Journal of Computers Vol. 31 No. 4, 2020, pp. 1-13
doi:10.3966/199115992020083104001

An Image Classification Method for Digital Breast Tomosynthesis Based on Combined Texture Feature Extraction

You-Ming Wang*, Jia-Qi Miao, Han-Mei Zhang

School of Automation, Xi'an University of Posts and Telecommunications, Xi'an 710121, China
xautroland@126.com

Received 18 November 2018; Revised 17 April 2019; Accepted 8 May 2019

Abstract. A new combined algorithm is proposed for the texture feature extraction of digital breast tomosynthesis (DBT) image based on gray level co-occurrence matrix (GLCM), Tamura and Relief algorithms. The disadvantage of the GLCM algorithm is that the global information including the pixel dependencies among the textures can not be fully utilized. The presentation of Tamura algorithm can overcome the shortcomings of the GLCM algorithm on the way of integrating the characteristics of human visual information into the process of feature extraction. In this paper, the texture features are screened by Relief algorithm, which is set as a criterion to select the optimal features. The DBT image is preprocessed to reduce the noise and increase the contrast by bilateral filtering, contrast-limited adaptive histogram equalization and L0 gradient filtering. The moving window is used to traverse the image. In each moving window, the energy, contrast, correlation, entropy, inverse difference moment (IDM) and correlation of the GLCM algorithm and the features of the Tamura algorithm including the coarseness, contrast and directivity are extracted, respectively. Each set of Tamura texture features is integrated into GLCM texture features to construct a fused texture feature space and the fused texture features are filtered for the optimal features by Relief algorithm. The filtered texture features are input into the support vector machine (SVM) for the image classification. Compared with other extraction algorithms, the proposed algorithm can improve the accuracy of feature extraction and recognition. The results show that the classification accuracy and efficiency of the proposed algorithm are much improved compared with SVM, GLCM, Tamura and GLCM-Tamura algorithms.

Keywords: digital breast tomosynthesis image, feature selection, gray level co-occurrence matrix, support vector machine

1 Introduction

Breast cancer is one of the main causes of cancer death for women worldwide. There are some limitations of traditional breast photography in the analysis of the texture and density of breasts. The mammography is projection X-ray images, in which the tissue layer of the breast is superimposed constantly. The digital breast tomosynthesis (DBT) [1] image is considered as a standard three-dimensional (3D) imaging modality that is reconstructed from low-dose X-rays in different angles. DBT imaging technology is based on the mammography, which reduces the tissue overlap of mammography and improves the detection rate of breast cancer. The texture features of DBT images are usually extracted in the classification of breast tissue and its two-dimensional grayscale features are used to represent detailed feature in-formation of the image.

Recently, many classification methods have been emerged, such as ID3 algorithm [2-3], C4.5 algorithm [4-5], CART algorithm [6-7], BP algorithm [8-9], Bayesian classification [10], k-NN algorithm [11], Rough set approach [12], fuzzy set method [12] and support vector machine (SVM) [14], Yao et al.

* Corresponding Author

carried out a series of studies on the granularity computation for classification problems. The association rule mining is an important research field in data mining [15-17]. The algorithms of associative classification mainly include CBA [18], ADT [19], CMAR [20], etc. Ye and Li proposed a novel classification method CCA-S based on the clustering method and the K-NN method [21]. SVM is an effective data mining tool for classification, clustering, and time series analysis. The clustering-based SVM (CB-SVM) [22-24] algorithm is the development of traditional SVM, which uses hierarchical clustering for SVM to accelerate the processing speed of SVM to large-scale data. Schwenk and Bengio proposed AdaBoost method to control the classification accuracy [25].

The gray level co-occurrence matrix (GLCM) algorithm [26-27] is a popular texture-based feature extraction method, which can reflect the spatial distribution of gray levels and determine the texture relationship between pixels according to the second-order statistics in DBT images [28, 39]. If GLCM algorithm is used to extract the texture features of the DBT image, the global image information that contains information about all parts or all pixels will not be fully utilized, which will result in a large amount of redundant information and storage space. Due to the lack of human visual information, it is difficult to study the pixel dependencies between textures and improve the classification accuracy of the GLCM algorithm [29].

The presentation of Tamura algorithm [30-31] can overcome the shortcomings of the GLCM algorithm on the way of integrating the characteristics of human visual information into the process of feature extraction. Tamura et al. proposed six texture features similar to human visual features on the basis of a number of psychological experiments including coarseness, contrast, directionality, line-likeness, regularity and roughness [32]. The advantage of Tamura algorithm is that the appropriate human vision and texture feature parameters are extracted, which can detailedly describe pathological images.

Generally, many image features are employed by traditional methods for the improvement of classification accuracy. However, with the increase of feature dimension, Hughes phenomenon, namely “dimension disaster”, is often caused [33]. If all the set features are involved in the classification, not only the operation becomes complex and the processing speed is greatly reduced, but also the classification accuracy will be reduced in the case of limited training samples. Therefore, how to select a small number of optimal features from a large number of original features in classification has become very important in the image classification. It is necessary to build a robust model for the feature selection from a large amount of data or the images. Among feature selection techniques, Relief algorithm [34] is a typical filtering feature selection algorithm, which has the characteristics of fast operation speed and strong generalization ability. The basic idea of this method is to assign different weights to features according to the correlation of each feature and category. When the weights are less than a certain threshold, the features will be removed.

In this paper, an algorithm based on the combination of GLCM and Tamura texture features is presented to extract texture features of breast tissues. The features are filtered by feature selection rules and the extracted features are classified by support vector machine. The organization of the paper is as follows. Section 2 introduces the principle and methods including the image pretreatment, GLCM and Tamura texture features. The feature selection, support vector machine, combined texture feature extraction algorithm are also described as the bases for the image processing. Relief algorithm is set as a criterion to select the optimal features. The effectiveness of the proposed method is verified and discussed in Section 3 by numerical experiments. Lastly, the conclusion is made in Section 4.

The Principle and Methods

2.1 Image Pretreatment

The blur, the deformation and noise of the image are needed to be preprocessed before the texture feature extraction. The bilateral filtering method [35] is utilized to achieve smooth filtering of images and protect the edge information of images. The contrast limited adaptive histogram equalization (CLAHE) [36] is performed to divide the image into many unrelated parts and acquire the grayscale value of the image according to the bilinear interpolation. The L0-normimage smoothing algorithm (L0 smooth) [37] is used to removes small non-zero gradients and enhance the edges of the image. In order to obtain the information of the DBT image, the region of interest (ROI) is applied, where the lesion area in the image is marked and the background area is removed [38].

2.2 GLCM Texture Features

The GLCM algorithm is a common method for statistical feature extraction, where the texture is represented by calculating the combined conditional probability $p(i,j,d,\theta)$ between the grayscale levels of the image. Suppose the size of the original image $f(x,y)$ is $M \times N$, the probability $p(i,j,d,\theta)$ can be described by

$$p(i,j,d,\theta) = \#\left\{\left((x,y),(x+dx,y+dy)\right) \mid f(x,y)=i, f(x+dx,y+dy)=j, dx=\cos\theta, dy=\sin\theta\right\}, \quad \textbf{(1)}$$

where # is the number of locations that meet the above requirements, (x,y) is the position of the pixel in the image, i is the gray value of the pixel (x,y), j is the gray value of the pixel $(x+dx,y+dy)$, d is the relative distance between $f(x,y)$ and $f(x+dx,y+dy)$, θ is the angle between the line between $f(x,y)$, $f(x+dx,y+dy)$ and the horizontal axis on the angel of $0°,45°,90°,135°$, respectively. It is found that only 5 texture features are irrelevant including angle second moment, contrast, correlation, entropy and inverse moment [39].

2.3 Tamura Texture Features

Tamura et al. proposed six texture features similar to human visual features on the premise of a large number of psychological experiments, including coarseness, contrast, directionality, line-likeness, regularity, and roughness. However, only three texture features are irrelevant and usually used for image classification.

Coarseness. First, it is necessary to calculate the average brightness of each pixel in a moving window with rectangular area $2^k \times 2^k$.

$$A_k(x,y) = 2^{-2k} \sum_{i=x-2^{k-1}}^{x+2^{k-1}-1} \sum_{j=y-2^{k-1}}^{y+2^{k-1}-1} g(i,j), \quad \textbf{(2)}$$

where (x,y) is the position of the pixel in the image, $g(i,j)$ is the gray value of the pixel (i,j) and k is the range of pixels.

The intensity change of each pixel in the horizontal and vertical directions can be calculated in the form

$$E_{k,x}(x,y) = \left| A_k\left(x+2^{k-1},y\right) - A_k\left(x-2^{k-1},y\right) \right|, \quad \textbf{(3)}$$

$$E_{k,x}(x,y) = \left| A_k\left(x,y+2^{k-1}\right) - A_k\left(x,y-2^{k-1}\right) \right|, \quad \textbf{(4)}$$

The coarseness F_{crs} is determined by the average value of $R_{best}(x,y)$ as

$$F_{crs} = \frac{1}{M \times N} \sum_{x=1}^{M} \sum_{y=1}^{N} R_{best}(x,y), \quad \textbf{(5)}$$

where $R_{best}(x,y)$ is the window and k is to determine optimal window size, $R_{best}(x,y) = 2^k$, $k = 0,1,\cdots,5$.

Contrast. F_{con} is the global metrics of the entire image.

$$F_{con} = \frac{S^2}{\sqrt{\alpha_4}}, \quad \textbf{(6)}$$

where α_4 is the fourth moment of gray and S^2 is the variance of gray.

Directivity. The gradient vector of the pixel is calculated. The modulus and direction of the gradient vector can be represented by Eqs. (7) and (8), respectively.

$$|\Delta G| = |\Delta H| + |\Delta V|, \tag{7}$$

$$|\Delta H|, \tag{8}$$

where the horizontal gradient $|\Delta H|$ equals to the deviation of 3 gray values between the left and right pixels and the vertical gradient $|\Delta V|$ is the deviation of 3 gray values between up and down pixels.

The histogram of the local-marginal probability is defined in the form of

$$Q_\alpha(k) = \frac{N_\alpha(k)}{\sum_{i=0}^{N-1} N_\alpha}, \tag{9}$$

where $N_\alpha(k)$ is the number of pixels when $|\Delta G| \geq t$, $\frac{(2k-1)}{2N} \leq \alpha \leq \frac{(2K+1)}{2N}$ and t is the threshold.

The characteristics of the histogram reflect the strength of the texture direction, which can be denoted as

$$F_{dir} = 1 - rN_P \sum_{p=1}^{N_P} \sum_{\phi \in w_p} \left((k - k_p)^2 Q_I(k) \right), \tag{10}$$

where p is a peak in the histogram, N_P is all the peaks in the histogram, w_p is the distance between the valley floor on both sides of the peak p, k_p is the center position of the crest and r is the normalization factor.

2.4 Feature Selection

In the process of image recognition, there are redundant information among a number of image features. Thus, it is necessary to select the optimal image features [40]. The general selection process shown in Fig. 1 includes the following sections: generating a subset of features, evaluating the quality of feature subset; stopping criteria and verification. The purpose of feature selection is to select a set of optimal features from a set of feature sets and the number of feature sets is more than the optimal features. Generally, the branch and bound selection algorithm [41] and Relief algorithm are used for feature selection.

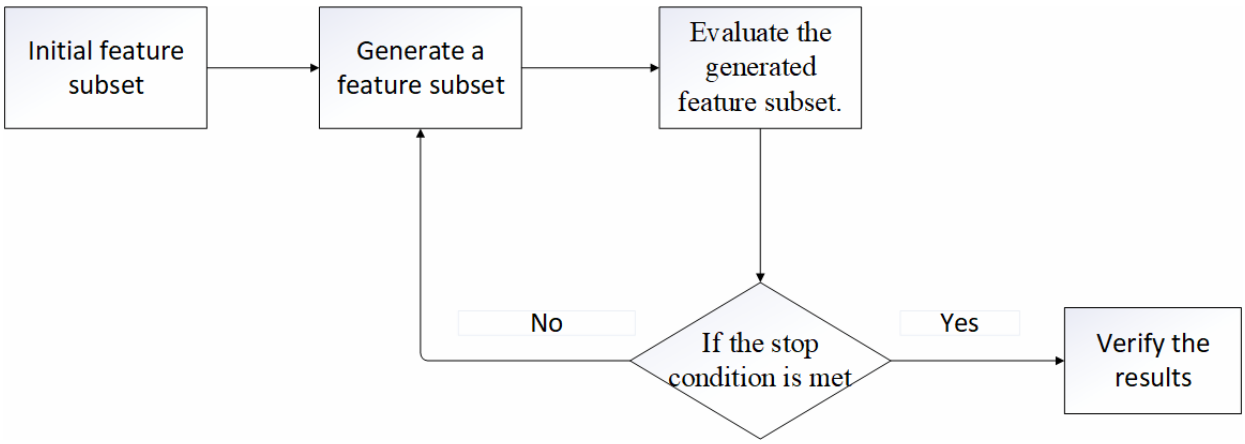


Fig. 1. Flowchart of feature selection

Branch and bound algorithm. The branch and bound algorithm [42] can be used to solve the pure integer planning problem or the mixed integer-programming problem. The feature of the algorithm is that each substep can be carried out depending on the problem, the available software tools and the skill of

the designer of the algorithm. In the algorithm, a branch can be deleted in the case if its bound is not greater than the allowed error. For the solution of an optimization problem, a systematic search is performed, which is called the branch and definition. The iterative division of the solution space into small subsets is called branching. An initial lower bound for the solution set of the minimum value problem is calculated, which is called delimitation. If the target value of a known feasible solution set cannot reach the current limit, the subset is discarded, which is called pruning.

Sequential forward selection. Sequential forward selection (SFS) is a classic feature selection algorithm for high-dimensional images. Sequential forward selection (SFS) algorithm is a bottom-up search procedure, which is a convenient way to gain a combination of features. The SFS method begins with a set of features and a sequential way adding parameters and operates until the criterion of selection has reached a minimum or the parameters are added the model. The feature is selected among the remaining available features, which results in redundant features between the added features and the previous set of features.

Relief feature selection algorithm. Relief algorithm is a typical filtering algorithm, where the process of selecting feature samples and classification algorithm are independent. The Relief algorithm has strong versatility and low computational complexity, which is suitable for feature selection of large amount of feature samples. The principle of Relief algorithm is to allocate different weights to each feature according to the correlation between each feature. When the weight of the feature is smaller than the threshold, it is needed to remove the feature. Relief algorithm is a mature method of feature selection, which can get the best feature subset. D is the training data set, m is the randomly selected number of samples, and n is the number of features in the sample. The principle of Relief algorithm can be described as follows:

Step 1. The feature weights of the sample is initialized to be 0;

Step 2. The sample z_i is randomly selected from the sample set, and its nearest neighbor sample H is selected from the same sample set as the sample z_i, and its nearest neighbor sample K is selected in a sample set different from the category of sample z_i. The distance $dist(z_i, H)$ between z_i and H, and the distance $dist(z_i, K)$ between z_i and K are calculated, respectively.

Step 3. If $dist(z_i, H) < dist(z_i, K)$, the feature can be effectively distinguished between samples of the same category and different categories and the feature weight should be increased. If $dist(z_i, H) > dist(z_i, K)$, the feature can not effectively distinguish between samples of the same category and different categories and the feature weight should be reduced.

Step 4. Step 2 is repeated for m times and n feature weights will be obtained.

2.5 Support Vector Machine

Support vector machine (SVM) is a widely used machine learning algorithm based on statistical learning theory. The linear classification problem is solved by separation hyperplane to classify the training data. The training sample set is $\{(x_i, y_i) | i = 1, 2, \cdots, \mathrm{m}\}$, where $x_i \in R^n$ is the training sample and $y_i \in \{1, -1\}$ is the category of the input sample. The goal of the training is to find an optimal classification plane, which separates the two types of samples and minimizes the generalization error. The optimal classification plane is the hyperplane in the form of

$$f(x) = \omega \cdot x + b, \tag{11}$$

where x is the sample, ω is the weight vector, and b is the classification threshold.

When the data cannot be classified, the hyper-plane is needed to classify the data with the classification interval. The hyper-plane should satisfy the constraint condition $y_i(\omega \cdot x + b) \geq 1$, $i = 1, 2, \cdots, \mathrm{n}$.

The Lagrange function is introduced in the optimization theory and the objective function of SVM can be represented by

$$w(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j (x_i, y_i), \tag{12}$$

where α_i is the Lagrangian coefficient and satisfies the condition $\alpha_i \geq 0, (i = 1, 2, \cdots, n)$. The discriminant function is

$$f(x) = \operatorname{sgn}\left\{\left(\omega^* \cdot b^*\right)\right\} = \operatorname{sgn}\left\{\sum_{x_i \in SV} \alpha^* y_i \left(x_i, \mathrm{x}_j\right) + b^*\right\}, \tag{13}$$

where $\alpha_i^* = \left(\alpha_1^*, \alpha_2^*, \cdots, \alpha_l^*\right)^T$ is the optimal solution and α_i^* is zero for most of samples, b^* is the classification threshold. It is noted that the solution α_i^* determines the optimal classification plane when the value of α_i^* is not zero.

For nonlinear problems, the data in the low-dimensional feature space can be mapped to the high-dimensional space so that the linearly inseparable samples become linearly separable samples in the high-dimensional space.

The symmetric function of the product operation in a space is $k(x, y)$, which satisfies Eq. (14) for both $\varphi(x) \neq 0$ and $\int \varphi(x)^2 dx < \infty$,

$$\iint k(x, y)\varphi(x)\varphi(y)dxdy > 0, \tag{14}$$

where $k(x, y)$ is a kernel function.

When the optimal hyperplane is constructed with linear inseparability, The Eq. (14) can be represented by

$$W(\alpha) = \sum_{i=1}^{n} \alpha_i - \frac{1}{2} \sum_{i=1}^{n} \sum_{j=1}^{n} \alpha_i \alpha_j y_i y_j K\left(x_i \cdot x_j\right), \tag{15}$$

and the discriminant function is

$$f(x) = \operatorname{sgn}\left\{\left(\omega^* \cdot b^*\right)\right\} = \operatorname{sgn}\left\{\sum_{x_i \in SV} \alpha_i^* y_i K\left(x_i \cdot x_j\right) + b^*\right\}. \tag{16}$$

Because the discriminant function of SVM is determined by some support vectors and the complexity of calculation is independent of the dimensions of the feature space, it is an effective method to solve the problem in high-dimensional space.

2.6 Combined Texture Feature Extraction Algorithm

When GLCM algorithm is used to extract the texture features, the global information of the image will not be fully utilized and the characteristics of human visual information are not considered. The presentation of Tamura algorithm can overcome the shortcomings of the GLCM algorithm on the way of taking into account of the characteristics of human visual information. When the texture features are extracted by the Tamura algorithm, the texture features of the global image or local image can be represented.

In this paper, a feature extraction algorithm is proposed, which combines GLCM and Tamura texture features. In the algorithm, the coarse texture features and fine texture features are extracted by Tamura algorithm and GLCM calculation respectively, which means that the extracted coarse texture features are combined with the fine texture features for the image classification. For DBT images, the combined texture feature extraction algorithm can be described as follows.

Step 1. For the original image, the denoising process is performed and the contrast of the gradation is enhanced.

Step 2. The feature extraction starts from the top left corner of the image. The moving window is used to traverse the image. In each moving window, the energy, contrast, correlation, entropy, inverse difference moment (IDM) and correlation of the GLCM algorithm and the features of the Tamura algorithm including the coarseness, contrast and directivity are extracted, respectively.

Step 3. Each set of Tamura texture features is integrated into GLCM texture features to construct an 8-dimensional fused texture feature space.
Step 4. The fused texture features are filtered for the optimal features by Relief algorithm. The weight and threshold of the feature are calculated. The features are retained if the weight is greater than the threshold. Otherwise, the features are removed.
Step 5. The selected texture features are input into the support vector machine for the image classification.

The flowchart of combined texture feature algorithm is shown in Fig. 2.

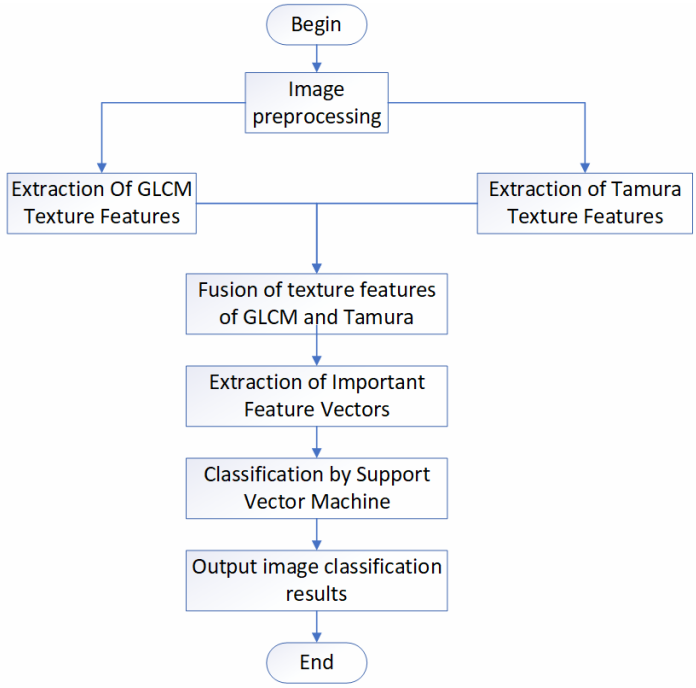


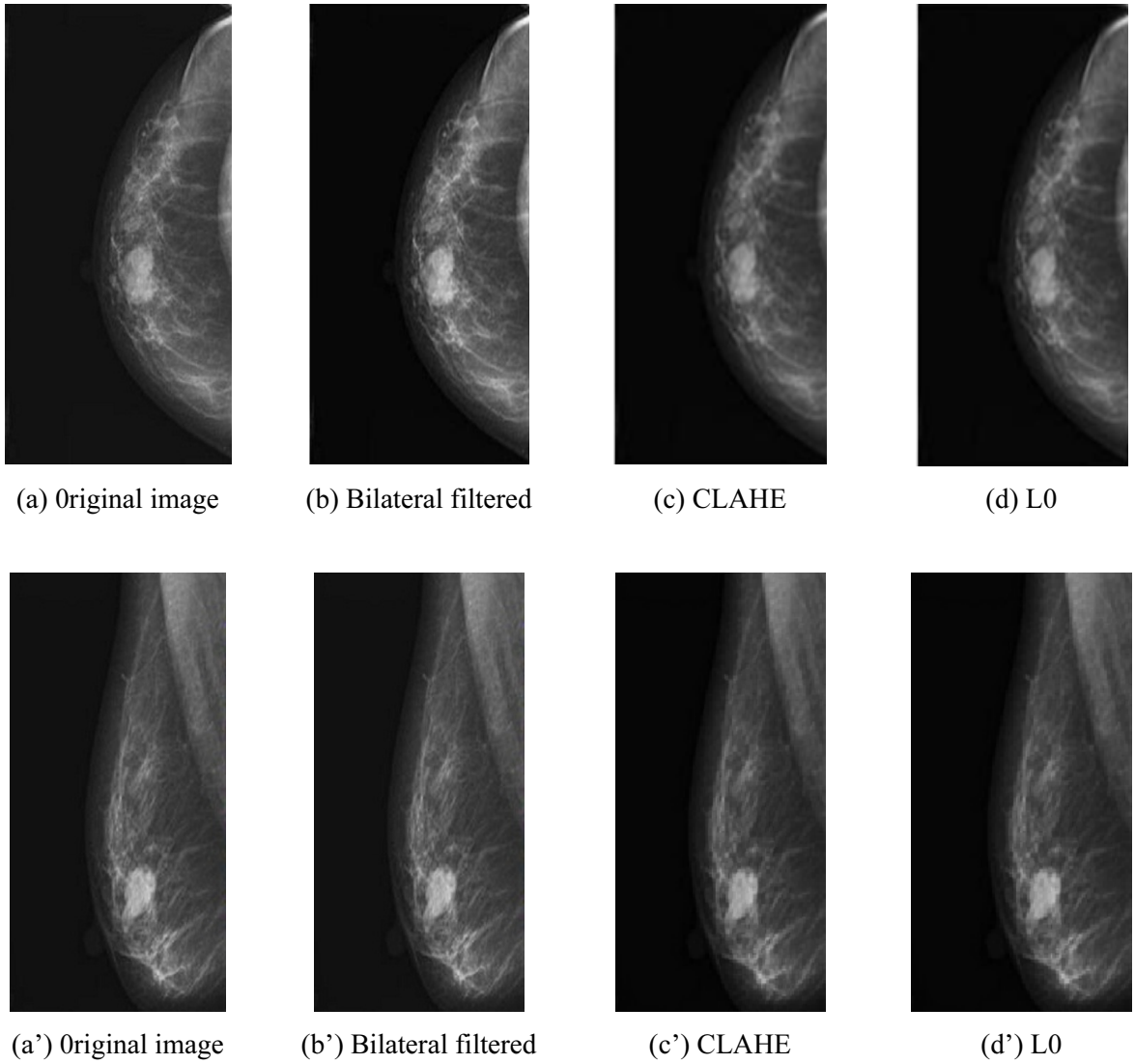
Fig. 2. Flowchart of combined texture feature extraction algorithm

3 Experimental Results

60 images were randomly selected from the DBT image database of the First Affiliated Hospital of Xi'an Jiaotong University, which can be classified as pathological images and normal images. 30 images were randomly selected as training samples and the other images were used as test samples. The classification accuracy of images is analyzed to evaluate the algorithm performance. The texture features of 60 DBT images are extracted by the proposed algorithm. The eigenvalues of 8 texture parameters of each image are obtained and the feature vector is optimally selected by Relief algorithm. The SVM is used to classify the image with a 6×60 matrix, which is composed by the optimal features of 60 images. Lastly, the matrix is input into SVM for the image classification.

3.1 Image Preprocessing

The window size of the bilateral filtering algorithm is set to be 4×4, the variance of the Gaussian function of the spatial domain is δ_θ, $\delta_\theta=2.0$, and the variance of the Gaussian function of the range is δ_γ, $\delta_\gamma=0.1$. The window's size of the CLAHE algorithm is 4×4, and the clipping value $\alpha=6$. In the L0 norm image smoothing algorithm, λ is the smoothness parameter, and β is iteration number, $\lambda=0.002$, $\beta=2.0$. The original image and the processed images are shown in Fig. 3.



(a) 0riginal image (b) Bilateral filtered (c) CLAHE (d) L0

(a') 0riginal image (b') Bilateral filtered (c') CLAHE (d') L0

Fig. 3. Original image and the preprocessed image

After the preprocessing of DBT image, the redundant details of the image are eliminated, the edge features of the image are retained and the contrast of the image is increased. The ROI area of original images is shown in Fig. 4. It is noted that the size of ROI area is 60×120 pixel.

(a) (b)

Fig. 4. Extraction of ROI area

3.2 Feature Extraction and Selection

The GLCM texture features of the ROI region are extracted. The size of the extraction window is 32×32, the step size is $d = 1$. The direction angle θ is given as $0°, 45°, 90°, 135°$, respectively. The mean and standard deviation of the texture features are calculated.

The extraction results of ROI area in Fig. 4(a) and Fig. 4(b) are provided in Table 1 and Table 2.

Table 1. GLCM texture parameter of ROI regions in Fig. 4(a)

Angle	Energy	Contrast	Entropy	IDM	Correlation
0°	1.3813	4.4961	1.0209	0.3510	1.2537
45°	0.0022	2.4923	0.1695	0.0182	1.2913
90°	1.1054	4.4375	1.0149	0.4121	1.3398
135°	0.0012	2.7350	0.1540	0.0621	1.2752
Mean	0.6225	3.0402	0.5898	0.2108	1.2900
Standard Deviation	0.5266	0.2530	0.0069	0.0397	0.0013

Table 2. GLCM texture parameter of ROI regions in Fig. 4(b)

Angle	Energy	Contrast	Entropy	IDM	Correlation
0°	0.2797	2.1207	3.9518	0.1015	0.1242
45°	0.1531	7.3965	4.5048	0.1166	0.0670
90°	1.2901	5.4718	4.2236	0.1433	0.0878
135°	0.1527	7.1999	4.4914	0.0919	0.0692
Mean	0.2189	5.5472	4.2929	0.1133	0.0871
Standard Deviation	0.0763	2.4425	0.1616	0.0005	0.0264

The Tamura texture features of the ROI region are extracted. The size of the extraction window is 32×32. The results can be found in Table 3.

Table 3. Tamura texture parameter of ROI regions

Window Size \ Texture Features	Coarseness	Contrast	Directivity
Fig. 4(a)	7.0816	29.064251	20.8766
Fig. 4(b)	9.4626	31.605051	16.7393

5 GLCM texture features and their mean values and 3 Tamura texture features are integrated into an 8-dimensional feature vector. The feature vectors of the 60 images are combined into an 8×60 matrix, which is used as the categorical dataset. If the feature dimension of the matrix is large, redundant information will be generated, which increases the program running time and reduces the classification accuracy. Therefore, it is necessary to select the features of the image and reduce the feature dimension based on Relief algorithm, which uses the weight to classify the performance of the feature representation. The classification weight is compared to a threshold if the classification weight is greater than the threshold. There are six optimal features obtained by Relief algorithm for feature selection. The six optimal features are energy, entropy, inverse difference moment (IDM), correlation, coarseness and contrast.

3.3 Results of Classification Accuracy Based on Four Different Methods of SVM

After the texture features of DBT images are extracted by the GLCM [43] and the Tamura [44] algorithm respectively, the feature features are input into the SVM [45] for the image classification. Table 4 describes the classification results of five texture feature extraction method based on support vector machines. The first method is to classify the texture features by traditional SVM. The second and third methods are to extract texture features using the GLCM algorithm and the Tamura algorithm, respectively. The fourth method is to extract the texture features by a combined of GLCM and Tamura algorithms. The fifth method is the use of Relief algorithm to select the optimal texture features after the processing of GLCM-Tamura algorithm, which can be described as FSGT algorithm. The texture

features obtained by the GLCM, Tamura, GLCM-Tamura and FSGT algorithms are input into the SVM for classification.

Table 4. Results of different feature extraction methods

Method	Classification Accuracy/%	Running time/s
SVM	60.87	17.8
GLCM	76.60	24.9
Tamura	56.32	25.6
GLCM-Tamura	88.23	36.5
FSGT	90.00	26.5

The process of the FSGT algorithm is given as follows. Firstly, the features are extracted by the mixed texture feature extraction algorithm, that is, the texture features of GLCM and Tamura are extracted respectively, and then their texture features are fused to build the texture eigenvectors. Then, the classification weight of each feature of the sample is calculated by Relief algorithm, where a set of characteristics with the maximum weight is selected and the features with large weight are retained as the output features. In the redundancy analysis of features, the threshold of correlation and weight of classification features is 0.9 and 2500, respectively. Finally, the eigenvectors are input into SVM for classification. The classification accuracy and the running time of each method are shown in Table 4.

The conclusions can be made as follows.

(1) The average classification accuracy of SVM, GLCM, Tamura, GLCM-Tamura, and FSGT is 60.87%, 76.60%, 56.32%, 88.23%, and 90.00% respectively. The average classification accuracy of GLCM-Tamura is higher than that of GLCM and Tamura, which indicates that the algorithm combining GLCM and Tamura texture features can improve the classification accuracy and obtain better classification results of DBT images. The average classification accuracy of FSGT is higher than GLCM-Tamura, which proves that the optimized feature selection can be achieved by Relief algorithm.

(2) The Running time of SVM, GLCM, Tamura, GLCM-Tamura, and FSGT is 17.8s, 24.9s, 25.6s, 36.5s, and 26.5s, respectively. It can be seen that GLCM-Tamura has the lowest computational efficiently. That is because the dimension of the feature vector of GLCM-Tamura is increased and a large amount of data operation is needed. The running time of FSGT is higher than that of GLCM-Tamura, which indicates that the feature selection method can effectively filter feature vectors, select the optimal features and reduce the amount of data.

4 Conclusion

A texture feature extraction algorithm is proposed to improve the classification accuracy of DBT images. Based on bilateral filtering, contrast-limited adaptive histogram equalization and L0 gradient filtering, the DBT image is preprocessed to filter out the noise and improve the grayscale contrast. The characteristics of human visual information are integrated into the texture feature extraction, where the extracted coarse texture features are combined with the fine texture features for the image classification. The Relief algorithm is used to filter and obtain an optimal feature subset and the SVM is used to classify feature vectors. The results show that the proposed algorithm can improve the accuracy of feature extraction and recognition and the features can effectively represent the texture of breast cancer. Compared with the traditional SVM, the GLCM, the Tamura and GLCM-Tamura algorithms, the image classification accuracy and efficiency of the proposed feature extraction method have been improved. It is promising that the proposed method can help doctors find potential tumors, effectively improve the early diagnosis of breast cancer and has great significance for the cure of breast cancer.

Acknowledgements

This work is supported by the Key Research and Development Program of Shaanxi Province of China (2018GY-018, 2019GY-086) and the New Star Team of Xi'an University of Posts and Telecommunications.

References

[1] X. Qin, G.I. Lu, Sechopoulos, Breast tissue classification in digital tomosynthesis images based on global gradient minimization and texture features, in: Proc. SPIE- the International Society for Optical Engineering, 2014.

[2] J.R. Quinlan, Induction of decision trees, Machine Learning 1(1986) 81-106.

[3] Z.J. Wang, A new way to choose splitting attribute in ID3 algorithm, in: Proc. 2017 IEEE 2nd Information Technology, Networking, Electronic and Automation Control Conference (ITNEC 2017), 2017.

[4] J.R. Quinlan, C4.5: Programs for Machine Learning, Morgan Kaufmann, San Mateo, CA, 1993.

[5] B. Rawal, R. Agarwal, Improving accuracy of classification based on C4.5 decision tree algorithm using big data analytics, in: H. Behera, J. Nayak, B. Naik, A. Abraham (Eds.), Computational Intelligence in Data Mining, Springer, Singapore, 2018, pp. 203-211. http://doi.org/443.webvpn.fjmu.edu.cn/10.1007/978-981-10-8055-5_19

[6] L. Breiman, J. Friedman, R. Olshen, Classification and Regression Trees. Wadsworth International Group, Monterey, CA, 1984.

[7] J.S. Liang, W. Zheng, Z.B. Yuan, Well group connectivity relations discriminate based on CART algorithm, Applied Mechanics and Materials 513-517(2014) 1252-1255.

[8] D.E. Rumelhart, G.E. Hinton, R.J. Williams, Learning internal representations by error propagation, in: D.E. Rumelhart, J.L. McClelland, CORPORATE PDP Research Group (Eds.), Parallel Distributed Processing: Explorations in the Microstructure of Cognition, Vol. 1: Foundations, MIT Press, Cambridge, MA, 1986, pp. 318-362.

[9] J. Meng, L. Wang, Research on the intrusion detection based on the improved BP algorithm, in: Proc. 2012 Fourth International Conference on Computational and Information Sciences, 2012.

[10] P. Langley, W. Iba, K. Thompson, An analysis of Bayesian classifiers, in: Proc. Tenth National Conference on Artificial intelligence, 1992.

[11] T.M. Cover, P.E. Hart, Nearest neighbor pattern classification, IEEE Transactions on Information Theory 13(1967) 21-27.

[12] Z. Pawlak, Rough classification, International Journal of Man-Machine Studies 20(5)(1984) 469-483.

[14] L.A. Zadeh, Fuzzy Sets, Fuzzy Logic, and Fuzzy Systems, World Scientific, River Edge, NJ, 1996.

[15] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, New York, NY, 2000.

[16] J.T. Yao, Y.Y. Yao, A granular computing approach to machine learning, in: Proc. the 1st International Conference on Fuzzy Systems and Knowledge Discovery, 2002.

[17] Y.Y. Yao, J.T. Yao, Induction of classification rules by granular computing, in: Proc. the Third International Conference on Rough Sets and Current Trends in Computing, 2002.

[18] Y.Y. Yao, J.T. Yao, Granular computing as a basis for consistent classification problems, in: Proc. PAKDD'02 Workshop on Toward the Foundation of Data Mining, 2002.

[19] B. Liu, W. Hsu, Y. Ma, Integrating classification and association rule mining, in: Proc. of the Fourth International Conference on Knowledge Discovery and Data Mining (KDD-98), 1998.

[20] K. Wang, S. Zhou, Y. He, Growing decision tree on support-less association rules, in: Proc. the Sixth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining (KDD'00), 2000.

[21] W. Li, J. Han, J. Pei, CMAR, Accurate and efficient classification based on multiple class-association rules, in: Proc. 2001 IEEE International Conference on Data Mining, 2002.

[22] N. Ye, X. Li, A machine learning algorithm based on supervised clustering and classification, in: Proc. International Computer Science Conference on Active Media Technology, 2001.

[23] H. Yu, J. Yang, J. Han, Classifying large data sets using SVMs with hierarchical clusters, in: Proc. the Ninth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, 2003.

[24] G.M. Xian, An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM, Expert Systems With Applications 37(10)(2010) 6737-6741.

[25] A. Schnall, M. Heckmann, Feature-Space SVM Adaptation for Speaker Adapted Word Prominence Detection, Computer Speech & Language 53(2018), DOI: 10.1016/j.csl.2018.06.001.

[26] H. Schwenk, Y. Bengio, Boosting neural networks, Neural Computation 12(8)(2000) 1869-1887.

[27] R.M. Haralick, K. Shanmugum, I.H. Dinstein, Textural features for image classification, IEEE Transactions on Systems, Man and Cybernetics 8(6)(1973) 610-621.

[28] J. Zhang, G.L. Li, S.W. He, Texture-based image retrieval by edge detection matching GLCM, in: Proc. IEEE International Conference on High Performance Computing & Communications, 2008.

[29] F. Albregtsen, B. Nielsen, H.E. Danielsen, Adaptive gray level run length features from class distance matrices, in: Proc. 15th International Conference on Pattern Recognition, 2000.

[30] M. Arebey, M.A. Hannan, R.A. Begum, H. Basri, Solid waste bin level detection using gray level cooccurrence matrix feature extraction approach, Journal of Environmental Management 104(2012) 9-18.

[31] Y.H. Xie, J.C. Wang, Study on the identification of the wood surface defects based on texture features, Optik - International Journal for Light and Electron Optics 126(19)(2015) 2231-2235.

[32] H.J. Tao, X.B. Lu, Smoky vehicle detection based on multi-scale block Tamura features, Signal, Image and Video Processing 12(6)(2018) 1-8.

[33] H. Tamura, S. Mori, T. Yamawaki, Textural features corresponding to visual perception, IEEE Transactions on Systems, Man and Cybernetics 8(6)(1978) 460-473.

[34] S.A. Medjahed, M. Ouali, Band selection based on optimization approach for hyperspectral image classification, The Egyptian Journal of Remote Sensing and Space Science 21(3)(2018) 413-418.

[35] L. Gao, T.T. Li, L.Z. Yao, F. Wen, Research and application of data mining feature selection based on relief algorithm, Journal of Software 9(2)(2014) 515-520.

[36] N. Rajpoot, I. Butt, A multiresolution framework for local similarity based image denoising, Pattern Recognition 45(8)(2012) 2938-2951.

[37] S. Wu, Q. Zhu, Y. Yang, Feature and contrast enhancement of mammographic image based on multiscale analysis and morphology, in: Proc. IEEE International Conference on Information and Automation, 2013.

[38] X. Cheng, M. Zeng, X. Liu, Feature-preserving filtering with L0 gradient minimization, Computers & Graphics 38(1)(2014) 150-157.

[39] M. Xian, Y.T. Zhang, H.D. Cheng, Fully automatic segmentation of breast ultrasound images based on breast characteristics in space and frequency domains, Pattern Recognition 48(2)(2015) 340-355.

[40] G. M. Xian, An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM, Expert Systems With Applications 37(10)(2010) 6737-6741.

[41] S. Vanaja, K.R. Kumar, Analysis of feature selection algorithms on classification: a survey, International Journal of Computer Applications 96(17)(2014) 28-35.

[42] S.L. Sun, Z. Chen, Z.L. Liu, An Improved Branch and Bound Algorithm, Software, 2011.

[43] A.A. Kadri, I. Kacem, K. Labadi, A branch-and-bound algorithm for solving the static rebalancing problem in bicycle-sharing systems, Computers & Industrial Engineering, 95(2016). DOI: 10.1016/j.cie.2016.02.002.

[44] Ş. Öztürk, B. Akdemir, Application of feature extraction and classification methods for histopathological image using GLCM, LBP, LBGLCM, GLRLM and SFTA, Procedia Computer Science 132(2018) 40-46.

[45] J.F. Jing, H.H. Zhang, J. Wang, P. Li, J. Jia, Fabric defect classification based on local binary patterns and Tamura method, Advanced Materials Research 4(2012) 562-564.

[46] N. Cristianini, J. Shawe-Taylor, An Introduction to Support Vector Machines, Cambridge University Press, New York, NY, 2000.