Mass Classification Based on Texture Features in Breast Ultrasound Image

H. D. Cheng and X. J. Shi

Department of Computer Science, Utah State University, 401B. Old Mail Hill, Logan UT 84322, USA

Abstract

Breast cancer occurs to over 8% women during their lifetime, and is the leading cause of death of women in the United States. Sonography is superior to the mammography in the ability to detect focal abnormalities in the dense breasts of adolescent women and has no side-effect. In this paper, the stepwise logistic regression method is investigated to select an optimal subset of texture features, and the back-propagation neural network is employed as the classifier to discriminate benign solid masses from malignant solid masses based on texture features. The experimental results showed that the proposed model results in high accuracy of mass detection in breast ultrasound images.

1. Introduction

Breast cancer occurs to over 8% women during their lifetime, and is the leading cause of death of women in the United States. It is estimated that there was approximately 215,990 new cases of invasive breast cancer and about 40,589 deaths were caused by the breast cancer in the United States in 2004 [1]. The breast cancer can be most effectively treated if it is detected at its early stage. Currently the most effective method for early detection and screening of breast cancers is mammography. However, reading mammography is a demanding job for radiologists, and reading ultrasound images will be the same. The judgments depend on training, experience, and subjective criteria. About 10 percentage of the breast cancer is missed by radiologists, and most of them are in dense breasts [2]. On the other hand, about two-thirds of lesions that were sent to biopsy are benign. The reasons for this high miss rate and low specificity of mammography are: (1) the low conspicuity of mammography lesions; (2) the noisy nature of the images; and (3) the overlying and underlying structures that obscure the features of ROIs [3]. Sonography has proven to be an important adjunct to mammography in the breast cancer detection and has been primarily useful for differentiating cysts from solid tumors. Furthermore, it has been shown that breast sonography is superior to the mammography in the ability to detect focal abnormalities in the dense breasts of adolescent women [4]. The accuracy rate of breast ultrasound can reach at 96-100% in the diagnosis of simple benign cysts, and the lesions characterized as benign cysts do not require biopsies, which are expensive and involve minor risks, for further evaluation [3]. To avoid these unnecessary surgical procedures, many researchers investigate the characteristics of the solid breast masses for discriminating between benign or malignant.

Computer aided diagnosis (CAD) systems may help radiologists in interpreting sonography for mass detection and classification. It is very important to develop CAD systems that can distinguish benign lesions from malignant lesions. The combination of the CAD scheme and experts' knowledge would greatly improve the detection accuracy of the abnormalities detection.

In this paper, the texture features of ROIs are analyzed; then the stepwise regression method is used to select an optimal subset of the features; finally a back-propagation neural network is employed as the classifier to discriminate benign and malignant tumors.

2 Material and methods

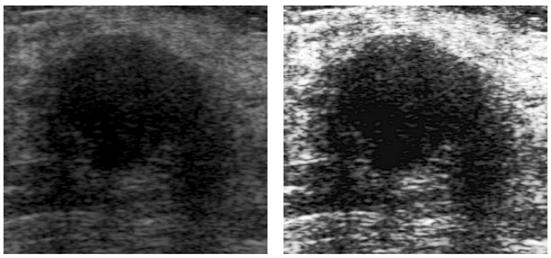
The ultrasound images originated from an ultrasonic scanner (VIVID7) manufactured by GE Medical Systems, the frequency ranges of the ultrasonic scanner are 5-12MHz [5]. The dataset consists of 35 cases (12 benign solid masses and 23 malignant solid masses). All ROIs were manually segmented by a trained observer.

2.1 Preprocessing

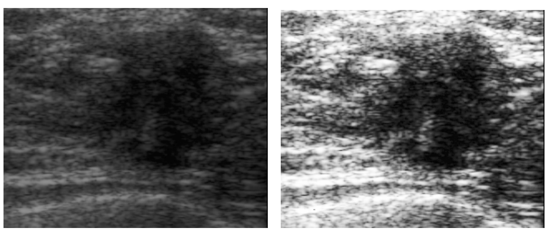
Preprocessing is an important issue in low-level image processing. The underlying principle of the preprocessing is to make image clearer. Preprocessing of breast ultrasound image is to produce reliable representations of breast tissue structures by enhancing the contrast. The effective method for enhancement must aim to enhance texture and features of masses. The reasons are: (1)

low-contrast of breast ultrasound images; (2) hard to read masses in breast ultrasound image.

The Multi-peak GHE method, our previous work [6] is very effective not only in enhancing the entire image but also in enhancing the image texture. It also makes the change of the order of gray levels of the original image completely controllable. Fig. 1 shows the performance of enhancing the ultrasound image:



(a) (b)



(a) (b)

Figure 1 (a) Original image; (b) the result of the enhancement by the method in [6].

2.2 Texture feature extraction and selection

Feature extraction and selection play a key role in pattern recognition. Good features and their optimum combination can greatly improve the accuracy in classifying the suspicious abnormalities into benign or malignant solid masses. Extracting efficient and effective features that can discriminate benign from malignant solid masses to the greatest extent is a challenging task.

The ideal image features can well characterize an object. The features should be very similar for objects in the same category while they are very different for objects in different categories. Image features can be extracted from the ROI characteristics.

2.2.1 Texture feature extraction

Recently, researchers and scientists employed a lot of features based on the characteristics of benign and/or malignant solid masses in breast ultrasound images to discriminate benign masses from malignant masses. The texture features are the most popular and important features in the breast ultrasound image analysis [3, 7-10]. In this paper, the texture features and fractal dimension are analyzed.

(1) Texture features

Texture features[10] are derived from the spatial gray level dependence (SGLD) matrices. SGLD matrices are used to measure the texture-context information. It is a 2-D histogram. An element of the SGLD matrix P(i,j,d,θ) is defined as the joint probability that the gray levels i and j occur separated by distanced d and along direction θ of the image. In order to simplify the computational complexity of the algorithm, the θ is often given as 0°, 45°, 90°, and 135°, and the distance d is often defined as the Manhattan or city block distance. The element P(i,j,d,θ) of the SGLD matrix can be expressed:

$$P(i,j,d,\theta)=\left\|\left\{\left((x_1,y_1),(x_2,y_2)\right)\right\}\right\|,$$

where (x_1, y_1) and (x_2, y_2) are satisfied with following condition:

$$if\left(\theta=0°\right) \quad then$$

$$\left|x_2-x_1\right|=d, y_2-y_1=0$$

$$if\left(\theta=45°\right) \quad then$$

$$\left(x_2-x_1=d, y_2-y_1=-d\right)$$

$$\text{or } \left(x_2-x_1=-d, y_2-y_1=d\right)$$

$$if\left(\theta=90°\right) \quad then$$

$$x_2-x_1=0, \left|y_2-y_1\right|=d$$

$$if\left(\theta=135°\right) \quad then$$

$$\left(x_2-x_1=d, y_2-y_1=d\right)$$

$$\text{or } \left(x_2-x_1=-d, y_2-y_1=-d\right)$$

where $i = I(x_1, y_1)$, $j = I(x_2, y_2)$, $I(x, y)$ is the intensity value of the pixel at the position (x, y), and ‖S‖ the number of the elements in the set S.

The textural features can be extracted from SGLD matrices with different distance d and direction θ.

In practice, given the distance d, four SGLD matrices, which are pertinent to the value of θ as 0°, 45°, 90°, and 135° respectively, are calculated. Hence a set of four values for each of 14 measures are obtained based on four SGLD matrices. For each measure with four values, we can get its mean and range. Therefore, a set of 28 textural features are extracted from these four matrices for a given distance d. Some of the features are strongly correlated with each other. A feature selection procedure may be applied to select a subset of the 28 features.

In our experiment, five SGLD matrices are calculated at five different distances (d = 1, 2, 4, 8, and 16 pixels) to produce 140 features.

(2) Fractal dimension

The fractal concept [11] is very useful to represent a statistical quality of roughness and self-similarity at different scales of most nature surfaces and/or curves [12]. This statistical quality is most often characterized in terms of the fractal dimension. The fractal dimension as a geometric feature has become quite popular in modeling image properties. Intuitively, the degree of roughness of the image texture is proportional to the fractal dimension. The definition of the fractal dimension is similar to the Hausdorff dimension [13]. Informally, self-similar objects with parameters N and s are described by a power law such as:

$$N = S^d$$

Thus

$$d = \frac{\log N}{\log s}$$

is defined as the "dimension" of the scaling law, known as the Hausdorff dimension.

There are many methods for estimating the fractal dimensions of a gray-level image [12-14]. The popular way to compute the fractal dimension is based on the box-count algorithm due to its simplicity [14].

2.2.2 Feature selection

More than 100 texture features are derived from a breast ultrasound image. But not all of the features are suitable for mass classification of breast ultrasound images. Also the correlation may exist among the features. Too many irrespective features and correlation among the features not only make the classifier complicated, but also will reduce the accuracy of the classification.

Stepwise regression is a statistic technique for choosing an optimal subset of explanatory variables. Based on the 141 features (140 texture features and the fractal dimension), stepwise regression produces an optimal subset of the features that comprised of 4 features: the angular second moment (d = 8), the range of difference entropy (d = 8), the sum average (d = 16), and fractal dimension.

2.3 Classification

There are several kinds of classifiers to classify the suspicious lesions into benign or malignant solid masses. The Artificial Neural Networks (ANNs) are the most popular. ANNs are the collection of mathematical models that imitate the properties of biological nervous system and the functions of adaptive biological learning. They are made of many processing elements that are highly interconnected together with weighted links that are similar to synapses. Unlike linear discriminant, ANNs usually use non-linear mapping functions as decision boundaries. The advantage of ANNs is their capability suitable to solve problems that are too complex to use conventional techniques, or hard to find algorithmic solutions.

3. Experimental results and discussion

It is important to notice that an objective comparison of the performance of different CAD systems is very difficult and even impossible due to the use of different databases. Even if a common database is used to test different systems, it could not guarantee that the comparison is valid and just.

Five descriptive statistics (accuracy, sensitivity, specificity, positive predictive value and negative predictive value) are the most generally used objective indices to estimate the performance of the classification results in clinical practice [7, 8]. Given TP (true-positive: the number of correctly diagnosed malignant cases), TN (true-negative: the number of correctly diagnosed benign cases), FP (false-positive: the number of incorrectly diagnosed malignant cases), and FN (false-negative: the number of incorrectly diagnosed benign cases), the five statistic indices are defined as follows.

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN}$$

$$Sensitivity = \frac{TP}{TP + FN}$$

$$Specificity = \frac{TN}{TN + FP}$$

$$Positive\ predictive\ value\ (PPV) = \frac{TP}{TP + FP}$$

$$Negative\ predictive\ value\ (NPV) = \frac{TN}{TN + FN}$$

One of the criteria to evaluate the reliable and valid CAD system is that the higher the five indices are, the more reliable and valid the CAD system is.

In our experiment, 35 breast ultrasound images (12 benign solid masses and 23 malignant solid masses) are analyzed. The leave-one-out method, also called cross-validation [15] is useful to test the performance with small sample size. The leave-one-out method examines the classification performance on each individual sample by removing it from the complete training set. In our experiment, for each sample, it is considered as the test sample, and the other 34 samples are considered as the training set. The test sample is classified based on the ANN model resulted from the training set.

Table 1 showed the result of our experiments, and Table 2 showed the comparison of performance between the proposed approach and the bootstrap technique proposed in [8].

From the definitions of the five indices, we know that the higher the specificity and PPV are, the lower the FP, and the higher the Sensitivity and NPV are, the lower the FN. By the comparison of performance between the proposed method and the approach in [8], three indices of accuracy, specificity, and PPV are higher, and other two indices are a little lower in the proposed approach than the approach [8]. We concludes that the proposed approach can classify the malignant solid masses more accurately than the approach in [8], and also can result in high accuracy of detecting the benign solid masses. In short word, the proposed approach is more acceptable to patients.

From these two tables, we conclude that:

1) All five indices are very high, greater than 83. It indicates the proposed method is quite reliable.
2) All malignant solid masses are classified correctly. It is very important to CAD systems.
3) Only two of benign solid masses are not classified correctly. It can be tolerant in practice in order to make all the malignant solid masses be diagnosed correctly.

Table 1 Result by the proposed approach

TP	FP	TN	FN
23	0	10	2

Table 2 Comparison of the performances

	Proposed approach	Approach in [8]
Accuracy (%)	94.3	87.07
Sensitivity (%)	92	95.35
Specificity (%)	100	79.10
PPV (%)	100	81.46
NPV (%)	83.3	94.64

References

[1] "American Cancer Society: Cancer Facts & Figures," 2005.

[2] V. Jackson, R. Hendrick, S. Feig, and D. Kopans, "Imaging of the Radiographically Dense Breast," *Radiology*, vol. 188, pp. 297-301, 1993.

[3] R. Sivaramakrishna, K. A. Powell, M. L. Lieber, W. A. Chilcote, and R. Shekhar, "Texture analysis of lesions in breast ultrasound images," *Computerized Medical Imaging and Graphics*, vol. 26, pp. 303-307, 2002.

[4] C. Sohn, J.-U. Blohmer, and U. M. Hamper, *Breast Ultrasound*, 1999.

[5] H. D. Cheng, L. Hu, J. Tian, and a. L. Sun, "A Novel Markov Random Field Segmentation Algorithm and Its Application to Breast Ultrasound Image Analysis," *in progress*.

[6] H. D. Cheng and X. J. Shi, "A simple and effective histogram equalization approach to image enhancement," *Digital Signal Processing*, vol. 14, pp. 158-170, 2004.

[7] 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. 679-686, 2003.

[8] D. R. Chen, W. J. Kuo, R. F. Chang, W. K. Moon, and C. C. Lee, "Use of the bootstrap technique with small training sets for computer-aided diagnosis in breast ultrasound," *Ultrasound in Medicine and Biology*, vol. 28, pp. 897-902, 2002.

[9] B. S. Garra, B. H. Krasner, S. C. Horii, S. Ascher, s. K. Mun, and R. K. Zeman, "Improving the Distinction Between Benign and Malignant Brest Lesions: the Value of Sonographic Texture Analysis," *Ultrasound Imaging*, vol. 15, pp. 267-285, 1993.

[10] R. M. Haralick, H. K. Shanmugam, and I. Dinstein, "Texture features for image classification," *IEEE Transaction on Systems, Man, and Cybernetics*, vol. SMC-3, pp. 610-621, 1973.

[11] B. B. Mandelbrot, *The Fractal Geometry of Nature*. San Francisco, CA: Freeman, 1982.

[12] M. Tuceryan and A. K. Jain, "Texture Analysis," in *The Handbook of Pattern Recognition and Computer Vision*, C. H. Chen, L. F. Pau, and P. S. P. Wang, Eds., 2nd ed: World Scientific Publishing Co., 1998, pp. 207-248.

[13] C. C. Chen, J. S. Daponte, and M. D. Fox, "Fractal Feature Analysis and Classification in Medical Imaging," *Ieee Transactions on Medical Imaging*, vol. 8, pp. 133-142, 1989.

[14] W.-S. Chen, S.-Y. Yuan, and C.-M. Hsieh, "Two algorithms to estimate fractal dimension of gray-level images," *Optical Engineering*, vol. 42, pp. 2452-2464, 2003.

[15] M. Nadler and E. P. Smith, *Pattern Recognition Engineering*. New York: Wiley and Sons, 1993.