

Classification of Breast Tumors Using Sonographic Texture Analysis

Ali Abbasian Ardakani, MSc, Akbar Gharbali, PhD, Afshin Mohammadi, MD

Objectives—The purpose of this study was to evaluate a computer-aided diagnostic system with texture analysis to improve radiologists' accuracy in identification of breast tumors as malignant or benign.

Methods—The database included 20 benign and 12 malignant tumors. We extracted 300 statistical texture features as descriptors for each selected region of interest in 3 normalization schemes (default, $\mu - 3\sigma$, and $\mu + 3\sigma$, where μ and σ were the mean value and standard deviation, respectively, of the gray-level intensity and 1%–99%). Then features determined by the Fisher coefficient and the lowest probability of classification error + average correlation coefficient yielded the 10 best and most effective features. We analyzed these features under 2 standardization states (standard and nonstandard). For texture analysis of the breast tumors, we applied principle component, linear discriminant, and nonlinear discriminant analyses. First-nearest neighbor classification was performed for the features resulting from the principle component and linear discriminant analyses. Nonlinear discriminant analysis features were classified by an artificial neural network. Receiver operating characteristic curve analysis was used for examining the performance of the texture analysis methods.

Results—Standard feature parameters extracted by the Fisher coefficient under the default and 3σ normalization schemes via nonlinear discriminant analysis showed high performance for discrimination between benign and malignant tumors, with sensitivity of 94.28%, specificity of 100%, accuracy of 97.80%, and an area under the receiver operating characteristic curve of 0.9714.

Conclusions—Texture analysis is a reliable method and has the potential to be used effectively for classification of benign and malignant tumors on breast sonography.

Key Words—breast tumors; breast ultrasound; computer-aided diagnosis; sonography; texture analysis

Received March 4, 2014, from the Student Research Committee (A.A.A.) and Department of Medical Physics, Faculty of Medicine (A.G.), Urmia University of Medical Sciences, Urmia, Iran; and Department of Radiology, Faculty of Medicine, Imam Khomeini Hospital, Urmia University of Medical Sciences, Urmia, Iran (A.M.). Revision requested May 1, 2014. Revised manuscript accepted for publication June 2, 2014.

Address correspondence to Akbar Gharbali, PhD, Department of Medical Physics, Faculty of Medicine, Urmia University of Medical Sciences, Urmia 1138, Iran.

E-mail: gharbali@yahoo.com

Abbreviations

A_c , area under the curve; ROI, region of interest

doi:10.7863/ultra.34.2.225

Breast cancer may be considered the most common type of cancer in women from both developed and developing countries. According to a scientific report, it is the one of the most common causes of death in women in the United States, after lung and bronchial cancers.¹ To increase treatment options and decrease the death rate, early detection and diagnosis of breast cancer are very critical issues. Early detection requires an accurate and reliable diagnostic procedure.^{2,3} At present, breast cancer is detected by a combined approach, including biopsy, physical examination, and imaging.^{4,5} In general, breast cancer should be pathologically proven before any treatment decision is made. Although biopsy

is the most common clinical approach to determine whether a tumor is benign or malignant, it is painful, incurs health care costs, and has the risk of infection and bruising. To reduce the need for biopsy and improve the accuracy of imaging for differentiating between benign and malignant tumors, computer-aided diagnostic systems have been developed.^{3,6}

Mammography and sonography are two primary imaging modalities for detecting breast cancer. Although mammography is the main imaging technique, it has some restrictions. On radiographic examinations, radiologists misread 10% to 30% of breast cancers (false-positive and false-negative values), especially in dense breasts.^{3,7–11} To provide additional information and enhance the performance of mammography, sonography is recommended. Sonography is a popular medical imaging technique because of its lower cost, real-time scanning, and lack of radiation. Owing to the recent advancements in scanner resolution, transducer design, and signal processing, breast sonography may be considered comparable with mammography in terms of performance.^{12,13} Sonograms have diverse gray-level intensities, and different tissues have considerably different textures. Although there is no precise definition of image texture, it is perceived by humans. Generally, the textures of images are complex visual patterns in the region of interest (ROI) that characterize the distribution of gray-level values, brightness, color, size, frequency, roughness, and regularity, among other factors. A texture may contain substantial information about the internal structure of the human tissue or organ. Since radiologists usually assess texture qualitatively, quantitative texture analysis is required for more accurate diagnosis.^{14–16}

There are 4 types of features that can be used to classify benign and malignant breast tumors on sonography: model-based, descriptor, texture, and morphologic features.³ Many studies have been conducted to distinguish between benign and malignant tumors. Chen et al¹⁷ used fractal features and a K-means classifier to classify benign and malignant breast tumors, with accuracy of 88.8%, sensitivity of 93.64%, and specificity of 84.29%. Wu and Moon¹⁸ reported accuracy, sensitivity, and specificity of 92.8%, 94.44%, and 91.67% respectively, when morphologic features were used. Chen et al¹⁹ used texture features for classification of benign and malignant breast tumors. They used principle component analysis to diminish the dimensions of features and an image retrieval technique for differentiating breast tumors with mean accuracy \pm SD of $92.5\% \pm 0.019\%$. Moon et al²⁰ used gray level co-occurrence matrix, shape, and ellipsoid-fitting features. They indicated that combining shape and ellipsoid-fitting features could achieve the best

performance, with accuracy, sensitivity, and specificity of 85%, 84.5%, and 85.5%. Zhou et al²¹ indicated that extracted texture features by shearlet transform is more effective for describing breast tumors on sonography compared to other features extracted using wavelet, curvelet, contourlet, and gray-level co-occurrence matrices. They achieved accuracy of $91.0\% \pm 3.8\%$, sensitivity of 92.5% to 6.6%, and specificity of 90.0% to 3.8% using a support vector machine classifier.

In this study, we used texture analysis to extract texture features from sonograms to differentiate benign tumors from malignant ones. The most important texture features in texture analysis are computed from statistical, model-based, structural, and transform methods. The most commonly used texture parameters in texture analysis come from 6 main categories: histogram (statistical class), absolute gradient (statistical class), run length matrix (statistical class), co-occurrence matrix (statistical class), autoregressive model (model class), and wavelet (transform class).^{15,22}

Materials and Methods

The sonographic database consisted of 32 biopsy-proven tumors (20 benign and 12 malignant). Sonographic examinations were performed with an Accuvix V20 scanner (Medison Co, Ltd, Seoul, Korea) and an L5-13IS (5–13-MHz) linear transducer. This study protocol was approved by the Student Research Committee Bureau of the Urmia University of Medical Sciences.

One sonogram per patient was input into MaZda version 4.6 software (Lodz University of Technology, Lodz, Poland) for texture analysis.^{16,23} In general, more than 91 nonoverlapping ROIs consisting of 56 benign and 35 malignant tumors were selected for discrimination and classification. Before feature extraction for individual ROIs, we applied 3 normalization schemes: (1) default, in which images had the same appearance and an intensity range from 1 to 2k, where k was the number of the bits per pixel; (2) $\mu \pm 3\sigma$, in which the image intensities were located inside the normalization range ($\mu - 3\sigma$ and $\mu + 3\sigma$, where μ and σ were the mean value and standard deviation, respectively, of the gray-level intensity inside the ROI, so the intensity levels outside the normalization range were not considered in further analysis of the ROI; and (3) 1% to 99%, in which the ROI gray-level range between the darkness level at which the accumulated histogram of the image was equal to 1% of its total to the brightness level at which the accumulated histogram was equal to 99% of its total. Then we extracted 300 texture features

based on the histogram, absolute gradient (spatial variation of gray-level values), run length matrix (counts of pixel runs with the specified gray-level value and length in a given direction), co-occurrence matrix (information about the distribution of pairs of pixels separated by a given distance and direction), autoregressive model (description of correlation between neighboring pixels), and wavelet (decomposition image frequency at different scales).

Because of the large number of features, they were not suitable for statistical analysis. We used 2 well-known automated reduction algorithms (Fisher and probability of classification error + average correlation coefficient) for selection of up to 10 texture features with the highest Fisher and lowest probability of classification error + average correlation coefficient features that showed the best discrimination between benign and malignant tumors. The Fisher algorithm uses a ratio of between-class variance to within-class variance. The probability of classification error + average correlation coefficient algorithm uses classification error probability combined with average correlation coefficients.^{24,25} Next, these features were analyzed by principle component analysis, linear discriminant analysis, and nonlinear discriminant analysis methods under 2 standardization states (standard and nonstandard).^{26–29} First-nearest neighbor classification was performed for the features resulting from the principle component and linear discriminant analyses. Nonlinear discriminant analysis features were classified by an artificial neural network.^{29,30}

At the end, receiver operating characteristic curve analysis was used for evaluating the performance of the applied texture analysis methods by calculating area under the curve (A_z), sensitivity, specificity, and overall accuracy.^{31–34} In this study, sensitivity and specificity represent the probability of a correct diagnosis of malignant tumors and the probability of a correct detection of benign tumors by the radiologist, respectively. The A_z value represents the correct classification ability of the test. Accuracy shows the percentage of cases that were correctly diagnosed. Figure 1 shows the computer-aided diagnostic processing steps.

Results

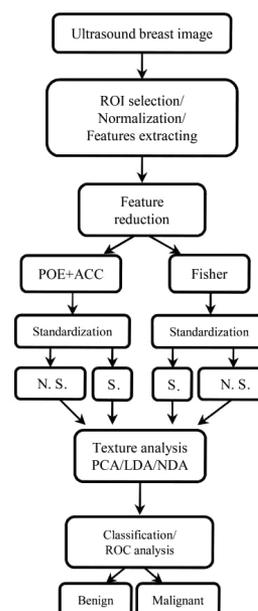
As mentioned previously, in this study we applied all available options in the MaZda program for texture analysis: 3 normalization schemes, 2 feature reduction methods, 2 standardization states, and 3 texture analysis methods. In default normalization, the features extracted by both the Fisher and probability of classification error + average correlation coefficient algorithms showed higher performance in nonlinear discriminant analysis than principle component

analysis and linear discriminant analysis, with sensitivity of 94.28%, specificity of 100%, and accuracy of 97.80% (Table 1). In 3σ normalization, nonlinear discriminant analysis showed the same discrimination performance only with features from the Fisher reduction method (Table 2). In 1% to 99% normalization, probability of classification error + average correlation coefficient features showed higher sensitivity, specificity, and accuracy (94.28%, 98.21% and 96.70%, respectively) in nonlinear discriminant analysis than principle component analysis and linear discriminant analysis (Table 3). Figures 2–4 show receiver operating characteristic curves plotted on the same graphs for discriminating benign from malignant breast tumors based on the normalization and standardization schemes. In general, the nonlinear discriminant analysis method, which had the greatest A_z value, had an advantage over principle component analysis and linear discriminant analysis in each state. Table 4 shows the best performance in this study.

Discussion

In this study, 3 normalizations, 2 reduction automated elimination algorithms, 2 standardization states, and 3 texture data analysis methods altogether provided 36 states per

Figure 1. Overview of the general breast sonographic texture analysis process. ACC indicates average correlation coefficient; LDA, linear discriminant analysis; NDA, nonlinear discriminant analysis; N.S., nonstandard; PCA, principle component analysis; POE, probability of classification error; and S., standard.



ROI case study. The best results were driven by 3 conditions (1) default normalization with features extracted by the Fisher algorithm and analyzed by nonlinear discriminant analysis; (2) 3σ normalization with features extracted by

the Fisher algorithm and analyzed by nonlinear discriminant analysis; and (3) default normalization with features extracted by the probability of classification error + average correlation coefficient algorithm and analyzed by nonlin-

Table 1. Summary of Performance for Different Features and Feature Reduction Methods in Default Normalization

Feature Reduction Method	Feature Analysis Method	SEN, %	SPC, %	ACCY, %	PPV, %	NPV, %	A_z
Fisher	N.S. PCA	71.42	82.14	78.02	71.42	82.14	0.7678
	S. PCA	85.71	92.85	90.10	88.23	91.22	0.8928
	N.S. LDA	88.57	92.85	91.20	88.57	92.85	0.9071
	S. LDA	88.57	92.85	91.20	88.57	92.85	0.9071
	NDA	94.28	100	97.80	100	96.55	0.9714
POE + ACC	N.S. PCA	71.42	82.14	78.02	71.42	82.14	0.7678
	S. PCA	88.57	96.42	95.60	93.93	93.10	0.9250
	N.S. LDA	88.57	92.58	91.20	88.57	92.85	0.9057
	S. LDA	88.57	92.58	91.20	88.57	92.85	0.9057
	NDA	94.28	100	97.80	100	96.55	0.9714

ACCY indicates accuracy; NPV, negative predictive value; PPV, positive predictive value; SEN, sensitivity; and SPC, specificity; other abbreviations are as in Figure 1.

Table 2. Summary of Performance for Different Features and Feature Reduction Methods in 3σ Normalization

Feature Reduction Method	Feature Analysis Method	SEN, %	SPC, %	ACCY, %	PPV, %	NPV, %	A_z
Fisher	N.S. PCA	82.85	85.71	84.61	78.37	88.88	0.8428
	S. PCA	80.00	83.92	82.41	75.67	87.03	0.8196
	N.S. LDA	82.85	91.07	87.91	85.29	89.47	0.8696
	S. LDA	82.85	91.07	87.91	85.29	89.47	0.8696
	NDA	94.28	100	97.80	100	96.55	0.9714
POE + ACC	N.S. PCA	77.14	85.71	82.41	77.14	85.71	0.8142
	S. PCA	80.00	89.28	85.71	82.35	86.20	0.8321
	N.S. LDA	85.71	92.85	90.10	88.23	91.22	0.8928
	S. LDA	85.71	92.85	90.10	88.23	91.22	0.8928
	NDA	94.28	98.21	96.70	97.05	96.49	0.9624

Abbreviations are as in Figure 1 and Table 1.

Table 3. Summary of Performance for Different Features and Feature Reduction Methods in 1% to 99% Normalization

Feature Reduction Method	Feature Analysis Method	SEN, %	SPC, %	ACCY, %	PPV, %	NPV, %	A_z
Fisher	N.S. PCA	85.71	87.50	86.81	81.08	90.74	0.8142
	S. PCA	85.71	92.85	90.10	88.23	91.22	0.8321
	N.S. LDA	80.00	89.28	82.41	82.35	87.71	0.8928
	S. LDA	80.00	89.28	82.41	82.35	87.71	0.8928
	NDA	94.28	96.42	95.60	94.28	96.42	0.9624
POE + ACC	N.S. PCA	85.71	85.71	85.71	78.94	90.56	0.8571
	S. PCA	85.71	87.50	86.81	81.08	90.74	0.8660
	N.S. LDA	91.42	91.07	91.20	86.48	94.44	0.9124
	S. LDA	91.42	91.07	91.20	86.48	94.44	0.9124
	NDA	94.28	98.21	96.70	97.05	96.49	0.962

Abbreviations are as in Figure 1 and Table 1.

ear discriminant analysis. The receiver operating characteristic curve analysis indicated that all of these conditions had similar performance in differentiating between benign and malignant tumors (Figure 5).

Feature standardization has only a small effect on principle component analysis, and it leads to an improvement in performance. It has no impact on linear discriminant analysis. To decrease the training time with the artificial

Figure 2. Receiver operating characteristic curves for each texture analysis method in default normalization: **A**, Fisher features; **B**, probability of classification error + average correlation coefficient features. Abbreviations are as in Figure 1.

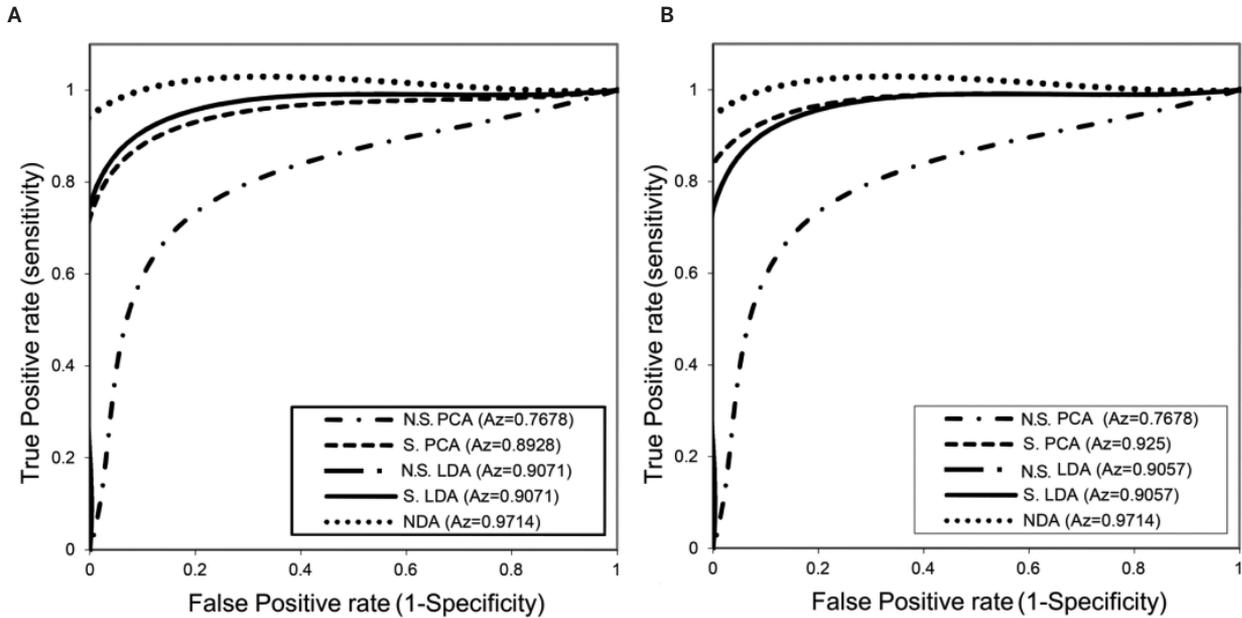
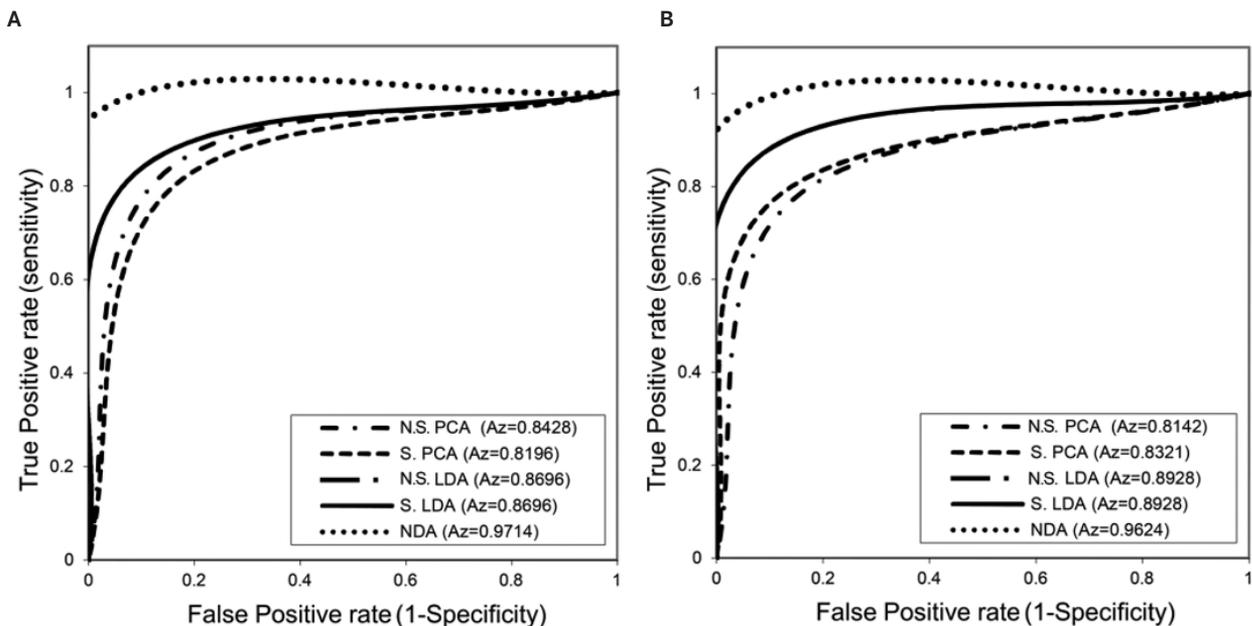


Figure 3. Receiver operating characteristic curves for each texture analysis method in 3σ normalization: **A**, Fisher features; **B**, probability of classification error + average correlation coefficient features. Abbreviations are as in Figure 1.



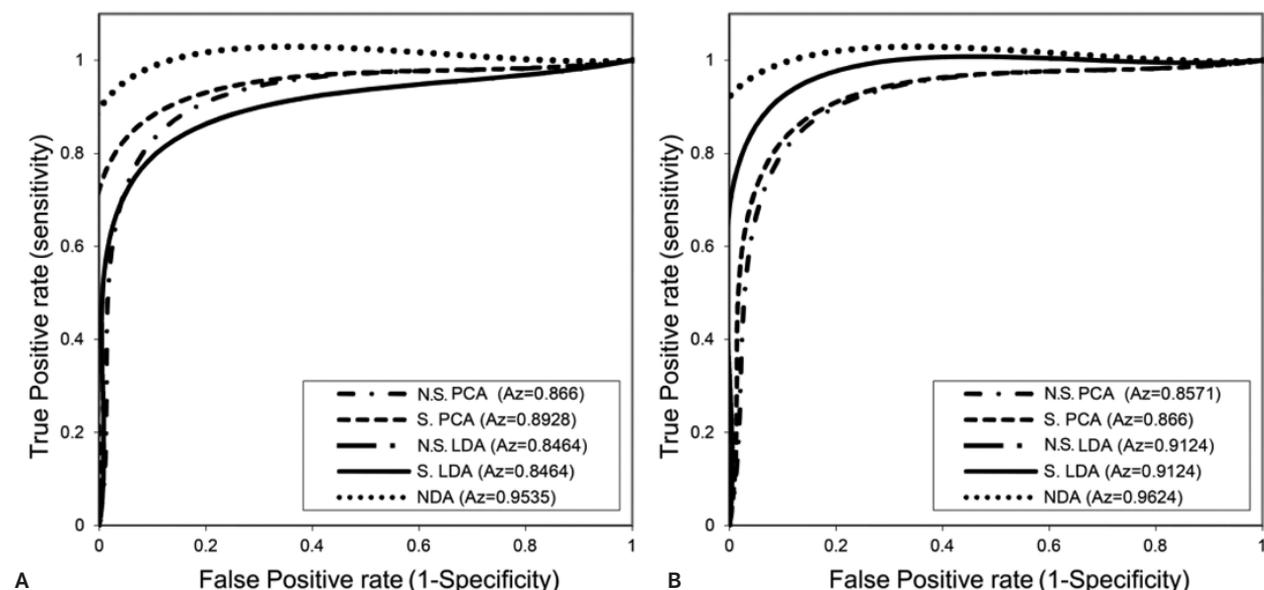


Figure 4. Receiver operating characteristic curves for each texture analysis method in 1% to 99% normalization: **A**, Fisher features; **B**, probability of classification error + average correlation coefficient features. Abbreviations are as in Figure 1.

Table 4. Summary of Best Performance

Feature Reduction Method	Feature Analysis Method	Normalization	SEN, %	SPC, %	ACCY, %	PPV, %	NPV, %	A _z
Fisher	NDA	Default	94.28	100	97.8	100	96.55	0.9714
POE + ACC	NDA	Default	94.28	100	97.8	100	96.55	0.9714
Fisher	NDA	3σ	94.28	100	97.8	100	96.55	0.9714

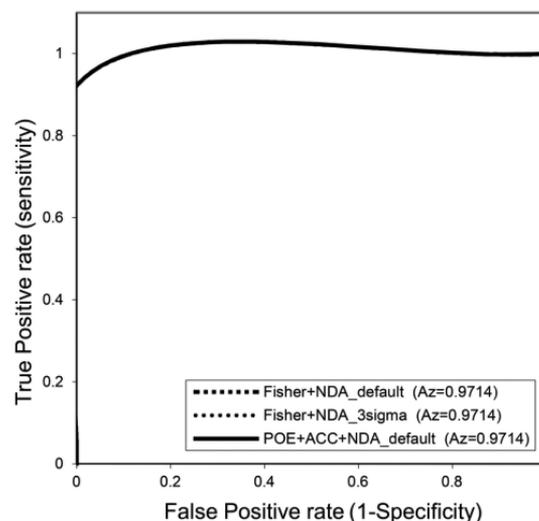
Abbreviations are as in Figure 1 and Table 1.

neural network classifier, all input feature parameters were standardized previously in nonlinear discriminant analysis.

The proposed method demonstrated more reliable performance in comparison to all studies^{17–21} already conducted on the subject of sonographic differentiation between benign and malignant breast cancers, with sensitivity ranging from 84.5% to 94.44%, specificity ranging from 84.29% to 91.67%, and accuracy ranging from 85% to 92.8%.

Some limitations of this study should be clearly noted. First, the data group was small; to overcome this limitation, several ROIs for each tumor were selected. Further investigation with a larger data set is needed. Second, feature combination tools were not available in the MaZda program. For example, averaging of run length matrix features of 4 different orientations was hard to perform with MaZda. Third, since lower frequencies are used for obese breasts, those images had lower spatial resolution than others. Thus, further investigation with separate uniform groups is needed.

Figure 5. Receiver operating characteristic curves for the best results. All 3 methods had the same performance. Abbreviations are as in Figure 1.



The MaZda software was developed in 1998 for the purpose of automatic texture analysis on magnetic resonance imaging.²² We used it for sonography because of its superiority over other texture analysis methods by providing more than 300 features that are essential for texture pattern recognition. Generally, our results indicate that an automatic texture analysis method with application of MaZda software can provide useful information that contributes to discrimination of breast cancers on sonography and also has the potential to help radiologists in detection and classification of benign and malignant breast tumors.

References

1. Brant WE, Helms C. *Fundamentals of Diagnostic Radiology*. Philadelphia, PA: Lippincott Williams & Wilkins; 2012.
2. Sivaramakrishna R, Gordon R. Detection of breast cancer at a smaller size can reduce the likelihood of metastatic spread: a quantitative analysis. *Acad Radiol* 1997; 4:8–12.
3. Jalalian A, Mashohor SBT, Mahmud HR, Saripan MIB, Ramli ARB, Karasfi B. Computer-aided detection/diagnosis of breast cancer in mammography and ultrasound: a review. *Clin Imaging* 2013; 37:420–426.
4. Coughlin SS, Ekwueme DU. Breast cancer as a global health concern. *Cancer Epidemiol* 2009; 33:315–318.
5. Esserman LJ. New approaches to the imaging, diagnosis, and biopsy of breast lesions. *Cancer J* 2002; 8(suppl 1):S1–S14.
6. Jesneck JL, Lo JY, Baker JA. Breast mass lesions: computer-aided diagnosis models with mammographic and sonographic descriptors. *Radiology* 2007; 244:390–398.
7. Gordon PB, Goldenberg SL. Malignant breast masses detected only by ultrasound: a retrospective review. *Cancer* 1995; 76:626–630.
8. Beyer T, Moonka R. Normal mammography and ultrasonography in the setting of palpable breast cancer. *Am J Surg* 2003; 185:416–419.
9. Crystal P, Strano SD, Shcharynski S, Koretz MJ. Using sonography to screen women with mammographically dense breasts. *AJR Am J Roentgenol* 2003; 181:177–182.
10. Berg WA. Supplemental screening sonography in dense breasts. *Radiol Clin North Am* 2004; 42:845–851, vi.
11. Nothacker M, Duda V, Hahn M, et al. Early detection of breast cancer: benefits and risks of supplemental breast ultrasound in asymptomatic women with mammographically dense breast tissue—a systematic review. *BMC Cancer* 2009; 9:335.
12. Gerlach B, Holzgreve W. Comparison of X-ray mammography and sonomammography of 1209 histologically verified breast diseases. In: *Breast Ultrasound Update*. Basel, Switzerland: Karger Verlag; 1994:40–50.
13. Cheng XY, Akiyama I, Itoh K, et al. Breast tumor diagnosis system using 3 dimensional ultrasonic echography. Paper presented at: 19th Annual International Conference of the IEEE; October 30–November 2, 1997; Chicago, IL.
14. Materka A. Texture analysis methodologies for magnetic resonance imaging. *Dialogues Clin Neurosci* 2004; 6:243–250.
15. Castellano G, Bonilha L, Li LM, Cendes F. Texture analysis of medical images. *Clin Radiol* 2004; 59:1061–1069.
16. Szczypiński PM, Strzelecki M, Materka A, Klepaczko A. MaZda: a software package for image texture analysis. *Comput Methods Programs Biomed* 2009; 94:66–76.
17. Chen DR, Chang RF, Chen CJ, et al. Classification of breast ultrasound images using fractal feature. *Clin Imaging* 2005; 29:235–245.
18. Wu WJ, Moon WK. Ultrasound breast tumor image computer-aided diagnosis with texture and morphological features. *Acad Radiol* 2008; 15:873–880.
19. Chen DR, Huang YL, Lin SH. Computer-aided diagnosis with textural features for breast lesions in sonograms. *Comput Med Imaging Graph* 2011; 35:220–226.
20. Moon WK, Shen YW, Huang CS, Chiang LR, Chang RF. Computer-aided diagnosis for the classification of breast masses in automated whole breast ultrasound images. *Ultrasound Med Biol* 2011; 37:539–548.
21. Zhou S, Shi J, Zhu J, Cai Y, Wang R. Shearlet-based texture feature extraction for classification of breast tumor in ultrasound image. *Biomed Signal Processing Control* 2013; 8:688–696.
22. Materka A, Strzelecki M. *Texture Analysis Methods: A Review*. COST B11 Report. Lodz, Poland: Lodz University of Technology, Institute of Electronics; 1998.
23. Hájek M, Dezortova M, Materka A, Lerski R (eds). *Texture Analysis for Magnetic Resonance Imaging*. Prague, Czech Republic: Med4Publishing; 2006.
24. Schürmann J. *Pattern Classification: A Unified View of Statistical and Neural Approaches*. New York, NY: John Wiley & Sons; 1996.
25. Mucciardi AN, Gose EE. A comparison of seven techniques for choosing subsets of pattern recognition properties. *IEEE Trans Comput* 1971; 100:1023–1031.
26. Webb AR. *Statistical Pattern Recognition*. Hoboken, NJ: John Wiley & Sons; 2003.
27. Fukunaga K. *Introduction to Statistical Pattern Recognition*. Philadelphia, PA: Elsevier; 1990.
28. Krzanowski WJ, Krzanowski W. *Principles of Multivariate Analysis*. Oxford, England: Oxford University Press; 2000.
29. Duda RO, Hart PE, Stork DG. *Pattern Classification*. Hoboken, NJ: John Wiley & Sons; 2012.
30. Anderson JA, Rosenfeld E. *Neurocomputing*. Vol 2. Cambridge, MA: MIT Press; 1993.
31. Metz CE. Some practical issues of experimental design and data analysis in radiological ROC studies. *Invest Radiol* 1989; 24:234–245.
32. Van Erkel AR, Pattynama PMT. Receiver operating characteristic (ROC) analysis: basic principles and applications in radiology. *Eur J Radiol* 1998; 27:88–94.
33. Obuchowski NA. Receiver operating characteristic curves and their use in radiology. *Radiology* 2003; 229:3–8.
34. Soreide K, Körner H, Soreide JA. Diagnostic accuracy and receiver-operating characteristics curve analysis in surgical research and decision making. *Ann Surg* 2011; 253:27–34.