$$

$$1 \leq p, q \leq b, \forall \{(k, l), (m, n)\} \in R$$

where $\#$ denotes the joint frequency of the entry (p, q) within the ranklet co-occurrence matrix RCM and b is the number of quantization levels.

As previously studied in [21, 30, 31], we adopt twelve GLCM-based texture features (codes) as given in Appendix I for texture extraction. After we quantized the ranklet coefficients into b levels, each texture code can be calculated from the normalized ranklet GLCMs ($\theta = \{0^\circ, 45^\circ, 90^\circ, 135^\circ\}$); besides, 1-pixel displacement distance (i.e., $d = 1$) between two points as used in [26, 34] within ROIs is adopted for computing the co-occurrence matrices. Previously, Masotti *et al.* [26] proposed to calculate the mean of these *invariant* angular features for texture classification. Therefore, we would derive each representative texture code by averaging the angular features, which can be defined as [30]

$$F_{i,r}^t = \frac{1}{4} \sum_{j=1}^4 F_{i,r}^t(\theta_j) / 4, 1 \leq i \leq 12, \theta_j \in \{0^\circ, 45^\circ, 90^\circ, 135^\circ\} \quad (4)$$

After we derived the feature codes for each BUS image, the multi-resolution feature extraction approach is applied to different resolution of ranklet image and the calculated texture codes are merged to produce the representative compact

textures for further texture analysis. In this paper, we adopt 1-pixel displacement ($d=1$), averaged angular feature (i.e., averaged four canonical angles) and 256 quantization levels ($b=256$) as our *default* texture descriptor. Also, we would adopt several texture descriptors as suggested in [20] to verify the diagnostic capability of the remarkable texture analyses [20, 26, 34].

C. Detailed Implementations of the Texture Analyses

To evaluate the stability and robustness of the texture analyses using ranklet transform for breast sonographic diagnoses, the GLCM-based textural features using original image (*origin*), multi-resolution wavelet images (*wavelets*) and multi-resolution ranklet images (*ranklets*) are extracted for texture analyses in this study, respectively.

In our proposed texture analysis based on *ranklets*, we decompose each BUS image into *four* ranklet resolutions (i.e., $r=\{4, 8, 16, 32\}$) and corresponding *three* orientations (i.e., $t=\{H, V, D\}$) to produce 12 ranklet transformed images (*ranklets*). After computing the ranklet GLCMs for deriving the texture codes, we will produce 144 (i.e., 4 resolutions \times 3 orientations \times 12 texture codes) texture features to form the compact image texture representation for a BUS image.

Similarly, we also apply the multi-resolution approach based on *wavelets* [17, 32, 34, 35] for texture analysis as well. Since we apply multi-resolution (multi-scale) wavelet decomposition for an input BUS image via Haar wavelet transform [17], we can generate wavelet transformed images (*wavelets*) for further GLCMs texture extraction. To objectively compare the diagnostic performance with the *ranklets*, we decompose the BUS images into *four* scales while the image size in specific scale j would shrink into the quarter of that in scale $j+1$ ($j=-1, -2, -3, -4$). $j=0$ indicates the original image and lower-resolution images while $j<0$). Instead of particularly choosing specific subband for extracting the textures [17], we derive *three* frequency subbands in each scale; namely, *HH* (Diagonal subband, D), *HL* (Horizontal subband, H), *LH* (Vertical subband, V) to generate 12 wavelet transformed images (i.e., W_j , $-1 \leq j \leq -4$, $t=\{H, V, D\}$) and 144 texture features as those of the *ranklets*. We note that the derived image with low frequency *LL* subband in scale j is the down-sampled version of that in scale $j+1$ from which we cannot derive the wavelet coefficients.

Conventional GLCM-based texture analyses extracted textures from original BUS images for tumor diagnosis [20, 21, 30, 31]. Since merely *one* image scale is used for feature extraction, we can derive *twelve* texture features from each BUS image for training the classifier. Compared to texture analyses via multi-resolution wavelet transform and ranklet transform, the diagnostic performance can be regarded as the performance baseline. Particularly, we note the coefficients within the extracted ROIs for the *three* compared texture analyses are scaled to the same range [0, 1] (which is also adopted in [20]).

Ultimately, we clarify the gray-scale *invariant* property of the BUS image pre-processed with varied contrast settings. Thus, two non-linear monotonic gray-scale transformation filters; namely, gamma correction and histogram equalization [36] are applied to produce those enhanced images. The *ranklets* R_8 and

wavelets W_1 are derived from an input image I or I filtered by three different non-linear monotonic transformations (i.e., the histogram equalization and the gamma correction with gamma value set to 0.5 and 1.5, respectively). As depicted in Fig. 3, the *ranklets* derived from original input image or enhanced images are gray-scale *invariant* as compared to those of *wavelets*. In other words, we can further extract the gray-scale *invariant* texture features from input BUS images for robust tumor diagnoses whether which distribution of texture representation collected from BUS databases are addressed. The experiments would demonstrate the effectiveness of the tumor diagnosis via robust GLCM texture extraction from the *ranklets*.

D. Texture Analysis for Breast Ultrasound Diagnosis using Support Vector Machine (SVM)

The Support Vector Machines (SVMs) have been widely used for pattern recognition research fields [13, 26, 37-39] due to the high classification accuracy and capability in dealing with high dimensional data [39, 40]. In SVM model training, we aim to solve the following convex quadratic optimization problem:

$$\begin{aligned} \min_{w,b,\xi} \quad & \frac{1}{2} w^T w + C \sum_{i=1}^N \xi_i \\ \text{s.t.} \quad & y_i (w^T \Phi(x_i) + b) \geq 1 - \xi_i, \\ & \xi_i \geq 0, i = 1, \dots, N \end{aligned} \quad (5)$$

In (5), the parameter b is the off-set of the constructed hyperplane and N is the number of training tumor cases. Furthermore, the parameter y_i is the class label and x_i is the multi-resolution texture features of the training sample, respectively. $\Phi(x_i)$ is the multi-resolution texture features represented in the transformed space (if linear kernel is adopted, $\Phi(x_i)$ is the identity transformation to x_i), w is the derived normal vector to classification hyperplane of the objective quadratic function and C controls the tradeoff between the model complexity and training error. Before training the SVM classifier, we perform feature normalization which scales each feature dimension to the same range of [0, 1] as suggested in [19]. To effectively deal with the parameter selection problem, we apply the grid search on the parameter selection [19] for constructing the SVM models. The parameter of the model can be determined by performing the k -fold cross-validation ($k=10$) using the training data with varied input parameters; the model parameter with the best classification performance would be chosen as the model parameter [19, 41]. To clarify the kernel selection issue for SVM classifier, we carry out the experiments using linear and non-linear kernel (RBF kernel) for texture analysis. We use the parameter settings for linear and RBF kernels: 1. C automatically selected by the k -fold cross validation (for linear SVMs) 2. $C = 1$ and $\gamma = 1/\#$ of features (default setting for non-linear SVM in [19]) 3. (C, γ) automatically selected by the k -fold cross validation (non-linear). It is worth noting that, we do *not* limit using specific kernel for data discrimination. Nevertheless, the experiments have shown the effectiveness of adopting *linear* SVM model for tumor classification in terms of the promising diagnostic performances and the efficiency of that which only requires one to train and store the SVM solution vector (i.e., w

in eq. 5) for future classification.

E. Statistical Analysis

The binary SVM [19] is used to classify the tumors as a malignant case or a benign one based on the proposed GLCM texture features. The probability of each tumor sample predicted by the SVM model lies between 0 and 1. We choose a threshold of 0.5 to classify the benign and malignant tumors while conducting the experiments. A tumor is classified as a malignant case if the predicted probability is equal to or larger than 0.5; otherwise, the tumor is regarded as a benign case. Furthermore, we adopt two evaluation schemes for performance comparison, namely, LOO-CV and cross-platform training/testing schemes. If the training set and testing set are the same database, then the leave-one-out cross-validation (LOO-CV) method (i.e., LOO-CV scheme) [42] is adopted for objective performance evaluation. Contrarily, the classifier trained from the training database is used to evaluate the diagnostic performance of the testing database (i.e., cross-platform training/testing scheme).

Diagnostic performance of the binary SVM model based on the proposed texture features for classifying the breast tumors is evaluated with accuracy (ACC), sensitivity (SENS), specificity (SPEC), positive and negative predictive values (PPV and NPV). Moreover, the receiver operating characteristic (ROC) curves are obtained by using ROCKIT software (C. Metz; University of Chicago, Chicago, IL, USA) and the area under the ROC curve (AUC) is adopted as one of the indicator of diagnostic performance. The AUC value can be derived by adjusting different thresholds of the class probability and can be generated by the ROCKIT software. In addition, statistical analyses except ROC (AUC) are performed by using the software (SPSS, version 16 for Windows; SPSS, Chicago, IL, USA).

IV. EXPERIMENTS

Experiments are conducted using *three* different breast sonographic platforms for performance comparisons based on the proposed texture analyses (i.e., *origin*, *wavelets* and *ranklets*). To clarify the issues about the selection of scales for the proposed multi-resolution features and that of SVM kernels, we carry out the experiments with aforementioned parameter settings for building the SVM models. Furthermore, the texture descriptor as used in [26] (i.e., 1-pixel displacement and averaged angular feature) with 256 quantization levels (*default* texture descriptor) is adopted. As shown in Fig. 4, we observed using *two* resolutions (or image scales) for *wavelets* and *ranklets* with linear SVM kernel can produce promising diagnostic performance for tumor classification (i.e., the first two scales are used for *wavelets* while resolutions {4, 8} are adopted for *ranklets*). Though we can use more training features for objects classification to increase the model complexity and simultaneously decrease the training errors [28, 29]. Nevertheless, the training linear SVM models would over-fit the training data with limited observed samples while using too much features for training the classifier [43]. For objectively comparing the diagnostic performance, we adopt *two*

resolutions (i.e., 72 texture features) for feature extraction and *linear* SVM kernel for building the training models to verify the stability and robustness of the proposed texture analysis via ranklet transform. Note that all algorithms except SVM classifiers (LIBSVM package [19]) were implemented in MATLAB R2010a on an Intel quad-core PC with 2.23 GHz processor and 2G RAM. We also list the computation time for deriving a ranklet pixel with specific resolutions in Appendix II.

A. Stability of the Texture Analysis via Ranklet Transform

To verify the stability of the proposed texture analysis, we derived the area under the ROC curve (i.e., AUC values) to demonstrate the stable diagnostic performances of *ranklets*. In Fig. 4, AUC values for the *three* databases via ranklet transform are 0.918 (95% confidence interval [CI], 0.848 to 0.961), 0.943 (95% CI, 0.906 to 0.968) and 0.934 (95% CI, 0.883 to 0.961), respectively; whereas for texture analysis via wavelet transform are 0.847 (95% CI, 0.762 to 0.910), 0.922 (95% CI, 0.878 to 0.958) and 0.867 (95% CI, 0.798 to 0.914), respectively. The experiments demonstrate the stability of the texture analysis via *ranklets*, which is observed to be less sensitive to the breast sonographic platforms.

Considering the performance evaluation with LOO-CV bootstrap scheme might be upward bias [44], we adopt 0.632+ bootstrap estimators [45] (as also used in [20]) with 500 independent bootstrap samples to evaluate the diagnostic performance. To build each independent bootstrap sample, we randomly resample the training dataset with replacement, and the test samples are selected which are not included in the training set. Furthermore, we consider the combination of rotation angle (i.e., averaged angular feature and 90 rotation angle) and quantization level (i.e., 32, 64 and 256) as suggested in [20, 26] for computing the co-occurrence matrices. The average AUC values calculated from the 500 bootstrap samples for the *three* texture analyses of each database are shown in Fig. 5. The experiments show that texture analysis using original images (*origin*) are basically consistent with those reported in [20] and the best diagnostic performances of *wavelets* and *ranklets* are obtained while *default* texture descriptor is adopted (as also used in [26]). To be more specific, we list the detailed diagnostic performance using *default* texture descriptor for the *three* texture analyses in Table II.

To the best of our knowledge, there is no study aiming at searching for the optimal GLCM texture descriptors using *wavelets* or *ranklets* for BUS tumor classification. The averaged angular feature of texture descriptor as used in [26, 30] seems to perform well for *rotation-invariant* ranklet transform instead of merely considering specific rotation angle for texture analysis. Moreover, [20, 46] have shown the effectiveness of coefficients quantization for improving the classification performance, the ranklet transform quantizes (decomposes) the gray values of the BUS images into ranklet coefficients which is robust to the speckle noise existed in BUS images [20]. There is worth to note, we do *not* assume the *default* texture descriptor adopted in this paper would be the optimal. Particularly, we aim to examine how the *variant* textures would influence the diagnostic performance while applying *varied* texture representations of collected BUS data for texture analyses. Our experiments show

the usefulness of extracting *invariant* textures for developing a stable CAD system.

B. Robustness of the Texture Analysis via Ranklet Transform

To verify the robustness of the texture analyses using *ranklets* for tumor diagnosis, we conduct the experiments to derive the AUC values of cross-platform training/testing combination between each database. There is worth noting that, we resize the BUS images using bicubic interpolation down-sampling from different databases into the same pixel resolution before texture feature extraction to validate the correctness of cross-platform training/testing scheme. More specifically, all images from the two databases (i.e., Database B and C) are resized to a lower image resolution (0.12 mm/pixel) as Database A. The robust texture analysis for breast sonographic tumor diagnosis concerns one of the possible scenarios while the training data of new BUS platform are unavailable and patient cases need to be re-collected for months for training a new SVM model. We have to note that the parameter of the training *linear* SVM model with input data from the training platform is tuned as suggested in [19]. Afterward, the trained SVM model is adopted to evaluate the test data from the testing platform. As shown in Table III, the diagnostic performances (i.e., AUC values) of *ranklets* outperform those of *origin* and *wavelets* whenever which combinations of training/testing platforms are performed. Furthermore, we observed that the diagnostic performance would suffer from *performance degradation* while the cross-platform training/testing scheme is applied (LOO-CV scheme is the performance baseline). The percentages of the *performance degradation* (based on AUC values) of the test databases for *ranklets* are between 0.93% and 6.50% ($4.77\%\pm2.25\%$), for *wavelets* are between 2.90% and 21.74% ($11.70\%\pm6.62\%$) and for *origin* are between 1.92% and 10.28% ($7.20\%\pm3.43\%$). We note that the texture analysis using *ranklets* presents a better generalization of SVM model in classifying BUS images collected from different sonographic platforms.

To clarify the issue about the performance degradation while comparing the *two* adopted evaluation schemes (i.e., cross-platform training/testing and LOO-CV schemes). One of the possible reasons might be related to the different gray-scale distribution between *two* databases. Thus, we further calculate the intensity mean and standard deviation for each tumor within ROI. And then, the mean and standard deviation (i.e., Means and SDs as shown in Table IV) for the intensity mean and standard deviation calculated from tumor ROI can be obtained. Table IV demonstrates that if the cross-platform training/testing scheme is applied, the trained SVM model is not properly for evaluating the test database with differentiable Means and SDs between training and testing database, which results in significant *performance degradation* (especially for *wavelets* which particularly consider the frequency information existed in the BUS image). Since there are no researches aiming to investigate the relationship between *varied* texture patterns and diagnostic performance using BUS images, our experiments obviously demonstrate that the diagnostic performances using *invariant* texture from *ranklets* are less sensitive to the breast sonographic platforms, which is relatively robust to address the

BUS databases with different distributions of texture representation.

C. Statistical Analysis for the Texture Analyses

We conduct a *z*-test [47, 48] on the AUC values of the experiments for the two aforementioned evaluation schemes to prove whether the diagnostic performance of *ranklets* statistically outperform the other two methods. The difference of the diagnostic performance between two methods is statistically significant if the *p*-value is less than 0.05. Table V indicates the diagnostic performance of *ranklets* outperforms the other two methods in most of the evaluations which shows the efficacy for tumor diagnosis while applying texture analysis via ranklet transform.

Moreover, we use AUC values generated by the 0.632+ bootstrap estimators with *default* texture descriptor to perform the statistical analysis for comparing the *three* texture analyses. Before we perform the significance test between the methods, the Kolmogorov-Smirnov test is applied to test the normality of the distribution of the AUC values generated by each evaluated group. Due to all the methods of each database present normal distribution, the *F*-test is further used for verifying the equality of variances between two groups. The Welch's *t*-test (*p*<0.05) is substituted for the Student's *t* test when the hypothesis of equal variances is rejected, which all the compared groups present unequal variance. Based on the 0.632+ bootstrap estimators, Table VI states the texture analysis using *ranklets* statistically outperform that using *wavelets* and *origin*, respectively.

V. DISCUSSION AND CONCLUSION

Conventional texture analyses for tumor diagnosis aim at extracting textural features from gray-scale BUS images [13, 15, 16, 20, 21] or gray-scale images represented in the frequency domain [17]. However, most prior work did not consider the stability and robustness when designing CAD for practical tumor diagnosis applications. Gomez-Flores *et al* [20] dedicated the selection of effective texture descriptors for deriving informative GLCM texture features for classifying the BUS images, and their best AUC performance can achieve 0.87 ± 0.02 using 0.632+ bootstrap estimators. In this paper, our experiments demonstrated that the GLCM-based texture analysis using multi-resolution features via wavelet transform [17, 32] further improved the diagnostic performance. Nevertheless, addressing *varied* texture representations of BUS images from different sonographic platforms would produce gray-scale *variant* features and affect the consistency of the diagnostic performance. While the robust and stable properties are desired for texture analyses, we developed a robust feature extraction scheme via ranklet transform [26, 29] to derive the gray-scale *invariant* features for this purpose.

Different from wavelet transform, the ranklet transform of BUS images merely consider the corresponding ranks of the gray values and the gray-scale *invariant* texture features can be extracted to train the robust classifier for classifying the masses. Our experiments show the diagnostic performances of the

texture analysis via ranklet transform can achieve best and consistent performance using the LOO-CV scheme and without significant performance variations for cross-platform train/test scheme. We also observed that the diagnostic performance would degrade remarkably using *wavelets* and *origin*. In addition, due to the operator-dependent issue in BUS, it is not clear whether training a classifier using data with *varied* texture representations collected by different operators on a sonographic platform is applicable. That is, the parameters of BUS systems would lead to increased difficulties in providing quantitative and qualitative tumor diagnosis. Since some studies [22, 49] proposed to enhance the texture patterns existing in BUS via fuzzy logic or to estimate the parameters of the ultrasonic devices with log-compressed K-distribution, texture analyses with gray-scale *variant* feature for these methods via original images or wavelet transform cannot be guaranteed to achieve robust and stable tumor diagnoses in different BUS platforms as well.

In this paper, we proposed to extract gray-scale *invariant* features via ranklet transform for designing a cross-platform and practical CAD application. Our experiments confirmed that we produced stable and robust tumor diagnosis while GLCM-based texture analysis using multi-resolution ranklet transform was applied. For improving BUS tumor classification, more effective GLCM-based textures as used in [20] and the range of angular features as suggested in [30] will be introduced in our further works. On the other hand, texture analysis using local binary patterns (LBPs) [23, 24] based on multi-resolution approach is known to result in gray-scale *invariant* features as well, which will also be considered and compared in our ongoing study. Finally, our future research directions also include the investigation of feature selection techniques, which have been widely used in pattern recognition for improving recognition performance based on robustness requirement assumption (i.e., the selected features set are useful across different ultrasonic devices).

APPENDIX I

Twelve texture features adopted in this work are extracted from the GLCMs as described in Table VII. The symbols used in this paper and detailed description are listed in Table VIII.

APPENDIX II

We test our proposed ranklet transform to derive the computation time with various input ranklet resolutions (i.e., $r = \{2, 4, 8, 16, 32, 64\}$). The computation time is expressed in microsecond (μs) per ranklet point (rp), which is shown in Fig.6.

REFERENCES

- [1] P. M. Shankar, C. W. Piccoli, J. M. Reid, J. Forsberg, and B. B. Goldberg, "Application of the compound probability density function for characterization of breast masses in ultrasound B scans," *Physics in Medicine and Biology*, vol. 50, pp. 2241-2248, May 21 2005.
- [2] K. J. W. Taylor, C. Merritt, C. Piccoli, R. Schmidt, G. Rouse, E. Rubin, D. Georgian-Smith, F. Winsberg, B. Goldberg, and E. Mendelson, "Ultrasound as a complement to mammography and breast examination to characterize breast masses," *Ultrasound in Medicine and Biology*, vol. 28, pp. 19-26, Jan 2002.
- [3] A. T. Stavros, D. Thickman, C. L. Rapp, M. A. Dennis, S. H. Parker, and G. A. Sisney, "Solid breast nodules: use of sonography to distinguish between benign and malignant lesions," *Radiology*, vol. 196, pp. 123-34, Jul 1995.
- [4] W. A. Berg, J. D. Blume, J. B. Cormack, E. B. Mendelson, D. Lehrer, M. Bohm-Velez, E. D. Pisano, R. A. Jong, W. P. Evans, M. J. Morton, M. C. Mahoney, L. H. Larsen, R. G. Barr, D. M. Farria, H. S. Marques, K. Boparai, and A. Investigators, "Combined screening with ultrasound and mammography vs mammography alone in women at elevated risk of breast cancer," *Jama-Journal of the American Medical Association*, vol. 299, pp. 2151-2163, May 14 2008.
- [5] S. S. Kaplan, "Clinical utility of bilateral whole-breast US in the evaluation of women with dense breast tissue," *Radiology*, vol. 221, pp. 641-9, Dec 2001.
- [6] P. Crystal, S. D. Strano, S. Shcharynski, and M. J. Koretz, "Using sonography to screen women with mammographically dense breasts," *American Journal of Roentgenology*, vol. 181, pp. 177-182, Jul 2003.
- [7] D. R. Chen, R. F. Chang, W. J. Kuo, M. C. Chen, and Y. L. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound Med Biol*, vol. 28, pp. 1301-10, Oct 2002.
- [8] Y. L. Huang and D. R. Chen, "Support vector machines in sonography: application to decision making in the diagnosis of breast cancer," *Clin Imaging*, vol. 29, pp. 179-84, May-Jun 2005.
- [9] N. Piliouras, I. Kalatzis, N. Dimitropoulos, and D. Cavouras, "Development of the cubic least squares mapping linear-kernel support vector machine classifier for improving the characterization of breast lesions on ultrasound," *Comput Med Imaging Graph*, vol. 28, pp. 247-55, Jul 2004.
- [10] R. F. Chang, W. J. Wu, W. K. Moon, and D. R. Chen, "Automatic ultrasound segmentation and morphology based diagnosis of solid breast tumors," *Breast Cancer Res Treat*, vol. 89, pp. 179-85, Jan 2005.
- [11] M. André, M. Galperin, L. Olson, K. Richman, S. Payrovi, and P. Phan, "Improving the Accuracy of Diagnostic Breast Ultrasound," in *Acoustical Imaging*, vol. 26, R. Maev, Ed., ed: Springer US, 2002, pp. 453-460.
- [12] W. A. Berg, L. Gutierrez, M. S. NessAiver, W. B. Carter, M. Bhargavan, R. S. Lewis, and O. B. Ioffe, "Diagnostic accuracy of mammography, clinical examination, US, and MR imaging in preoperative assessment of breast cancer," *Radiology*, vol. 233, pp. 830-49, Dec 2004.
- [13] Y. L. Huang, K. L. Wang, and D. R. Chen, "Diagnosis of breast tumors with ultrasonic texture analysis using support vector machines," *Neural Computing & Applications*, vol. 15, pp. 164-169, Apr 2006.
- [14] R. F. Chang, W. J. Wu, W. K. Moon, and D. R. Chen, "Improvement in breast tumor discrimination by support vector machines and speckle-emphasis texture analysis," *Ultrasound in Medicine and Biology*, vol. 29, pp. 679-86, May 2003.
- [15] A. V. Alvarenga, W. C. Pereira, A. F. Infantosi, and C. M. Azevedo, "Complexity curve and grey level co-occurrence matrix in the texture evaluation of breast tumor on ultrasound images," *Med Phys*, vol. 34, pp. 379-87, Feb 2007.
- [16] W. K. Moon, Y. W. Shen, C. S. Huang, L. R. Chiang, and R. F. Chang, "Computer-Aided Diagnosis for the Classification of Breast Masses in Automated Whole Breast Ultrasound Images," *Ultrasound in Medicine and Biology*, vol. 37, pp. 539-548, Apr 2011.
- [17] D. R. Chen, R. F. Chang, W. J. Kuo, M. C. Chen, and Y. L. Huang, "Diagnosis of breast tumors with sonographic texture analysis using wavelet transform and neural networks," *Ultrasound in Medicine and Biology*, vol. 28, pp. 1301-1310, Oct 2002.
- [18] N. N. Tsiparas, S. Golemati, I. Andreadis, J. S. Stoitsis, I. Valavanis, and K. S. Nikita, "Comparison of Multiresolution Features for Texture Classification of Carotid Atherosclerosis From B-Mode Ultrasound," *Ieee Transactions on Information Technology in Biomedicine*, vol. 15, pp. 130-137, Jan 2011.
- [19] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, pp. 1-27, 2011.
- [20] W. Gomez-Flores, W. Pereira, and A. Infantosi, "Analysis of Co-occurrence Texture Statistics as a Function of Gray-Level Quantization for Classifying Breast Ultrasound," *IEEE Trans Med Imaging*, Jun 28 2012.
- [21] D. A. Clausi, "An analysis of co-occurrence texture statistics as a function of grey level quantization," *Canadian Journal of Remote Sensing*, vol. 28, pp. 45-62, Feb 2002.
- [22] A. Takemura, A. Shimizu, and K. Hamamoto, "Discrimination of Breast

- Tumors in Ultrasonic Images Using an Ensemble Classifier Based on the AdaBoost Algorithm With Feature Selection," *Ieee Transactions on Medical Imaging*, vol. 29, pp. 598-609, Mar 2010.
- [23] T. Ojala, M. Pietikainen, and T. Maenpaa, "Multiresolution gray-scale and rotation invariant texture classification with local binary patterns," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 971-987, Jul 2002.
- [24] M. M. S. Matsumoto, C. M. Sehgal, and J. K. Udupa, "Local binary pattern texture-based classification of solid masses in ultrasound breast images," pp. 83201H-83201H, 2012.
- [25] F. Smeraldi and J. Bigun, "Retinal vision applied to facial features detection and face authentication," *Pattern Recognition Letters*, vol. 23, pp. 463-475, Feb 2002.
- [26] M. Masotti and R. Campanini, "Texture classification using invariant ranklet features," *Pattern Recognition Letters*, vol. 29, pp. 1980-1986, Oct 15 2008.
- [27] E. L. Lehmann and H. J. M. D'Abreu, *Nonparametrics : statistical methods based on ranks*, Rev. 1st ed. New York: Springer, 2006.
- [28] M. Masotti, "A ranklet-based image representation for mass classification in digital mammograms," *Med Phys*, vol. 33, pp. 3951-61, Oct 2006.
- [29] F. Bianconi, A. Fernandez, E. Gonzalez, and J. Armesto, "Robust color texture features based on ranklets and discrete Fourier transform," *Journal of Electronic Imaging*, vol. 18, Oct-Dec 2009.
- [30] R. M. Haralick, Shanmuga.K, and I. Dinstein, "Textural Features for Image Classification," *Ieee Transactions on Systems Man and Cybernetics*, vol. Smc3, pp. 610-621, 1973.
- [31] L. K. Soh and C. Tsatsoulis, "Texture analysis of SAR sea ice imagery using gray level co-occurrence matrices," *Ieee Transactions on Geoscience and Remote Sensing*, vol. 37, pp. 780-795, Mar 1999.
- [32] S. G. Mallat, "A Theory for Multiresolution Signal Decomposition - the Wavelet Representation," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 11, pp. 674-693, Jul 1989.
- [33] G. E. Mailloux, M. Bertrand, R. Stampfler, and S. Ethier, "Local Histogram Information-Content of Ultrasound B-Mode Echographic Texture," *Ultrasound in Medicine and Biology*, vol. 11, pp. 743-750, 1985.
- [34] M. M. Mokji and S. A. R. A. Bakar, "Gray Level Co-Occurrence Matrix Computation Based On Haar Wavelet," presented at the Proceedings of the Computer Graphics, Imaging and Visualisation, 2007.
- [35] N. Linder, J. Konsti, R. Turkki, E. Rahtu, M. Lundin, S. Nordling, C. Haglund, T. Ahonen, M. Pietikainen, and J. Lundin, "Identification of tumor epithelium and stroma in tissue microarrays using texture analysis," *Diagnostic Pathology*, vol. 7, Mar 2 2012.
- [36] J. Benson, "Is screening mammography safe for high-risk patients?," *Lancet Oncol*, vol. 7, pp. 360-2, May 2006.
- [37] K. I. Kim, K. Jung, S. H. Park, and H. J. Kim, "Support vector machines for texture classification," *Ieee Transactions on Pattern Analysis and Machine Intelligence*, vol. 24, pp. 1542-1550, Nov 2002.
- [38] X. J. Peng and D. Xu, "Bi-density twin support vector machines for pattern recognition," *Neurocomputing*, vol. 99, pp. 134-143, Jan 1 2013.
- [39] R. F. Chang, W. J. Wu, C. C. Tseng, D. R. Chen, and W. K. Moon, "3-D snake for US in margin evaluation for malignant breast tumor excision using Mammotome," *IEEE Trans Inf Technol Biomed*, vol. 7, pp. 197-201, Sep 2003.
- [40] E. Fabisewska, I. Grabska, K. Jankowska, E. Wesolowska, and W. Bulski, "Comparison of results from quality control of physical parameters and results from clinical evaluation of mammographic images for the mammography screening facilities in Poland," *Radiat Prot Dosimetry*, vol. 147, pp. 206-9, Sep 2011.
- [41] J. Shawe-Taylor and N. Cristianini, *Kernel methods for pattern analysis*. Cambridge, UK ; New York: Cambridge University Press, 2004.
- [42] E. Alpaydin, *Introduction to machine learning*. Cambridge: Mass: MIT Press, 2004.
- [43] G. W. Heiman, *Basic statistics for the behavioral sciences*, 6th ed. Belmont, CA: Wadsworth Cengage Learning, 2011.
- [44] B. Sahiner, H. P. Chan, and L. Hadjiiski, "Classifier performance prediction for computer-aided diagnosis using a limited dataset," *Med Phys*, vol. 35, pp. 1559-1570, Apr 2008.
- [45] B. Efron and R. Tibshirani, "Improvements on cross-validation: The .632+ bootstrap method," *Journal of the American Statistical Association*, vol. 92, pp. 548-560, Jun 1997.
- [46] R. Xu, X. Zhao, X. Li, and C. I. Chang, "Target detection with improved image texture feature coding method and support vector machine," *International Journal of Intelligent Technology*, vol. 1, pp. 47-56, 2006.
- [47] C. H. Brase and C. P. Brase, *Understanding basic statistics*, 5th ed. Belmont, CA: Brooks/Cole, Cengage Learning, 2010.
- [48] J. A. Hanley and B. J. Mcneil, "A Method of Comparing the Areas under Receiver Operating Characteristic Curves Derived from the Same Cases," *Radiology*, vol. 148, pp. 839-843, 1983.
- [49] Y. H. Guo, H. D. Cheng, J. H. Huang, J. W. Tian, W. Zhao, L. T. Sun, and Y. X. Su, "Breast ultrasound image enhancement using fuzzy logic," *Ultrasound in Medicine and Biology*, vol. 32, pp. 237-247, Feb 2006.

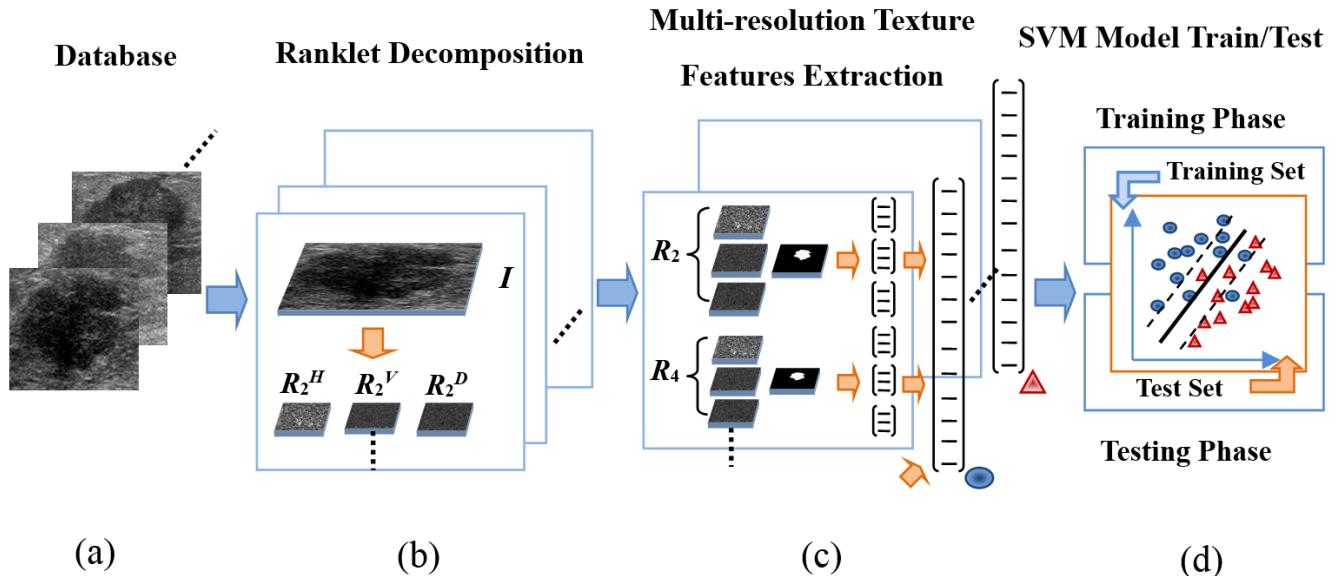


Fig. 1 The framework of the proposed multi-resolution texture analysis via ranklet transform. (a) Database (b) A BUS image I is decomposed into multiple ranklet images (i.e., $R_2^H, R_2^V, R_2^D \dots$) using ranklet transform (c) Multi-resolution ($R_2, R_4 \dots$) texture features are extracted from the *ranklets* to form the compact texture representation for each BUS image (d) SVM classifier is adopted to classify each BUS image from collected database into a benign tumor or a malignant one.

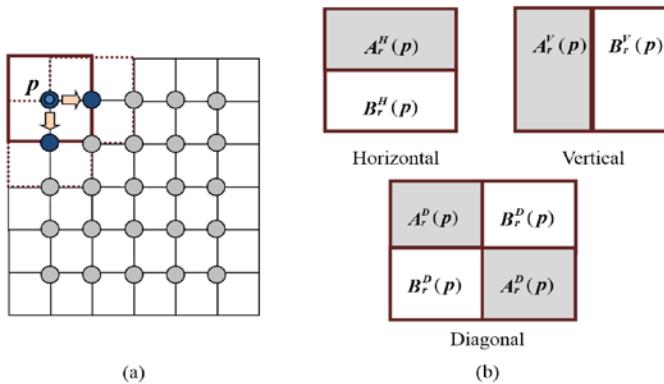


Fig. 2 (a) The diagram for determining the ranklet coefficient of an arbitrary point p and resolution r in a moving square crop (b) The geometrical representation of the three orientations horizontal (H), vertical (V) and diagonal (D) directions of the square crop with subset $A_r^t(p)$ and $B_r^t(p)$.

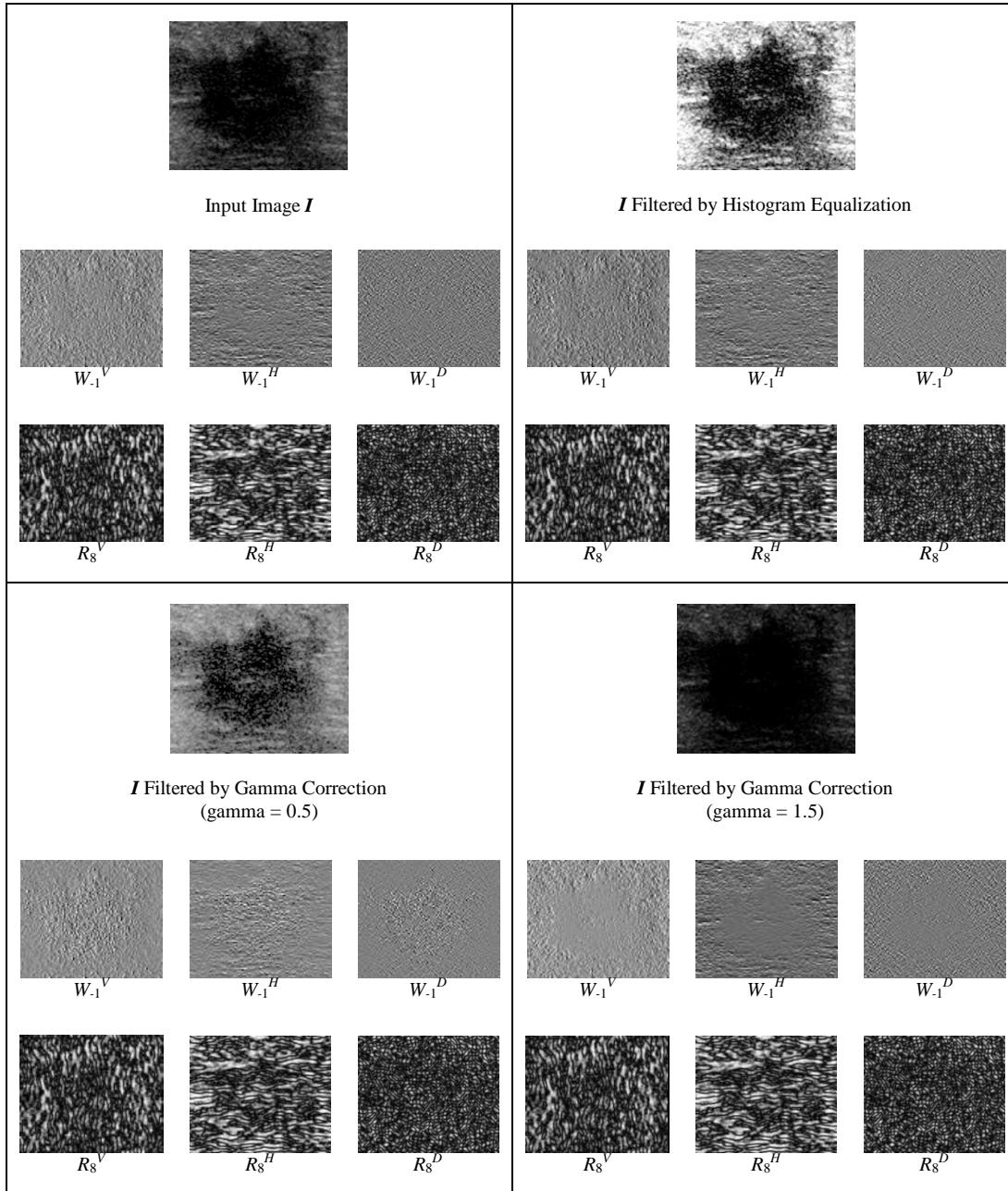


Fig. 3 The ranklets (R_8) and wavelets (W_1) with corresponding three orientations horizontal (H), vertical (V) and diagonal (D) are derived from an input BUS image I , I filtered by histogram equalization and gamma correction, respectively. Note that the coefficients of the wavelet and ranklet images are scaled for clear visualization.

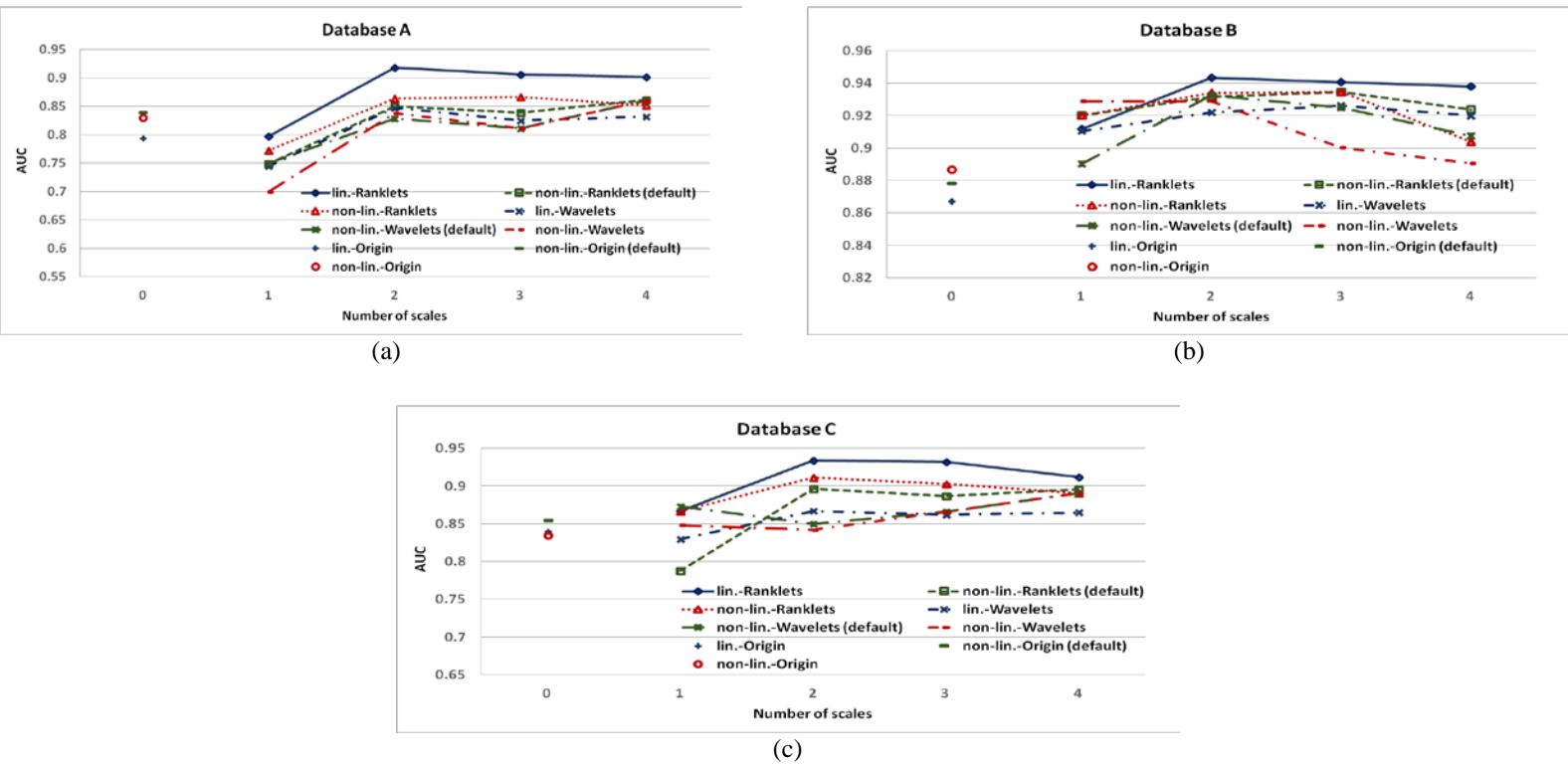


Fig. 4 The diagnostic performance (AUC) of *origin*, *wavelets* and *ranklets* using different number of scales (features) for training the SVM classifiers of each database. The selected parameter C for linear SVMs of the three databases are (a) {2, 2, 32} (b) {16, 16, 8} (c) {8, 8, 8}, respectively. Note that the parameter set for the three methods is expressed as {*origin*, *wavelets*, *ranklets*}.

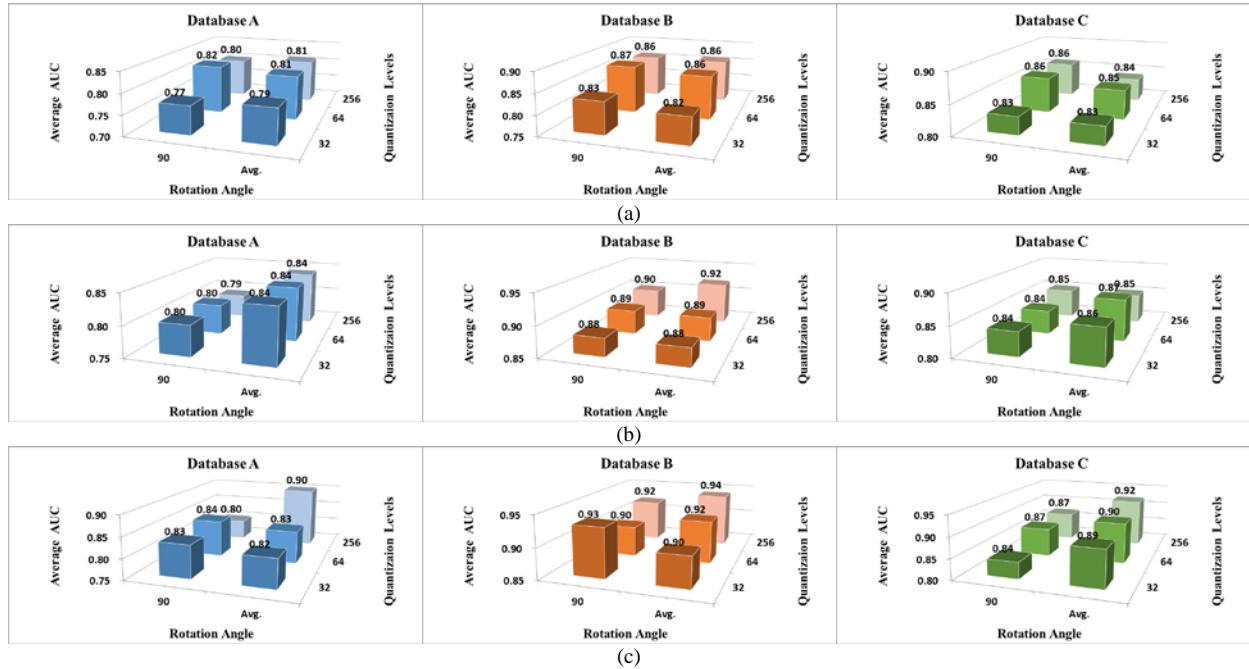


Fig. 5 Average AUC values calculated from 500 bootstrap samples of each database for the three texture analyses (a) *origin* (b) *wavelets* (c) *ranklets*.

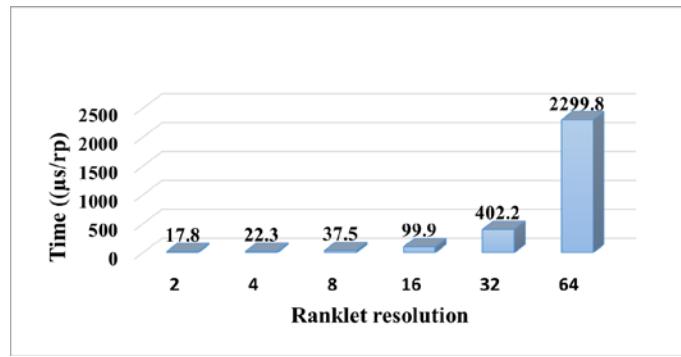


Fig. 6 The computation time ($\mu\text{s}/\text{rp}$) of the ranklet transform with specific resolution $r = \{2, 4, 8, 16, 32, 64\}$

TABLE I
TUMOR HISTOLOGICAL DISTRIBUTIONS OF THE COLLECTED DATABASES

Finding	Database A		Database B		Database C	
	No. of Cases	Tumor Size (cm)	No. of Cases	Tumor Size (cm)	No. of Cases	Tumor Size (cm)
Benign						
Fibroadenoma	44	1.65 ± 0.65	72	1.47 ± 0.55	61	1.83 ± 0.73
Fibrocystic change	32	1.03 ± 0.50	53	0.97 ± 0.52	40	1.25 ± 0.55
Papilloma	2	1.45 ± 0.07	8	0.94 ± 0.40	3	1.30 ± 0.36
Malignant						
Infiltrating ductal carcinoma	33	1.53 ± 0.70	46	2.12 ± 0.79	35	2.21 ± 0.73
Ductal carcinoma in situ	3	1.43 ± 0.49	14	2.09 ± 0.69	20	2.19 ± 0.58
Invasive tubular carcinoma	2	1.90 ± 0.42	0	0	2	2.35 ± 0.35
Total	116	1.37 ± 0.66	193	1.47 ± 0.80	161	1.80 ± 0.76

TABLE II
ROC TEXTURE ANALYSES (MEAN \pm STANDARD DEVIATION) FOR ORIGIN, WAVELETS AND RANKLETS WITH DEFAULT TEXTURE DESCRIPTOR AS WELL AS MINIMUM PREDICTION ERROR ($\hat{\epsilon}$) CALCULATED FROM 500 INDEPENDENT BOOTSTRAP SAMPLES

Database	Method	AUC	ACC (%)	SENS (%)	SPEC (%)	PPV (%)	NPV (%)	$\hat{\epsilon}$
A	Origin	0.81 ± 0.03	74.28 ± 2.27	63.93 ± 5.78	79.39 ± 3.14	60.85 ± 5.22	81.50 ± 3.09	0.213
	Wavelets	0.84 ± 0.03	79.45 ± 1.73	70.54 ± 4.23	83.76 ± 2.39	68.20 ± 4.18	85.23 ± 2.32	0.174
	Ranklets	0.90 ± 0.02	81.68 ± 1.69	69.66 ± 4.63	87.55 ± 2.15	73.49 ± 4.00	85.36 ± 2.39	0.146
B	Origin	0.86 ± 0.03	78.75 ± 2.21	66.53 ± 5.99	84.33 ± 2.89	66.20 ± 4.91	80.00 ± 2.95	0.169
	Wavelets	0.92 ± 0.02	84.23 ± 1.71	74.99 ± 4.74	88.38 ± 2.12	74.66 ± 3.82	88.60 ± 2.18	0.118
	Ranklets	0.94 ± 0.02	86.35 ± 1.64	79.56 ± 4.44	89.35 ± 2.11	77.08 ± 4.18	90.67 ± 2.12	0.097
C	Origin	0.84 ± 0.03	76.49 ± 2.28	67.74 ± 5.30	81.31 ± 3.32	67.11 ± 4.43	81.80 ± 3.19	0.188
	Wavelets	0.85 ± 0.02	77.14 ± 1.74	69.27 ± 3.87	81.48 ± 2.51	67.63 ± 3.82	82.60 ± 2.43	0.193
	Ranklets	0.92 ± 0.02	84.58 ± 1.70	81.50 ± 3.66	86.19 ± 2.44	76.27 ± 3.97	89.52 ± 2.22	0.119

Note. The best performance for each database is highlighted with bold.

TABLE III
DIAGNOSTIC PERFORMANCE EVALUATION (AUC VALUES AND 95% CI ARE LISTED) OF CROSS-PLATFORM TRAINING/TESTING COMBINATION BETWEEN EACH COLLECTED DATABASE

Train Database	Method	Test Database		
		B	C	B+C
A	Origin (95% CI)	0.783 (0.717-0.846)	0.823* (0.769-0.892)	0.786 (0.739-0.835)
	Wavelets (95% CI)	0.722† (0.6283-0.788)	0.752† (0.663-0.816)	0.729† (0.679-0.782)
	Ranklets (95% CI)	0.934* (0.896-0.965)	0.877 (0.825-0.929)	0.876* (0.837-0.909)
B		A	C	A+C
	Origin (95% CI)	0.724 (0.633-0.816)	0.807* (0.736-0.869)	0.764† (0.707-0.817)
	Wavelets (95% CI)	0.757† (0.643-0.828)	0.842* (0.784-0.906)	0.832* (0.779-0.879)
	Ranklets (95% CI)	0.867 (0.792-0.929)	0.873 (0.817-0.922)	0.875* (0.825-0.909)
C		A	B	A+B
	Origin (95% CI)	0.709† (0.614-0.806)	0.789 (0.720-0.860)	0.765† (0.708-0.821)
	Wavelets (95% CI)	0.795 (0.677-0.864)	0.785† (0.712-0.854)	0.808† (0.731-0.843)
	Ranklets (95% CI)	0.859 (0.780-0.913)	0.913* (0.871-0.949)	0.891* (0.855-0.925)

* The performance degradation is less than 5% for cross-platform training/testing scheme (regarding LOO-CV scheme as performance baseline).

† The performance degradation is larger than 10% for cross-platform training/testing scheme (regarding LOO-CV scheme as performance baseline).

TABLE IV
MEAN (Means) AND STANDARD DEVIATION (SDs) FOR EACH DATABASE WITH INTENSITY MEAN AND STANDARD DEVIATION CALCULATED FROM EACH TUMOR ROI

Database	Pathology Type	Means (mean ± standard deviation)	SDs (mean ± standard deviation)
A	Benign (n=78)	47.04±15.30	19.79±6.30
	Malignant (n=38)	38.44±8.840	20.02±5.13
	All (n=116)	44.22±14.09	19.87±5.92
B	Benign (n=133)	82.93±18.54	26.99±5.07
	Malignant (n=60)	74.51±16.89	25.16±4.23
	All (n=193)	80.31±18.42	26.42±4.89
C	Benign (n=104)	57.55±16.68	26.98±7.80
	Malignant (n=57)	52.91±15.02	26.28±6.21
	All (n=161)	55.91±16.21	26.73±7.26

TABLE V
THE P-VALUE OF THE Z-TEST ON THE AUC VALUE WHILE APPLYING CROSS-PLATFORM TRAINING/TESTING OR LOO-CV SCHEME BETWEEN RANKLETS AND WAVELETS, RANKLETS AND ORIGIN

Train Database	Test Database	Ranklets vs. Wavelets	Ranklets vs. Origin
A	A	0.040*	0.006*
	B	<0.001*	<0.001*
	C	<0.001*	0.152
	B+C	<0.001*	<0.001*
B	A	0.008*	<0.001*
	B	0.397	0.008*
	C	0.438	0.009*
	A+C	0.026*	<0.001*
C	A	0.136	<0.001*
	B	<0.001*	<0.001*
	C	0.042*	0.012*
	A+B	<0.001*	<0.001*

* Indicates the performance difference between two methods is statistically significant.

† If train and test database are the same, the LOO-CV scheme is applied.

TABLE VI
THE P-VALUE OF THE WELCH'S T-TEST ON THE DISTRIBUTION OF THE AUC VALUES BETWEEN RANKLETS AND WAVELETS, RANKLETS AND ORIGIN

Database	Ranklets vs. Wavelets	Ranklets vs. Origin
A	<0.0001*	<0.0001*
B	<0.01*	0.0000*
C	<0.0001*	0.0000*

* Indicates the performance difference between two methods is statistically significant.

TABLE I
TEXTURE FEATURES EXTRACTED FROM RANKLET GLCMS

Code	Feature	Equation	Ref.
$F_{1,r}^t(\theta)$	Energy	$\sum_{i=1}^b \sum_{j=1}^b RCM_{r,\theta}^t(i,j)^2$	[30]
$F_{2,r}^t(\theta)$	Entropy	$-\sum_{i=1}^b \sum_{j=1}^b RCM_{r,\theta}^t(i,j) \log_2(RCM_{r,\theta}^t(i,j))$	[30]
$F_{3,r}^t(\theta)$	Correlation	$\frac{\sum_{i=1}^b \sum_{j=1}^b (i - \mu_{x,r,\theta}^t)(j - \mu_{y,r,\theta}^t) RCM_{r,\theta}^t(i,j)}{\sigma_{x,r,\theta}^t \sigma_{y,r,\theta}^t}$	[30]
$F_{4,r}^t(\theta)$	Inverse Difference Moment Normalized	$\sum_{i=1}^b \sum_{j=1}^b \frac{1}{1 + \frac{(i-j)^2}{b}} RCM_{r,\theta}^t(i,j)$	[30]
$F_{5,r}^t(\theta)$	Inertia	$\sum_{i=1}^b \sum_{j=1}^b (i - j)^2 RCM_{r,\theta}^t(i - j)$	[30]
$F_{6,r}^t(\theta)$	Cluster Shade	$\sum_{i=1}^b \sum_{j=1}^b ((i - \mu_{x,r,\theta}^t) + (j - \mu_{y,r,\theta}^t))^3 RCM_{r,\theta}^t(i - j)$	[30]
$F_{7,r}^t(\theta)$	Cluster Prominence	$\sum_{i=1}^b \sum_{j=1}^b ((i - \mu_{x,r,\theta}^t) + (j - \mu_{y,r,\theta}^t))^4 RCM_{r,\theta}^t(i,j)$	[30]
$F_{8,r}^t(\theta)$	Haralick's Correlation	$\frac{\sum_{i=1}^b \sum_{j=1}^b (i \cdot j) RCM_{r,\theta}^t(i,j) - \mu_{x,r,\theta}^t \mu_{y,r,\theta}^t}{\sigma_{x,r,\theta}^t \sigma_{y,r,\theta}^t}$	[30]
$F_{9,r}^t(\theta)$	Autocorrelation	$\sum_{i=1}^b \sum_{j=1}^b (i \cdot j) RCM_{r,\theta}^t(i,j)$	[31]
$F_{10,r}^t(\theta)$	Dissimilarity	$\sum_{i=1}^b \sum_{j=1}^b (i - j) RCM_{r,\theta}^t(i,j)$	[31]
$F_{11,r}^t(\theta)$	Homogeneity	$\sum_{i=1}^b \sum_{j=1}^b \frac{1}{1 + (i - j)^2} RCM_{r,\theta}^t(i,j)$	[31]
$F_{12,r}^t(\theta)$	Maximum Probability	$\operatorname{argmax}_{i,j} RCM_{r,\theta}^t(i,j)$	[21]

TABLE II
 THE DESCRIPTION OF THE SYMBOLS USED IN THIS PAPER

Symbols	Description
n	The number of tumor cases
t	The orientation of the ranklet crop
r	The ranklet cropped resolution
θ	The rotation angle of the GLCM
b	The quantization levels of the GLCM
d	The displacement distance in pixels of the GLCM
$RCM_{r,\theta}^t(p, q)$	The (p, q) entry of the ranklet co-occurrence matrix with specified resolution r , orientation t and rotation angle θ
$\mu_{x,r,\theta}^t$	$\sum_{i=1}^b \sum_{j=1}^b i \cdot RCM_{r,\theta}^t(i, j)$
$\mu_{y,r,\theta}^t$	$\sum_{i=1}^b \sum_{j=1}^b j \cdot RCM_{r,\theta}^t(i, j)$
$\sigma_{x,r,\theta}^t$	$\sum_{i=1}^b \sum_{j=1}^b (i - \mu_{x,r,\theta}^t)^2 \cdot RCM_{r,\theta}^t(i, j)$
$\sigma_{y,r,\theta}^t$	$\sum_{i=1}^b \sum_{j=1}^b (j - \mu_{y,r,\theta}^t)^2 \cdot RCM_{r,\theta}^t(i, j)$