Computers in Biology and Medicine 72 (2016) 43–53

Contents lists available at ScienceDirect

Computers in Biology and Medicine

journal homepage: www.elsevier.com/locate/cbm

Breast mass classification on mammograms using radial local ternary patterns

Chisako Muramatsu [a,*], Takeshi Hara [a], Tokiko Endo [b,c], Hiroshi Fujita [a]

[a] Department of Intelligent Image Information, Graduate School of Medicine, Gifu University, 1-1 Yanagido, Gifu 501-1194, Japan
[b] Department of Advanced Diagnosis, Nagoya Medical Center, 4-1-1 Nakaku Sannomaru, Nagoya, Aichi 460-0001, Japan
[c] Department of Breast Surgery, Higashi Nagoya National Hospital, 5-101 Umemorizaka, Meitoku, Nagoya, Aichi 465-8620, Japan

ARTICLE INFO

Article history:
Received 23 October 2015
Received in revised form
7 March 2016
Accepted 15 March 2016

Keywords:
Local binary patterns
Local ternary patterns
Texture feature
Breast masses
Mammograms
Classification

ABSTRACT

Textural features can be useful in differentiating between benign and malignant breast lesions on mammograms. Unlike previous computerized schemes, which relied largely on shape and margin features based on manual contours of masses, textural features can be determined from regions of interest (ROIs) without precise lesion segmentation. In this study, therefore, we investigated an ROI-based feature, namely, radial local ternary patterns (RLTP), which takes into account the direction of edge patterns with respect to the center of masses for classification of ROIs for benign and malignant masses. Using an artificial neural network (ANN), support vector machine (SVM) and random forest (RF) classifiers, the classification abilities of RLTP were compared with those of the regular local ternary patterns (LTP), rotation invariant uniform (RIU2) LTP, texture features based on the gray level co-occurrence matrix (GLCM), and wavelet features. The performance was evaluated with 376 ROIs including 181 malignant and 195 benign masses. The highest areas under the receiver operating characteristic curves among three classifiers were 0.90, 0.77, 0.78, 0.86, and 0.83 for RLTP, LTP, RIU2-LTP, GLCM, and wavelet features, respectively. The results indicate the usefulness of the proposed texture features for distinguishing between benign and malignant lesions and the superiority of the radial patterns compared with the conventional rotation invariant patterns.

1. Introduction

Breast cancer is the most frequent cancer and constitute more than 20% of all but skin cancers in women worldwide [1]. Early detection is the key to reduce the number of cancer deaths and to improve patients' quality of lives. Mammography is considered an effective screening method for women with normal risk [2–4]. It is not easy, however, to read a large number of mammograms accurately and consistently in a limited time. It is known that about 30% of cancers are missed on mammograms and the reported positive biopsy rates range from 12% to 46% [5–9]. Even in multimodality reading, it is important to assess images of each modality independently and thoroughly. Studies have suggested that the computer-aided detection and diagnosis (CAD) can contribute to accurate diagnosis of mammograms [10–15].

Computerized detection of microcalcifications on mammograms has very high accuracy. On the other hand, computerized classification of malignant and benign lesions still has some room for improvement. A number of studies investigating computerized methods for differentiating between malignant and benign masses have been proposed [16,17]. Tan et al. [18], in their recent study, investigated a variety of different types of image features for classification of breast masses on mammograms. They found that the features related to mass shape, isodensity, and presence of fat were most frequently selected by their feature selection algorithm in a tenfold cross validation scheme. The reliability of these features depends on the accurate determination of mass contours. They also discussed the difficulty of accurate determination of spiculation features due to tissue overlap.

We have been investigating the similar-image retrieval method for assisting the radiologists' classification of breast lesions on mammograms [19–23]. Our proposed similarity measures coupled with machine learning method correlate well with the radiologists' subjective similarity for breast lesions, and the potential utility of the reference images was indicated by the observer studies. For determination of similarity of breast masses, the shape irregularity and margin characteristic features were found to be useful, suggesting the importance of mass outlines which were manually determined in our studies.

* Corresponding author.
E-mail address: chisa@fjt.info.gifu-u.ac.jp (C. Muramatsu).

http://dx.doi.org/10.1016/j.compbiomed.2016.03.007
0010-4825/

In clinical practice, it is not practical to obtain precise outlines of masses owing to busy routine work. In addition, manual outlines provided by different users would vary. Such a "personalized" system may be suitable for an image retrieval system, wherein preferred reference images that a particular user consider similar are selected. However, for a system providing the likelihood of malignancy of lesions, it is undesirable. Automatic segmentation of masses can provide consistent outputs but is not easy when tissues overlap with masses. Therefore, the objective of this study is to propose a method that does not require precise outlines of masses.

Recently several groups have proposed mass classification methods based on wavelets and texture features obtained from regions of interest (ROIs) or small patches [24–30]. Tan et al. proposed a computerized scheme to classify mammographic cases into cancer or benign cases on the basis of image features obtained from whole breast regions of four-view mammograms [24]. A large number of image features, including pixel value statistical features, cumulative projection histogram features, and textural features based on gray level run length and gray level co-occurrence matrix (GLCM), were considered. Although the classification accuracy in terms of the area under the receiver operating characteristic curve (AUC) was not very high, they concluded that such a scheme could become a supplemental system by potentially providing information different from the conventional lesion-based schemes. Eltoukhy et al. compared methods based on wavelet and curvelet transforms for classification of malignant, benign and normal ROIs [25]. Using the 100 largest coefficients from each of 4 decomposition levels as a feature vector, they obtained a high classification accuracy with a relatively small dataset with mixed abnormalities including masses, microcalcifications, architectural distortions, and asymmetries. do Nascimento et al. investigated mass classification methods using discrete wavelet transform (DWT) combined with different classifiers [26]. A large number of wavelet coefficients obtained from each decomposition level was reduced by the singular value decomposition to constitute a feature vector. They obtained a high classification performance using a polynomial classifier compared with the support vector machine, decision tree, and k-nearest neighbor classifiers.

Ergin et al. proposed a 3-class classification method, i.e., classification of normal, benign, and malignant cases, based on DWT and statistical features [27]. DWT was first applied, and features were extracted from the subbands using the histogram of oriented gradients, dense scale invariant feature transform, and local configuration pattern, which is local binary pattern (LBP) [28] combined with variance. Lahmiri et al. proposed a hybrid method combining DWT and Gabor filter for classification of medical images including breast cancer and normal mammograms [29]. Gabor filters with 4 directions and 3 scales were applied to the high frequency component of the wavelet transformed image, and the entropy and uniformity statistics were used as the feature vector. Reyad et al. also investigated different ROI-based features, including statistical, LBP, and multi-resolution analysis features, for classification of cancer and normal mammograms [30].

Chakraborty et al. proposed an angle co-occurrence matrices (ACM), which is a combination of Sobel filter and GLCM, for classification of benign and malignant masses [31,32]. ACM is derived using gradient magnitude and orientation to describe edge patterns. Their method, however, requires the mass segmentation, and ACM was computed in margin area using rubber-band straightening transform [33]. Other groups have also used LBP for false positive reduction in computerized mass detection [34,35]. The results from these studies indicate that the edge characteristics represented by the various features are useful in classification of mammograms. In some of these studies, edge orientation was not considered. For example, in [25], the largest coefficients were selected regardless of the pixel location, and in [27,35] rotation invariant features were employed. Other studies showed that the classification accuracy was decreased by use of rotation invariant features [30], and edge orientation is an important characteristic for lesion classification [31,32].

One of the major characteristics of malignant masses is spiculation. While benign masses have round or oval shapes with clear margin, malignant masses with spicula have jagged edges. Therefore, for the classification of benign and malignant lesions, edge orientation with respect to the center of a mass is important. In our preliminary study [36], we proposed a radial local ternary pattern (RLTP) and tested its potential utility with a small dataset. In that study, small sub ROIs (patches) were selected at characteristic points (close by mass outlines) for determination of RLTP, and the effect of different parameters and its superiority to regular LBP and local ternary pattern [37] were not investigated. In this study, we examined the usefulness of RLTP, performed extensive experiments to study the effect of different parameters, and compared with GLCM and DWT-based features for classification of benign and malignant masses on mammograms. The remainder of this paper is organized as follows. The image database is described in Section 2. The feature extraction and classification methods are described in Section 3. The experimental results and discussion are presented in Sections 4 and 5, respectively, and Section 6 concludes the paper.

2. Image database

Digital mammograms were obtained from Nagoya Medical Center, Nagoya, Japan, using one of the following systems: phase contrast mammography (PCM) system (Mermaid or Pureview, Konica Minolta Holdings, Inc.), direct conversion digital mammography system (Amulet, Fujifilm Corporation), and computed radiography (CR) systems (Mammomat 3000, Siemens, with C-Plate, Konica Minolta, or Profect, Fujifilm). The pixel sizes of the original images are 25 μm (PCM), 43.72 μm (C-Plate), and 50 μm (Amulet and Profect), and the grayscales are 10 bits (Profect), 12 bits (PCM and C-Plate), and 14 bits (Amulet). The study was approved by the institutional review board.

We employed the dataset used in the previous study [23]. For that study, mammograms with a history of biopsies were consecutively collected, and benign cases with follow-up ultrasound examinations were added from the same period. The total number of cases was 512, including the findings of microcalcifications and distortion. Images were retrospectively reviewed by radiologists, and square ROIs were obtained for mass lesions. The original radiologic and pathologic reports were available if needed. ROIs were extracted from both craniocaudal and mediolateral oblique views if the entire lesion was visible. In the previous study, masses with 9 pathologic types were used [23]. The database consists of 376 ROIs, including 195 benign and 181 malignant masses. The malignant cases were confirmed by biopsy or surgery, and benign cases were confirmed through biopsy or follow-ups by mammography and ultrasonography. For image analysis, the pixel size and grayscale of ROIs were unified to 50 μm and 10 bits, respectively, by linear interpolation. The size of the ROIs varied from 168×168 to 1888×1888 pixels.

3. Methods

3.1. RLTP-based features

LBP is a method to describe an image texture or local intensity information by a binary sequence [28]. The binary patterns are

determined for every pixel in the patch, and the histogram of the patterns is used as features for classification of the patch. The number of possible patterns, i.e., the number of the histogram bins, is 2^P for LBP, where P is the number of neighbor pixels to be compared with a pixel of interest. Ojala et al. suggested that a majority of the meaningful patterns can be represented by the fundamental "uniform" patterns and all the other patterns can be grouped into the "non-uniform" pattern, which greatly reduces the number of possible patterns to $P(P-1)+3$ [28].

LBP was originally proposed for textural pattern classification. For this purpose, rotation invariant LBP was designed to obtain the same feature value when an image is rotated [28]. However, as described earlier, for classification of benign and malignant breast masses, the edge orientation at the margin is an important characteristic. Therefore, we propose a rotation that a binary cord is shifted with respect to the mass center and defined by

$$RLBP_{P,R} = \left\{ ROR(LBP_{P,R}, s) \middle| s = \left\| \theta \cdot \frac{P}{2\pi} \right\| \right\} \quad (1)$$

$$\theta = \tan^{-1} \frac{y_m - y_c}{x_m - x_c} \quad (2)$$

where $ROR(LBP_{P,R}, s)$ performs a circular bit-wise right shift on the P-bit number s times, $\|x\|$ gives the integer nearest to x, and (x_c, y_c) and (x_m, y_m) are the coordinates of the pixel of interest and the pixel of mass center, respectively. In this study, we assumed that the center of a mass is located at the center of ROI. By applying the Eqs. (1) and (2), the binary sequence always began from the pixel closest to the center of a mass. Fig. 1 illustrates the coding of LBP and the relationships with variant LBPs, i.e., rotation invariant LBP (RI-LBP), RI-LTP, and RLTP, which is described below. We also investigated whether we could reduce the number of histogram bins by combining some patterns. After the sequence was rotated and aligned, P patterns were reduced to 4 patterns by combining the patterns corresponding to the edge in the inward, outward, clockwise, and counterclockwise directions, as shown in Fig. 2. In this figure, examples of "uniform" patterns when $P=8$ are shown, assuming that the shaded pixels correspond to the one closest to the mass center.

Since actual images generally contain some noise, LTP was proposed [37] to include a threshold, t. The obtained ternary code is split into 2 sets of binary codes corresponding to the positive and negative halves. In this study, the two histograms corresponding to the positive and negative patterns are combined by summation of corresponding bins. The pattern histograms were treated as the feature vectors and passed to classifiers.

Ojala et al. also suggested the inclusion of a local variance measure determined using P pixels [28]. In general, estimation of variance becomes more accurate by increasing the sample. Therefore, in this study, variance was determined in an area

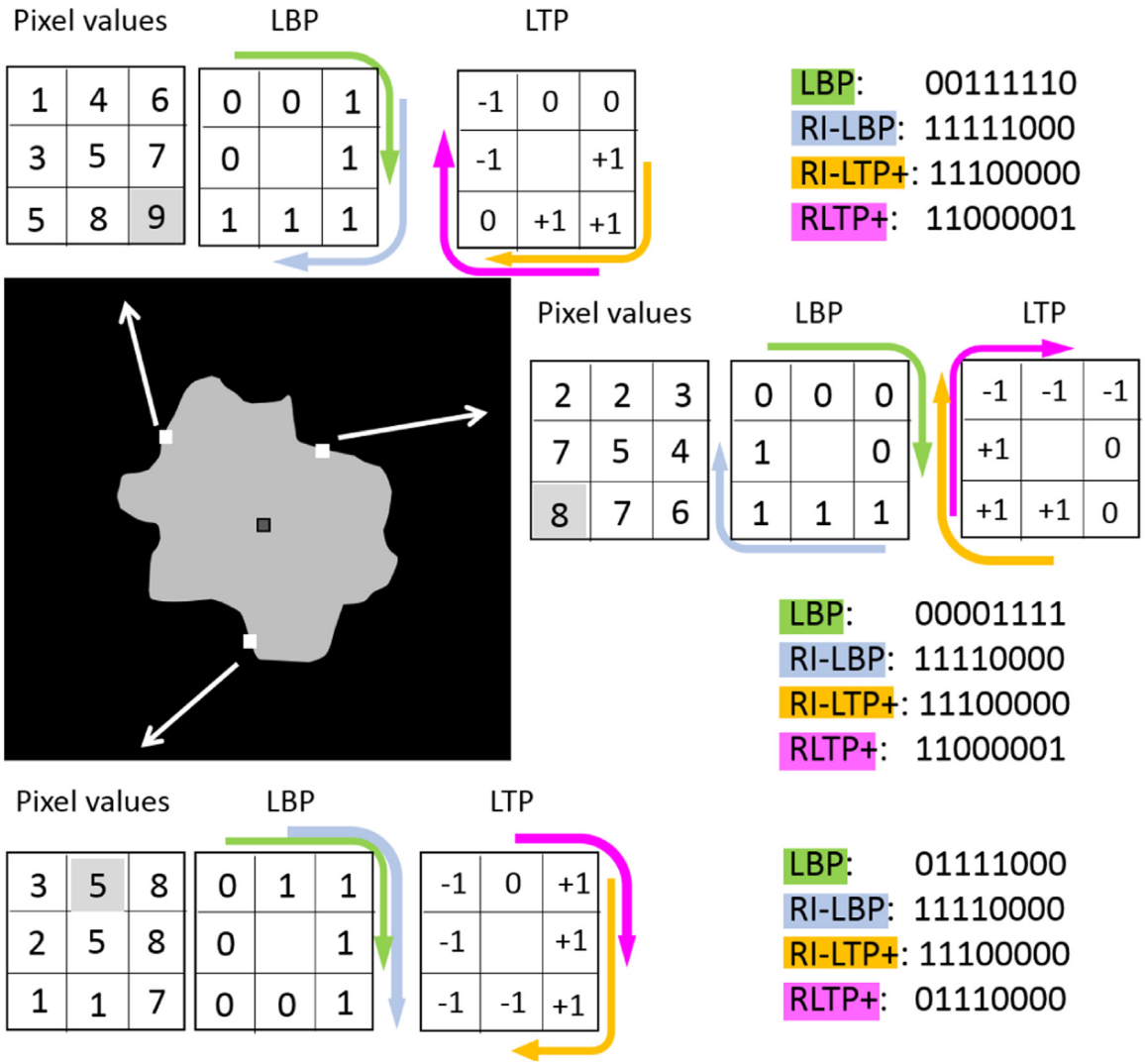

Fig. 1. Illustration of differences between LBP, RI-LBP, RI-LTP, and RLTP.

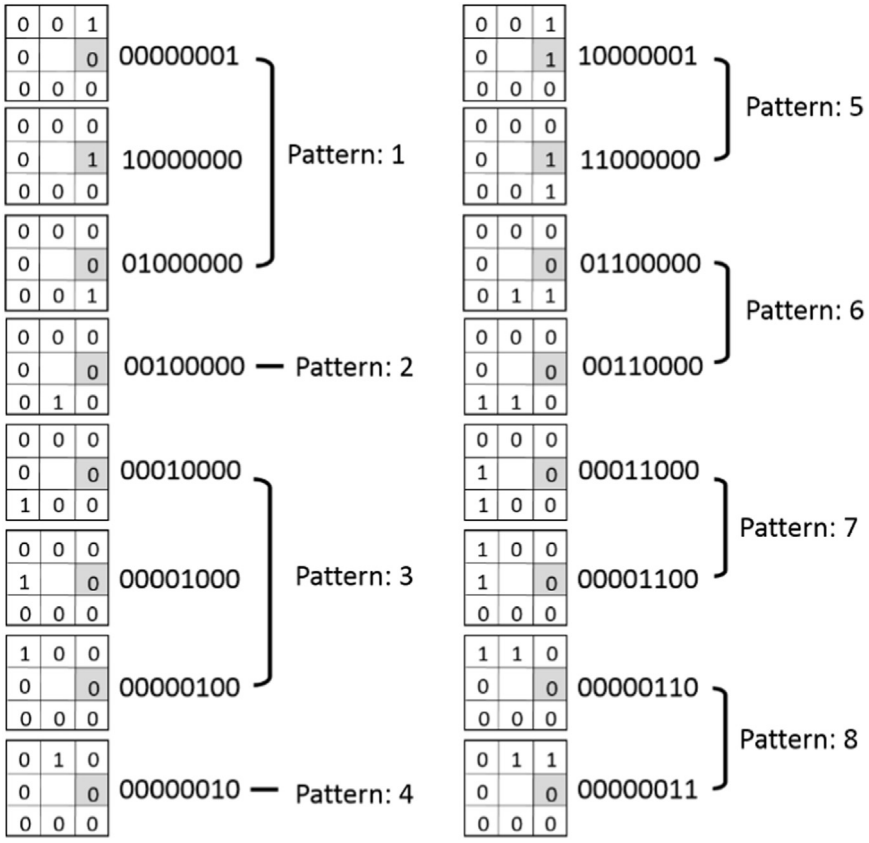

Fig. 2. Histogram bin reduction by combination of radial and tangential patterns.

instead of P sample points and defined as

$$VAR_R = \frac{1}{(2R+1)^2} \sum_{i=x_c-R}^{x_c+R} \sum_{j=y_c-R}^{y_c+R} \left(g_{i,j} - \mu\right)^2 \quad (3)$$

$$\mu = \frac{1}{(2R+1)^2} \sum_{i=x_c-R}^{x_c+R} \sum_{j=y_c-R}^{y_c+R} g_{i,j} \quad (4)$$

where $g_{i,j}$ is the pixel value at coordinate (i, j). As in [28], continuous values of VAR measures were quantized to B bins based on the percentile of combined data. The VAR histogram was concatenated to the LTP histogram and passed to the classifiers.

3.2. GLCM-based features

GLCM is another method that is frequently used for texture analysis. The matrix is computed to describe the distribution of pixel values in relation to the relative position of the pixels. The statistics of the matrix are employed as the texture features. In this study, Haralick's textural features [38] were determined and their ability for the classification of breast images was compared with that of RLTP. The feature set includes the angular second moment, contrast, correlation, sum of squares, inverse difference moment, sum average, sum variance, sum entropy, entropy, difference variance, difference entropy, two information measures of correlation, and dissimilarity. The definitions of these features are given in [38] and [39].

Four GLCMs were computed for the angles of 0°, 45°, 90° and 135°. In this study, we proposed RLTP to make binary pattern radially aligned. To obtain radially aligned GLCM, the ROIs were spread open by transforming the images to the polar coordinate system with respect to the mass center, so that the rows and columns of the image correspond to the radial and angular directions, respectively, as shown in Fig. 3. The GLCMs and corresponding textural features were determined with the original and the transformed images for comparison. In the transformed images, the pixels in the original ROIs were only used for the computation of the GLCMs. The features determined with the different angular matrices were handled separately constituting a 56-dimensional (14×4) feature vector, and also represented by the average and maximal values as a 28-dimensional (14×2) feature vector.

3.3. DWT-based features

DWT can also be used to represent image texture and has been increasingly used for image classification. In this study, a multilevel two-dimensional DWT was applied to the ROIs using Daubechies-4 function. As a result, an image is decomposed into low frequency subbands (LL), horizontal high frequency subbands (HL), vertical high frequency subbands (LH), and diagonal high frequency subbands (HH). The wavelet coefficients were treated as the features. For reducing the number of features, uses of the coefficients from the diagonal subbands (LL and HH) and the highest 100 coefficients from each subband were tested.

3.4. Classification

The classification of benign and malignant ROIs was performed using an artificial neural network (ANN), support vector machine (SVM) and random forest (RF) classifiers. For ANN, the three layered feedforward network with the backpropagation algorithm was used. The numbers of hidden units and training epochs were optimized experimentally. For SVM, linear, polynomial, radial, and sigmoid basis functions were tested with selected feature patterns. Overall, use of the linear basis function provided the best AUCs,

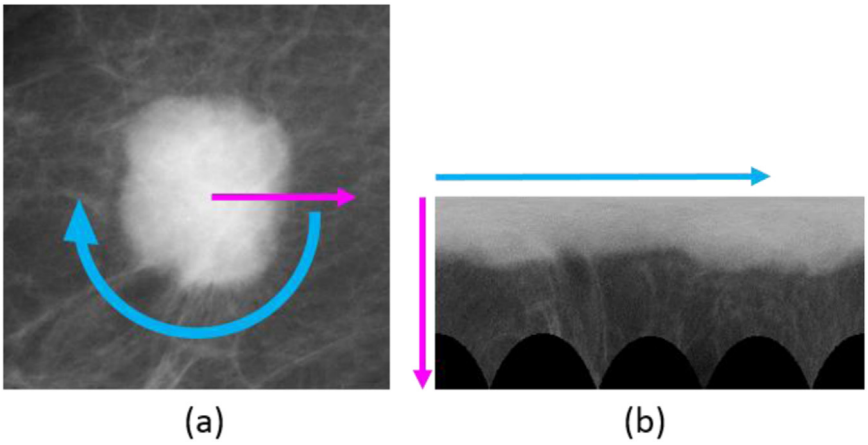

Fig. 3. Image transformation for GLCM computation.

and therefore, the linear kernel was employed in this study. For RF, the number of trees was set to 500 and the number of variables randomly sampled as candidates at each split was square root of the number of features. In order to reduce the high dimensions of feature vectors, principal component analysis (PCA) was applied, and the optimal number of input units (principal components, PCs) was selected by the experiment. Because of the relative robustness to the large input features and for providing a large number of variables for decision trees, the original feature vectors were also tested for RF. Leave-one-out cross validation was employed for evaluation, in which at each fold all the ROIs that belong to the same patient were removed as a test set and the remaining ROIs were used for training.

The ability of the proposed RLTP-based features and other features for classification between benign and malignant masses were evaluated using the area under the receiver operating characteristic (ROC) curve (AUC).

4. Experimental results

4.1. LTP parameters

For reducing the processing time and possibly reducing the effect of noise, the images were down sized by a factor of 2 in both directions by averaging. We also investigated the effect of image normalization. The images were subsampled or magnified by linear interpolation so that all ROIs have the same matrix size. By normalizing the ROI size, the effect of difference in lesion sizes may be reduced. The effect of grayscale reduction was also examined. Other parameters were kept the same, i.e., $R=3$, $P=8$, and the bin reduction to 4, for this comparison. The results of AUCs showed that the classification performance was improved because of the size reduction by 2 while the effect of bit reduction was small as shown in Table 1. Note that when the grayscale was reduced, the threshold value for LTP was also changed accordingly by a factor of 4. By image normalization, classification ability was generally improved.

Using the normalized ROIs, optimal R and P parameters were investigated. Combinations of ROI size and R parameter were tested as the optimal R may change with the ROI size. The results are shown in Table 2. For these comparison, $P=8$ is employed without and with bin reduction. In this study, using the ANN, ROI size of 200×200 pixels, $R=3$, and $P=8$ with the bin reduction to 4 directions provided the best result in terms of AUC. With these parameters, the number of patterns was 31, and the use of first 20 PCs provided the highest AUC. By using the SVM, ROI size of 250×250 pixels, $R=5$, and $P=8$ with the bin reduction to 4 directions provided the best AUC, whereas the AUC for ROI size of 300×300 pixels, $R=7$, and $P=8$ without the bin reduction was the best using the RF. For these results, the numbers of PCs employed were 19 and 16 for the SVM and RF, respectively. The results indicate that a small difference in ROI size does not strongly affect the result when an appropriate R is used for the ROI

Table 1
The effects of downsizing and grayscale reduction on AUC.

ROI size	Number of grayscale	AUC		
		ANN	SVM	RF
Original (50 μm/pixel)	8 bits	0.752	0.696	0.757
1/4 of original	8 bits	0.798	0.753	0.814
150 × 150	8 bits	0.890	0.874	0.843
200 × 200	**8 bits**	**0.900**	0.881	0.842
250 × 250	8 bits	0.882	0.874	0.825
300 × 300	8 bits	0.875	0.849	0.801
200 × 200	10 bits	0.884	0.877	0.842

Table 2
The effects of ROI size and R parameter with and without bin reduction.

ROI size	R	AUC					
		Bin reduction			No bin reduction		
		ANN	SVM	RF	ANN	SVM	RF
100 × 100	1	0.831	0.824	0.779	0.814	0.811	0.775
100 × 100	2	0.861	0.855	0.814	0.856	0.845	0.825
100 × 100	3	0.875	0.873	0.857	0.868	0.874	0.860
150 × 150	2	0.857	0.836	0.805	0.836	0.826	0.786
150 × 150	3	0.890	0.874	0.843	0.872	0.868	0.852
150 × 150	4	0.876	0.877	0.854	0.879	0.878	0.866
200 × 200	2	0.829	0.817	0.771	0.823	0.808	0.786
200 × 200	**3**	**0.900**	0.881	0.842	0.893	0.883	0.855
200 × 200	4	0.891	0.886	0.854	0.882	0.871	0.853
200 × 200	5	0.882	0.881	0.860	0.875	0.874	0.872
250 × 250	3	0.882	0.874	0.825	0.882	0.861	0.842
250 × 250	4	0.892	0.876	0.865	0.886	0.876	0.869
250 × 250	**5**	0.897	**0.895**	0.864	0.884	0.885	0.871
250 × 250	6	0.895	0.886	0.859	0.883	0.883	0.873
300 × 300	4	0.883	0.872	0.835	0.872	0.862	0.838
300 × 300	5	0.887	0.877	0.853	0.881	0.878	0.847
300 × 300	6	0.893	0.890	0.862	0.883	0.878	0.879
300 × 300	**7**	0.894	0.892	0.870	0.883	0.873	**0.881**
350 × 350	4	0.878	0.869	0.836	0.881	0.856	0.841
350 × 350	5	0.896	0.887	0.858	0.881	0.867	0.846
350 × 350	6	0.895	0.887	0.861	0.885	0.873	0.869
350 × 350	7	0.892	0.884	0.851	0.885	0.881	0.867

size. The classification performance was degraded when the ROI size was too small and/or R was not suitable for the ROI size. Table 3 presents the results for different P parameter. Although AUCs were lower for $P=4$ than for $P=8$, the results for $P=8$ with bin reduction were comparable with those without bin reduction. These results indicate that the bin reduction is effective in reducing the number of features without degrading the classification performance. Using ANN and SVM, comparable results were obtained with $P=16$; however, the calculation time was more than 100 times longer than that for $P=8$.

In this study, we proposed RLTP to take into account the orientation of edge patterns with respect to the mass center. Table 4 shows the comparison of the results using LTP, uniform LTP (U2-LTP), RI-LTP, RIU2-LTP, RLBP, and RLTP. For each classifier and LBP/LTP variant, the ROI size and other parameters were optimized and the best result was presented in the table. The ROC curves for these LTPs and LBP using the ANN, SVM, and RF are shown in Fig. 4. Table 5 presents the p values for statistical comparison of these results. With all three classifiers, the AUCs for RLTP are statistically significantly higher than those for other LTPs ($p < 0.01$ for all by ROCKIT v1.1, The University of Chicago [40]) and RLBP, except for the case using ANN ($p=0.06$). The results indicate the potential utility of the proposed RLTP for distinguishing between benign and malignant masses. Fig. 5 shows examples of RLTP images of benign and malignant masses representing the RLTP pattern of each pixel. Darker pixels correspond to the edges in radial direction, while lighter pixels correspond to the edges in direction perpendicular to the radial direction, some of which correspond to spicula. These images show the difference in edge patterns around the mass margin between benign and malignant masses. The corresponding RLTP histograms are also included. The histogram bins 2 to 8, 9 to 15, 16 to 22, and 23 to 29 correspond to inward edges, outward edges, clockwise edges, and counter clockwise edges, respectively. In benign masses, frequencies of radial edges are larger, whereas in malignant images frequencies of non-radial edges are larger.

By including the variance features, the results were almost unaffected in terms of AUC. Fig. 6 shows the variance images of the same ROIs in Fig. 5. For these examples, the variances were quantized to 10 bins. These images show that the variances are larger in malignant ROIs. The AUC by the variance alone using the ANN was 0.742. The results suggest that the variance histogram may be useful in distinguishing between benign and malignant ROIs. However, when combined with the RLTP histogram, the contribution of the variance was minimal in this study.

Table 3
The effect of P parameter.

ROI size	R	P	Bin reduction	Number of bins	AUC		
					ANN	SVM	RF
200×200	3	4	No	15	0.853	0.845	0.807
200×200	3	8	Yes	31	0.900	0.881	0.842
200×200	3	16	Yes	63	0.881	0.859	0.822
200×200	3	16	No	243	0.863	0.852	0.798
250×250	5	4	No	15	0.859	0.862	0.822
250×250	5	8	Yes	31	0.897	0.895	0.864
250×250	5	16	Yes	63	0.887	0.878	0.852
250×250	5	16	No	243	0.883	0.871	0.853
300×300	7	4	No	15	0.861	0.861	0.832
300×300	7	8	Yes	31	0.894	0.892	0.870
300×300	7	16	Yes	63	0.900	0.895	0.872
300×300	7	16	No	243	0.897	0.874	0.868

Table 4
The result for different LBP/LTP patterns.

Pattern	Number of patterns	AUC		
		ANN	SVM	RF
LTP	256	0.773	0.765	0.712
U2-LTP	59	0.766	0.783	0.759
RI-LTP	36	0.768	0.763	0.727
RIU2-LTP	10	0.780	0.778	0.757
RLBP	31/59	0.875	0.856	0.833
RLTP	**31/59**	**0.900**	**0.895**	**0.881**

*RLBP: radial LBP, U2-LTP: uniform LTP, RI-LTP: rotation invariant LTP, RIU2-LTP: rotation invariant uniform LTP, and RLTP: radial LTP

4.2. Comparison with other features

For the computation of GLCM, the number of gray levels was reduced to 5, 6, 7 or 8 bits, and the distance parameter, d, of 1, 2, 3, and 4 was tested. The AUC results are shown in Fig. 7 and the best AUCs are summarized in Table 6. As shown in the figure and table, different classifiers provided the best AUCs for different combination of ROI size and d parameter; however, the difference in AUCs was small among tested parameters except for the bit reduction to 5 bits. For LTP-based features, higher AUCs were obtained with the RLTP and RLBP than with the regular LTP and RI-LTP. Similarly, AUCs were higher for the features determined in the polar transformed images than in the original images (Table 6). These results support the importance of the texture pattern orientation for distinguishing between benign and malignant masses. Classification performance was higher for the RLTP-based features compared with that of the GLCM-based features.

Table 7 shows the classification performances using the wavelet coefficients for different subbands and composition levels. The highest AUC was obtained using the highest 100 coefficients from each subband of 3-level decomposition with ROIs of 128×128 pixels by ANN, although comparable AUCs were obtained by SVM and RF. By selecting the highest coefficients, the most characteristic parts (edges) are reflected in the feature disregarding the location. The process is somewhat similar to shifting the binary code. The result would be the same when the image is rotated. Therefore, this process can be considered as achieving rotation invariance. However, it does not take into account the edge orientation with respect to the mass center as with RLTP and GLCM in polar transformed images. It might be a possible reason for the lower AUCs compared with those of RLTP and GLCM-based features.

5. Discussion

We investigated the ROI-based features that do not require precise segmentation of lesions for classification between benign and malignant masses on mammograms. The proposed RLTP-based features obtained a high classification performance of 0.90 in AUC. The classification result may be useful for assisting radiologists' diagnoses of breast lesions with a minimal input of placing a square ROI. The system can provide consistent results for lesions overlapped with fibrograndular tissue where automated segmentation is challenging. The ROI-based features may supplement the conventional shape based features.

In pattern classification problems, it would be desirable that an image is classified to the same group when it is rotated. For facilitating such classification, rotation invariant features were proposed [28]. However, in distinguishing between benign and malignant masses, edge orientation with respect to the direction toward the mass center is important to characterize circumscribed and spiculated margins. In this study, we proposed RLTP that not

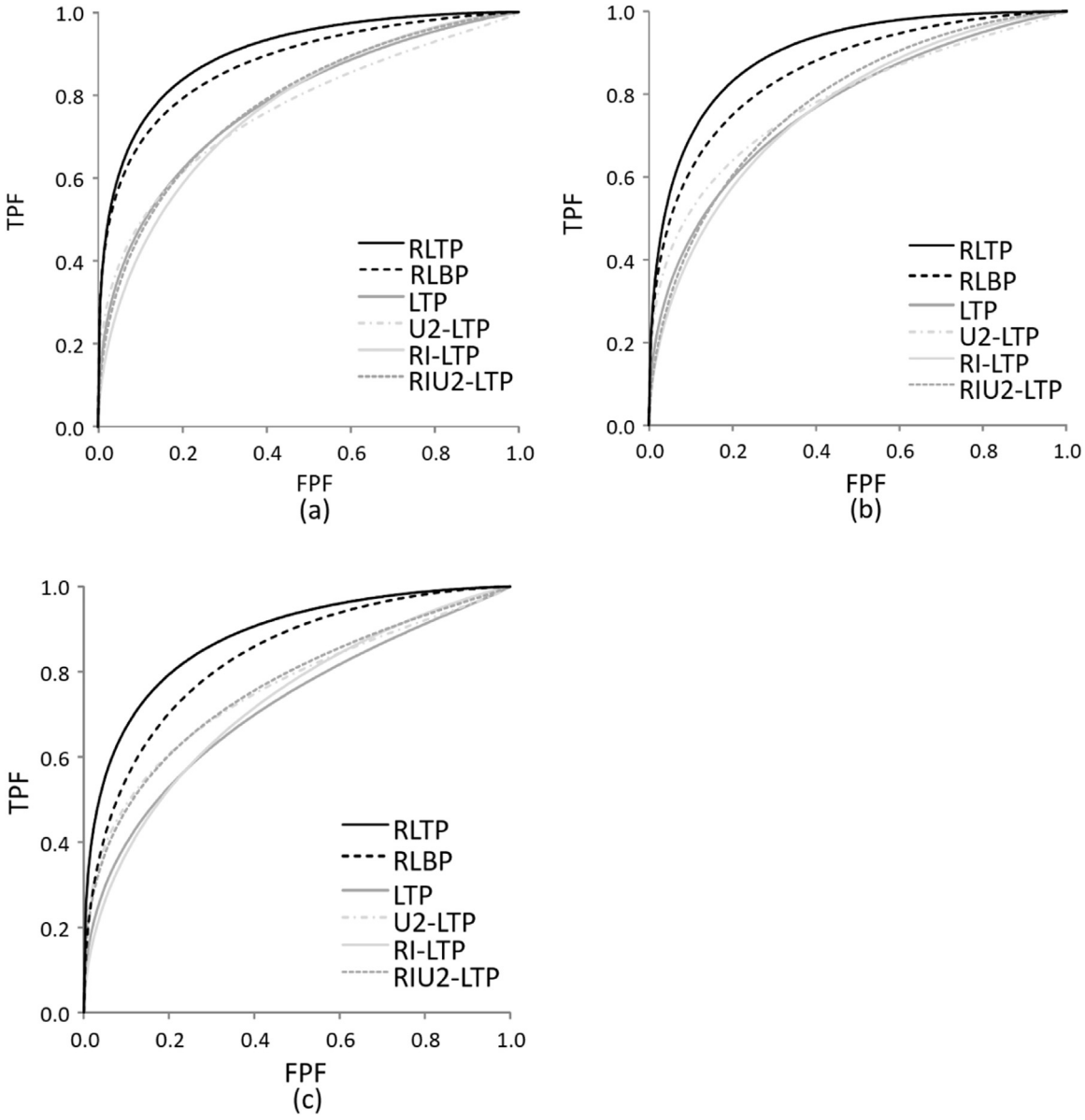

Fig. 4. ROC curves of classification results for different types of LBPs. The results by (a) ANN, (b) SVM, and (c) RF.

Table 5
P-values for difference in AUCs for different LBP/LTP patterns.

	U2-LTP	RI-LTP	RIU2-LTP	RLBP	RLTP
LTP	0.44	0.55	0.95	0.0001	< 0.0001
	0.23	0.84	0.34	0.0009	< 0.0001
	0.03	0.69	0.05	< 0.0001	< 0.0001
U2-LTP		0.94	0.53	< 0.0001	< 0.0001
		0.34	0.89	0.004	< 0.0001
		0.09	0.99	0.004	< 0.0001
RI-LTP			0.47	< 0.0001	< 0.0001
			0.13	0.0006	< 0.0001
			0.06	0.0001	< 0.0001
RIU2-LTP				0.0001	< 0.0001
				0.008	< 0.0001
				0.007	< 0.0001
RLBP					0.06
					0.003
					0.004

*First, second, and third rows in each cell correspond to the *p*-values for comparisons using ANN, SVM, and RF

only takes into account the pattern orientation with respect to the center of masses but also is robust to the image rotation. The best AUCs obtained for RLTP using ANN, SVM and RF were 0.900, 0.895, and 0.881, respectively, whereas the best AUCs obtained for RIU2-LTP were 0.780, 0.778, and 0.757, respectively. The results showed a higher classification performance for RLTP over the RIU2-LTP ($p < 0.0001$). There were no statistical significant differences between AUCs for RLTP obtained using different classifiers ($p = 0.68$ for ANN vs SVM, $p = 0.06$ for ANN vs RF, and $p = 0.13$ for SVM vs RF).

In this study, classification performance for RLTP-based features (AUC = 0.90) was superior to those for GLCM-based features (AUC = 0.86; $p = 0.01$) and DWT-based features (AUC = 0.83; $p = 0.0001$). In the literature, there have been several studies indicating the utility of DWT and other multiresolutional transforms for classification of mammograms [25–27,29,30]. Some of them showed very high classification performance. The results of these studies cannot be directly compared with those in this study because of the different datasets used. With DWT, there seems to

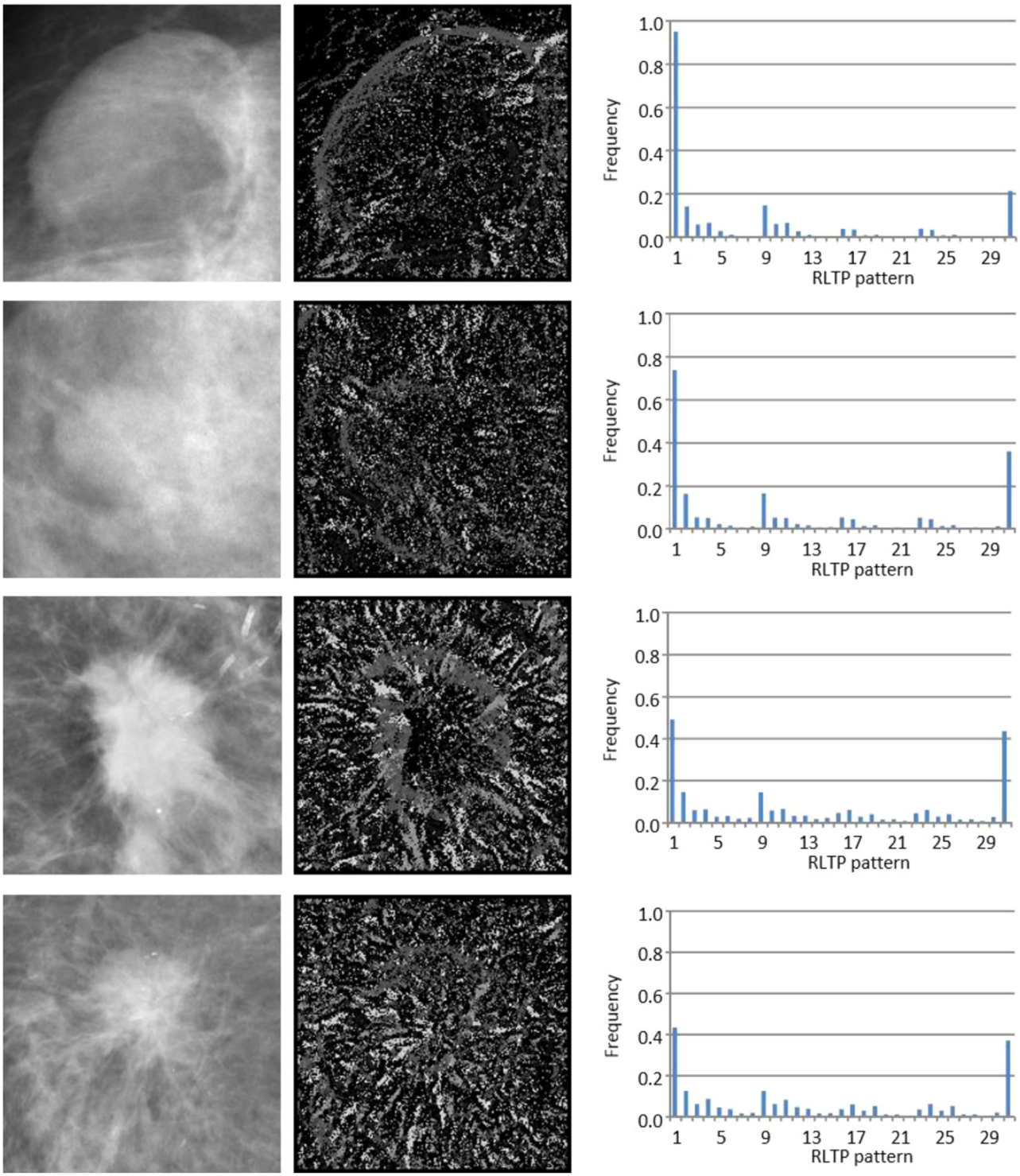

Fig. 5. RLTP images for benign (top two rows) and malignant (bottom two rows) masses and the corresponding RLTP histograms.

be no concrete consensus as to which wavelet function is used, which subbands are used and how it is used, e.g., as preprocessing or features. In this study, we applied one of the popular functions, Daubechies-4, and tested the use of the coefficients directly as features for mass classification. Because of the large number of features, the highest coefficients were selected before applying PCA; it may have a similar effect as obtaining the rotation invariant DWT features. The AUC was slightly higher than RI-LTP and GLCM in Cartesian coordinate images but slightly lower than RLTP and GLCM in polar transformed images. The results may be improved by use of other classifiers; however, exploration of the suitable classifier is beyond the scope of this study. There may be other effective ways of using DWT, such as the preprocessing approach as in [29]. A combination of DWT and RLTP and GLCM-based features should be investigated in the future.

Although we performed the test using a leave-one-case-out cross validation, the training and test datasets are not completely independent. The best parameters and the number of PCs were determined based on the cross validation results. The effectiveness of the proposed method and the optimal parameters must be validated with the independent dataset.

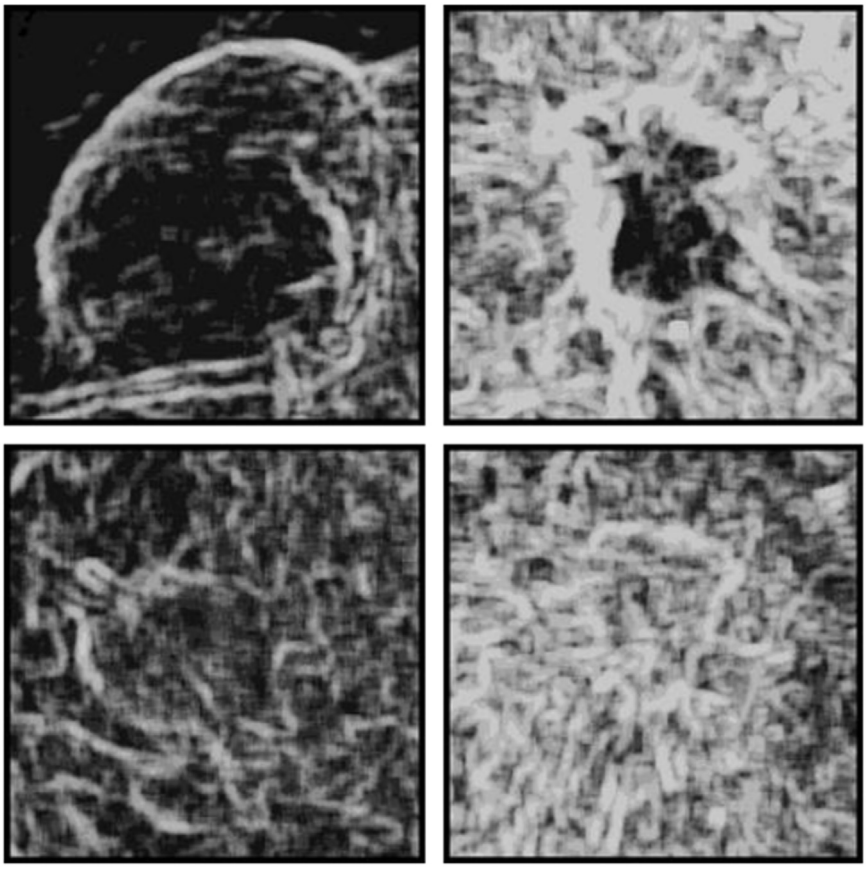

Fig. 6. Variance images for the ROIs in Fig. 5.

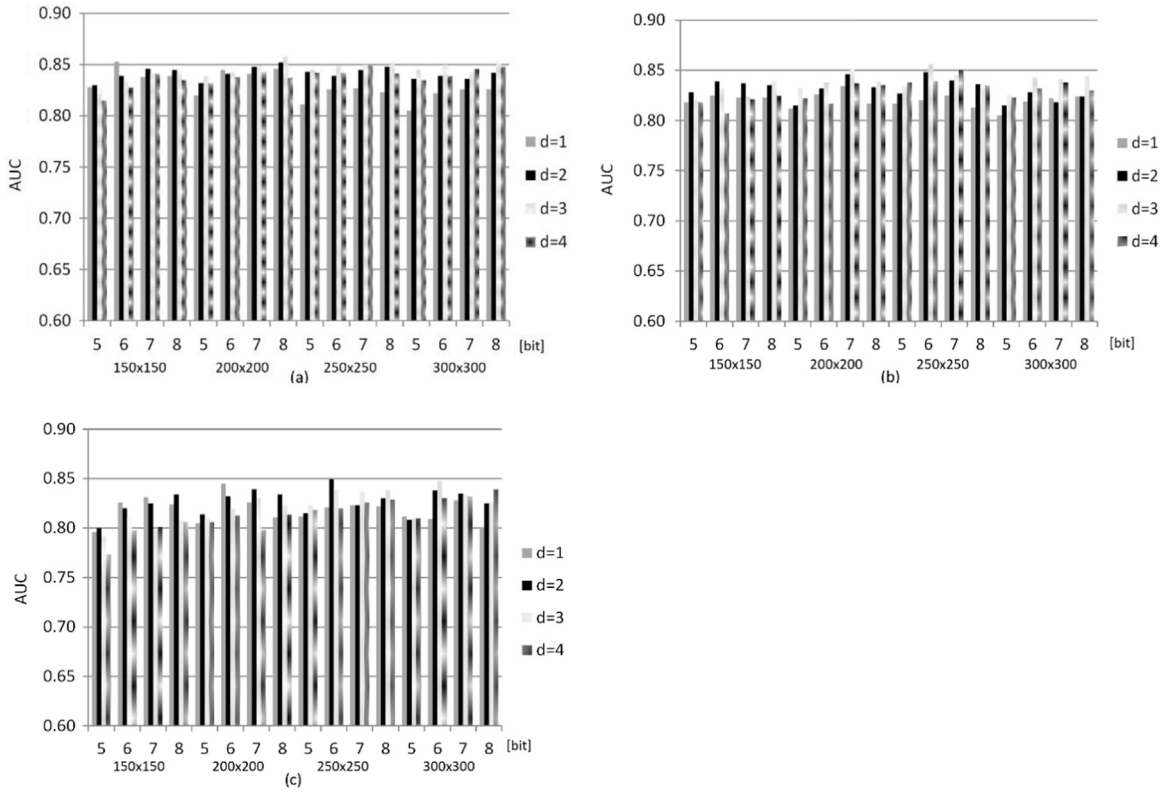

Fig. 7. AUC results for GLCM with different parameters using (a) ANN, (b) SVM, and (c) RF.

Table 6
The best AUC for each ROI size and classifier for GLCM-based features.

ROI size	Polar coordinate			Cartesian coordinate		
	ANN	SVM	RF	ANN	SVM	RF
150 × 150	0.853	0.839	0.834	0.773	0.753	0.764
200 × 200	**0.858**	0.852	0.845	0.769	0.752	0.757
250 × 250	0.852	**0.856**	**0.849**	0.772	0.754	0.764
300 × 300	0.852	0.844	0.848	0.774	0.760	0.773

Table 7
Classification results for DWT-based features.

ROI size	Number of decomposition	Subbands	Number of coefficients	AUC		
				ANN	SVM	RF
64 × 64	1	LL, HH	2048	0.785	0.819	0.762
64 × 64	2	LL, HH	1536	0.751	0.771	0.782
128 × 128	1	LL, HH	8192	0.795	0.806	0.773
128 × 128	2	LL, HH	6144	0.763	0.786	0.775
128 × 128	3	LL, HH	5632	0.679	0.760	0.734
64 × 64	1	100 highests from all	400	0.775	0.766	0.756
64 × 64	2	100 highests from all	700	0.822	0.789	0.770
128 × 128	1	100 highests from all	400	0.797	0.752	0.760
128 × 128	2	100 highests from all	700	0.768	0.750	0.737
128 × 128	**3**	**100 highests from all**	**1000**	**0.827**	**0.822**	0.810
256 × 256	1	100 highests from all	400	0.773	0.743	0.769
256 × 256	2	100 highests from all	700	0.766	0.750	0.762
256 × 256	3	100 highests from all	1000	0.750	0.742	0.739
256 × 256	**4**	**100 highests from all**	**1300**	0.825	0.814	**0.813**

6. Conclusion

ROI-based image features without the need of precise segmentation of lesions may be useful as a part of CAD system for diagnosis of mammograms. The experimental results showed the high classification ability of the proposed RLTP for ROIs with benign and malignant masses. RLTP classification could provide supplemental information to the conventional contour-based features. The utility of the proposed features should be further evaluated with a large independent database.

Conflict of interest

Tokiko Endo: Institutional research support, FUJIFILM Holdings Corporation

Others: Nothing to disclose

Summary

Texture features are useful for pattern recognition and classification. Unlike conventional features for classification of lesions, such as shape and contrast features, determination of texture features does not require the precise segmentation of the lesions. Therefore, such region of interest (ROI)-based features can be advantageous where manual segmentation is time-consuming and accurate automatic segmentation could be difficult by the background. In general, an image is expected to be classified to the same category when it is rotated. Therefore, rotation invariant texture features have been proposed. However, in the classification of lesions on medical images, edge orientation may be an important feature to distinguish between the lesions with well-defined and circumscribed margin and those with spiculated margin. In this study, we investigated the ROI-based features and propose a new variant of local binary patterns (LBP), which takes into account the pattern orientation with respect to the center of lesions and still is robust to image rotation. The usefulness of the proposed features was tested for classification of benign and malignant masses on mammograms and compared with other texture features. Our database consists of the square ROIs including 181 malignant and 195 benign masses. The proposed radial local ternary pattern (RLTP), conventional LTP, rotation invariant (RI) LTP, Haralick's texture features using gray level co-occurrence matrix (GLCM), and wavelet features were determined from the ROIs. After applying the principal component analysis to reduce the number of features, ROIs were classified by use of an artificial neural network, a support vector machine, or random forest classifier. Based on the receiver operating characteristic (ROC) analysis, the areas under the ROC curves for the RLTP, LTP, RIU2-LTP, GLCM, and wavelet features were 0.90, 077, 0.78, 0.86, and 0.83, respectively. The results indicated the superiority of the proposed feature over the conventional rotation invariant features and the potential utility of the feature for classification of breast lesions on mammograms.

Acknowledgment

Authors are grateful to Mikinao Oiwa, M.D., Ph.D., Misaki Shiraiwa, M.D., Ph.D., and Min Zhang, Ph.D. for their contribution in the study. This study was partly supported by the Grants-in-Aid for Scientific Research for Young Scientists (No. 26860399) by Japan Society for the Promotion of Science and Grant-in-Aid for Scientific Research on Innovative Areas (No. 26108005) by Ministry of Education, Culture, Sports, Sciences and Technology in Japan.

References

[1] American Cancer Society, Global Cancer Facts & Figures 2nd Edition, American Cancer Society, Atlanta, 2011.
[2] L. Tabar, G. Fagerberg, S.W. Duffy, N.E. Day, A. Gad, O. Grontoft, Update of the Swedish two-county program of mammographic screening for breast cancer, Radiol. Clin. N. Am. 30 (1992) 187–210.
[3] S. Shapiro, W. Venet, P. Strax, L. Venet, R. Roeser, Selection, follow-up, and analysis in the health insurance plan study: a randomized trial with breast cancer screening, J. Natl. Cancer Inst. Monogr. 67 (1985) 65–74.
[4] L.L. Humphrey, M. Helfand, B.K.S. Chan, S.H. Woolf, Breast cancer screening: a summary of the evidence for the U.S. preventive services task force, Ann. Intern. Med. 137 (2002) E-347-367.
[5] F.M. Hall, J.M. Storella, D.Z. Silverstone, G. Wyshak, Nonpalpable breast lesions: Recommendations for biopsy based on suspicion of carcinoma at mammography, Radiology 167 (1988) 353–358.
[6] D.B. Kopans, R.H. Moore, K.A. McCarthy, D.A. Hall, C.A. Hulka, G.J. Whitman, P. J. Slanetz, E.F. Halpern, Positive predictive value of breast biopsy performed as a result of mammography: there is no abrupt change at age 50 years, Radiology 200 (1996) 357–360.
[7] E.A. Sickles, D.L. Migioretti, R. Ballard-Barbash, B.M. Geller, J.W.T. Leung, R. D. Rosenberg, R. Smith-Bingman, B.C. Yankaskas, Performance benchmarks for diagnostic mammography, Radiology 235 (2005) 775–790.
[8] D. Gur, L.P. Wallace, A.H. Klym, L.A. Hardesty, G.S. Abrams, R. Shah, J. H. Sumkin, Trends in recall, biopsy, and positive biopsy rates for screening mammography in an academic practice, Radiology 235 (2005) 396–401.
[9] R.D. Rosenberg, B.C. Yankaskas, L.A. Abraham, E.A. Sickles, C.D. Lehman, B. M. Geller, P.A. Carney, K. Kerlikoske, D.S.M. Buist, D.L. Weaver, W.E. Barlow, R. Ballard-Barbash, Performance benchmarks for screening mammography, Radiology 241 (2006) 55–66.
[10] T.W. Freer, M.J. Ulissey, Screening mammography with computer-aided detection: prospective study of 12,860 patients in a community breast center, Radiology 220 (2001) 781–786.
[11] R.L. Birdwell, P. Bandodkar, D.M. Ikeda, Computer-aided detection with screening mammography in a university hospital setting, Radiology 236 (2005) 451–457.
[12] T.E. Cupples, J.E. Cunningham, J.C. Reynolds, Impact of computer-aided detection in a regional screening mammography program, AJR 185 (2005) 944–950.
[13] H.P. Chan, B. Sahiner, M.A. Roubidoux, T.E. Wilson, D.D. Adler, C. Paramagul, J. S. Newman, S. Sanjay-Gopal, Improvement of radiologists' characterization of mammographic masses by using coputer-aided diagnosis an ROC study, Radiology 212 (1999) 817–827.
[14] Z. Huo, M.L. Giger, C.J. Vyborny, C.E. Metz, Breast cancer: effectiveness of computer-aided diagnosis – observer study with independent database of mammograms, Radiology 224 (2002) 560–568.
[15] Y. Jiang, R.M. Nishikawa, R.A. Schmidt, C.E. Metz, M.L. Giger, K. Doi, Improving breast cancer diagnosis with computer-aided diagnosis, Acad. Radiol. 6 (1999) 22–32.
[16] R.M. Rangayyan, F.J. Ayres, J.E. Leo Desautels, A review of computer-aided diagnosis of breast cancer: toward the detection of subtle signs, J. Frankl. Inst. 344 (2007) 312–348.
[17] M. Elter, A. Horsch, CADx of mammographic masses and clustered microcalcifications: a review, Med. Phys. 36 (2009) 2052–2068.
[18] M. Tan, J. Pu, B. Zheng, Optimization of breast mass classification using sequential forward floating selection (SFFS) and a support vector machine (SVM) model, Int. J CARS 9 (2014) 1005–1020.
[19] C. Muramatsu, Q. Li, K. Suzuki, R.A. Schmidt, J. Shiraishi, G.M. Newstead, K. Doi, Investigation of psychophysical measure for evaluation of similar images for mammographic masses: preliminary results, Med. Phys. 32 (2005) 2295–2304.
[20] C. Muramatsu, Q. Li, R.A. Schmidt, J. Shiraishi, K. Doi, Investigation of psychophysical similarity measures for selection of similar images in the diagnosis of clustered microcalcifications on mammograms, Med. Phys. 35 (2008) 5695–5702.
[21] C. Muramatsu, R.A. Schmidt, J. Shiraishi, Q. Li, K. Doi, Presentation of similar images as a reference for distinction between benign and malignant masses on mammograms: analysis of initial observer study, J. Digit. Imaging 23 (2010) 592–602.
[22] C. Muramatsu, R.A. Schmidt, J. Shiraishi, T. Endo, H. Fujita, K. Doi, Usefulness of presentation of similar images in the diagnosis of breast masses on mammograms: comparison of observer performances in Japan and the USA, Radiol. Phys. Technol. 6 (2013) 70–77.
[23] C. Muramatsu, K. Nishimura, T. Endo, M. Oiwa, M. Shiraiwa, K. Doi, H. Fujita, Representation of lesion similarity by use of multidimensional scaling for breast masses on mammograms, J. Digit. Imaging 26 (2013) 740–747.
[24] M. Tan, J. Pu, B. Zheng, Reduction of false-positive recalls using a computerized mammographic image feature analysis scheme, Phys. Med. Biol. 59 (2014) 4357–4373.
[25] M.M. Eltoukhy, I. Faye, B.B. Samir, A comparison of wavelet and curvelet for breast cancer diagnosis in digital mammogram, Comput. Biol. Med. 40 (2010) 384–391.
[26] M.Z. do Nascimento, A.S. Martins, L.A. Neves, R.P. Ramos, E.L. Flores, G. A. Carrijo, Classification of masses in mammographic image using wavelet domain features and polynomial classifier, Expert. Syst. Appl. 40 (2013) 6213–6221.
[27] S. Ergin, O. Kilinc, A new feature extraction framework based on wavelets for breast cancer diagnosis, Comput. Biol. Med. 51 (2014) 171–182.
[28] T. Ojala, M. Pietikainen, T. Maenpaa, Multiresolution gray-scale and rotation invariant texture classification with local binary patterns, IEEE Trans. Pattern Anal. Mach., Intell. 24 (2002) 971–987.
[29] S. Lahmiri, M. Bouadoum, Hybrid discrete wavelet transform and Gabor filter banks processing for features extraction from biomedical images, J. Med. Eng. (2013) 13 pages.
[30] Y.A. Reyad, M.A. Berbar, M. Hussain, Comparison of statistical, LBP, and multiresolution analysis features for breast mass classification, J. Med. Syst. 38 (100) (2014) 15 pages.
[31] J. Chakraborty, A. Midya, S. Mukhopadhyay, A. Sadhu, Automatic characterization of masses in mammogram, Int. Conf. Biomed. Eng. Inform. (2013) 111–115.
[32] A. Midya, J. Chakraborty, Classification of benign and malignant masses in mammograms using multi-resolution analysis of oriented patterns, IEEE Int. Symp. Biomed. Imaging (2015) 411–414.
[33] R.S. Sahiner, H.P. Chan, N. Petrick, M.A. Helvie, M.M. Goodsitt, Computerized characterization of masses on mammograms: the rubber bank straightening transform and texture analysis, Med. Phys. 25 (1998) 516–526.
[34] X. Llado, A. Oliver, J. Freixenet, R. Marti, J. Marti, A textural approach for mass false positive reduction in mammography, Comput. Med. Image Graph. 33 (2009) 415–422.
[35] J.Y. Choi, Y.M. Ro, Multiresolution local binary pattern texture analysis combined with variable selection for application to false-positive reduction in computer-aided detection of breast masses on mammograms, Phys. Med. Biol. 57 (2012) 7029–7052.
[36] C. Muramatsu, M. Zhang, T. Hara, T. Endo, H. Fujita, Differentiation of malignant and benign masses on mammograms using radial local ternary pattern, in: H. Fujita, T. Hara, C. Muramatsu (Eds.), IWDM 2014, LNCS 8539, 2014, pp. 628–634.
[37] X. Tan, B. Triggs, Enhanced local texture feature sets for face recognition under difficult lighting conditions, IEEE Trans. Image Process. 19 (2009) 1635–1650.
[38] R.M. Haralick, K. Shanmugam, I. Dinstein, Textural features for image classification SMC-3, IEEE Trans. Syst. Man Cybern. (1973) 610–621.
[39] M.A. Tahir, A. Bouridane, F. Kurugollu, An FPGA based coprocessor for GLCM and Haralick texture features and their application in prostate cancer classification, Analog. Integr. Circuits Signal Process. 43 (2005) 205–215.
[40] ROC software, Available at: ⟨http://metz-roc.uchicago.edu/MetzROC/software#section-4⟩.