Application of SVM Based Optimization Algorithm in Medical Images Processing

Huiyan Jiang, Lingbo Zou
Software College
Northeastern University
Shenyang, P.R.China
{hyjiang@mail.neu.edu.cn, lingbo.zou.2009@gmail.com}

ABSTRACT: *Medical image processing is a very significant area of research as it has far reaching impact in health system. For pattern recognition, the Support vector machine is applied increasingly. The parameters selection plays a significant role in support vector machine(SVM). Through this research we have introduced a better scalable parameter optimization method based on traditional PSO optimizing algorithm by changing the fitness function in the traditional process. We have experimented it in real life and we came out with promising results which reflected in the ROC curves in medical images classification.*

Keywords: Medical Images Processing, Support Vector Machines, Fitness Function, ROC Curve

Received: 12 November 2011, Revised 27 December 2011, Accepted 31 December 2011

1. Introduction

Research on Support Vector Machine use in pattern recognition [1][2][3] and image processing in increasing currently. Support vector machine (SVM) uses a device called kernel function to map data in input space to higher dimensional feature space so that these data are linearly separable. There are various available kernels such as Gaussian, polynomial kernel and so on. Nowadays, support vector machines have been widely used in the area of image processing and artificial intelligence, because of its superior performance in classification and regression analysis.

Since the theory of support vector machines is proposed by Vapnik etc.[4] in the year of 1995, the method has been improved many times. As is recognized, support vector machine is promise for classification, especially in images classifications, many researches are focused on these areas recent years. In SVM kernels, there are always some parameters which affect the performance and generalization ability of SVM. Many kinds of parameters optimization methods have been used.

Shuang Liu etc.[5] has proposed a method that combines the support vector machine with genetic algorithm and applied it in IRIS dataset, which received a better classification accuracy. In 2004, L.Zhang, etc.[6] put forward Morlet wavelet kernel SVM and applied this model to the recognition of 2-D images of radar targets. Their results verify that the training speed of the new SVM model is slightly faster than the Gaussian kernel SVM, at the same time, it has higher classification accuracy. After that, Chong Liu etc.[7] proposed a Hyper-sphere SVM model and applied it in fault diagnosis of analogue circuits analysis. Then, Huiyan Jiang etc. [8] uses adaboost method to accelerate the speed of their support vector machine and greatly accelerated the speed of the model. In 2010, Xiangying Liu etc.[9] has applied an ant colony optimization algorithm on their support vector machine model to optimize the parameter of the SVM and the model have got a acceptable result in medical images classification. What's more, in previous research, grid search method[10] was used by [10] H.H.Gao to select C and σ for RBF kernel. Particle swarm optimization algorithm was also used in X.Y. Zhang and Y.L. Guo's paper[11] to optimize the parameter of support vector

machines. S.Z.Wang etc.[12] chose parameters for Marr wavelet kernel by virtue of modified chaotic particle swarm optimization. Finally, Huiyan Jiang etc.[13] improved the kernel function of SVM and optimized the parameters of their model by particle swarm optimizing method and achieved excellent results in medical images classification.

This paper is organized as follows. First, some basic theories and related work of PSO-SVM is introduced. Second, an improved PSO-SVM model is discussed in detail and some experimental results are given. At last, some useful conclusions are drawn.

2. Support vector Machines Classifier

The principle of SVM classifier is to construct an optimal separating hyper plane by maximizing the margin between different testing samples and the samples on the classification margin is called support vectors.

If the input dataset can be represented as $\{x_1, x_2,\ldots, x_m\}$, where m is the number of samples. Take the medical image samples as an example; we can assume that the samples have two classes. It is demonstrated that a separating hyper plane can be given as Equation (1).

$$f(x) = w^T x + b \tag{1}$$

In the formula above, w and b is used to determine the location and angle of the separating hyper plane. When determining w and b, a constraint must be set to ensure the effectiveness of the parameters. Then, the optimal hyper plane that separates the samples can be obtained by the following optimization problem:

$$\min \frac{1}{2}\|w\|^2 + C\sum_{i=1}^{m}\xi_i \; subject\ to\ y_i\, f(x_i) = y_i(w^T x_i + b) \geq 1 - \xi_i,\ \xi_i \geq 0 \ \ for\ i = 1, 2, ..., M \tag{2}$$

In the formula above, C is the error penalty parameter, and ξ_i is the slack variable. In solution of the problem, the optimization problem above can transfer to the form of Equation (3), a dual optimization problem.

$$\min \frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l} y_i\, y_j\, \alpha_i \alpha_j - \sum_{j=1}^{l}\alpha_j \; s.t. \sum_{i=1}^{l} y_i \alpha_i = 0,\ 0 \leq \alpha_i \leq C,\ i = 1, ..., l \tag{3}$$

After that, the decision function is shown in Equation (4). The positive class is labeled as +1, and the negative class is labeled as -1 correspondingly.

$$f(x) = sgn\left(\sum_{i=1}^{M} y_i\, \alpha_i\ (x_i \bullet x_j) + b\right) \tag{4}$$

Then, consider the situation of nonlinear classification, by using mapping function $\varphi(x)$, we map the original data into a higher dimensional feature space. The decision function can be written as Equation (5).

$$f(x) = sgn\left(\sum_{i=1}^{M} y_i\, \alpha_i K(x_i, x_j) + b\right) \tag{5}$$

Where, $K(x,x)$ must satisfy the Mercer's condition. The Mercer's condition is:

Theorem 1: any symmetric function in the input space can represent an inner product in feature space if

$$\iint K(x, x')g(x)g(x')dxdx' \geq 0,\ \forall g \neq 0 \text{ for which } \int g^2(\xi)d\xi \prec \infty \tag{6}$$

In Equation (6), g(x) is square-integrable function. Then $K(x,x)$ can be written as

$$K(x,x') = \langle \phi(x) \,.\, \phi(x^1) \rangle \tag{7}$$

There are several kinds of kernel in SVM have been made in use. In the training of support vector machines, the kernel selection is very important. A good kernel can help us map the sample vectors into higher dimensional space and solve the nonlinear problem easily. According to Mercer's theorem, the kernel function in SVM must be continuous and positive definite. Several forms of kernel functions widely used are shown in Table 1.

Name	$K(x, x)$
Linear kernel	$K(x, x) = x_i^T \bullet x_j$
Polynomial kernel	$K(x, x) = (\lambda x_i^T \bullet x_j + \gamma)^d$
Gaussian kernel	$K(x, x) = \exp(-\|x_i - x_j\|^2 / 2\sigma^2)$
Sigmoid kernel	$K(x, x) = tanh(\upsilon (x_i \bullet x_j) + c)$

Table 1. Several common kernels in use

3. Pso-svm optimization

3.1 Pso Algorithm Principle

Particle Swarm Optimization (PSO) algorithm is an important branch of swarm intelligence, which was proposed by J.Kennedy and R.C.Eberhart in 1995. It is widely used in the field of digital image processing, pattern recognition and operational research. PSO algorithm is based on swarms, which moves the individuals in the swarms to the better region according to its fitness towards the environment. In a D-dimensional search space, we regard every individual as a particle X_i, which has no volume. Then the position of X_i can be represented as $X_i = (x_{i1}, x_{i2}, \ldots, x_{iD})$, the best position of X_i can be represented as $V_i = (v_{i1}, v_{i2}, \ldots, v_{iD})$.

Accordingly, the flying velocity of each particle can be represented as $P_i = (p_{i1}, p_{i2}, \ldots, p_{iD})$, and $P_g = (p_{g1}, p_{g2}, \ldots, p_{gD})$ represents the best previous position of the swarm.

For every generation, velocity and position are updated each time according to the formula below.

$$\begin{cases} v_{i+1} = wv_i + c_1 r_1 (p_i - x_i) + c_2 r_2 (p_g - x_i) \\ x_{i+1} = x_i + v_i \end{cases} \quad (8)$$

In the formula, *i* represents iteration number; *w* is inertia weight; c_1, c_2 is learning factor; r_1 and r_2 are two random numbers in the range of [0,1].

For every generation, we use fitness function to determine best position of each particle and the best position of the whole swam.

The fundamental logical process of PSO method is shown in form of pseudocode as follows:

```
For each particle
    Initialize particle
END
Do
    For each particle
        Calculate fitness value
        If the fitness value is better than the best fitness value(pBest) in history
        Set current value as the new pBest
    END
    Choose the particle with the best fitness value of all the particles as the gBest
    For each particle
        Calculate particle velocity according equation
        Update particle position according equation
    END
While maximum iterations or minimum error criteria is not attained
```

3.2 PSO SVM Optimization

SVM is a useful tool in statistical learning theory. The parameter selection has a significant impact on the classification result i.e. the identification accuracy of support vector machines. However, how to select the best parameters for various SVM model is hard to determine and experts have a variety of opinions on such a key problem in pattern recognition.

Now, let us consider combining the particle swarm optimizing algorithm with the support vector machines. Our goal is to optimize the parameters of the support vector machine classifier by updating the every generation of particles, which we regard as a group of key parameters for SVM. And we use the classification accuracy of each group of particles as the fitness function; finally, we can easily get the trained model parameters for our support vector machine.

The steps of PSO parameter optimization algorithm are as follows:

*Step*1*: Initialize a group of particles, consider using the RBF kernel, the group is initialized as* $\{C, \sigma\}$*;*

*Step*2*: Calculate every fitness* [*i*] *for every group, that is the classification accuracy of the current parameters;*

*Step*3*: Find* P_i *and* P_g*;*

*Step*4*: Update position and velocity for each particle according to Equation* (8)*;*

*Step*5*: Recalculate every fitness* [*i*] *for every group, update* P_i *and* P_g*;*

*Step*6*: If terminate condition not satisfied, return to Step*3*;*
Else, finish algorithm.

4. Improved Pso-svm based image processing

4.1 Evaluation Indexes For Classifier

A good evaluation method is critical in medical images classification and clinical diagnose. It determines the goodness of the fitness function of parameters optimization method. There are several common evaluation methods of classifier. Consider crossing matrix as shown in Table 2, No. of actual positive samples is $P=TP + FN$, No. of actual negative samples is $N=FP + TN$. We can take full account of these indexes and build a new evaluation criteria relating to our fitness function.

4.2 ROC curve in image processing

Actual value	**Predicted positive examples(+1)**	**Predicted negative examples(-1)**
Positive examples(+1)	Correct positive examples (TP)	False negative examples (FN)
Negative examples(-1)	False positive examples (FP)	Correct negative examples (TN)

Table 1. No.of training and testing samples

From the Table, we can define True Positive Rate and False Positive Rate to form the ROC curve. The bigger the area under the curve is, the better result we will achieve.

4.3 Finess function optomization

In medical images processing and clinical diagnosis, a critic thing is to reduce the rate of misdiagnosis. It is unscientific to evaluate an algorithm only according to its classification accuracy. In large amount of experiments, we find that the ratio of True Positive Rate and False Positive Rate can accurately represent the accuracy of clinical diagnosis. The graphical representation is called receiver operating characteristic curve.

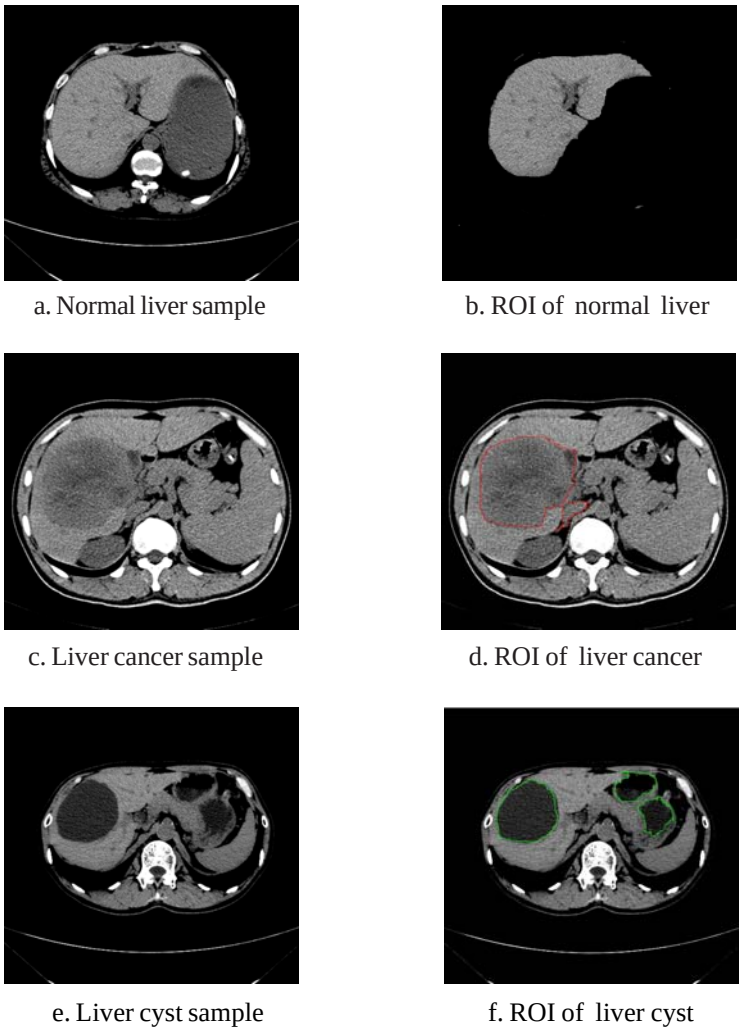
Consider combining the ratio mentioned above with particle swarm algorithm, in the training of support machines, we can use the ratio as the fitness function of PSO optimization, we can demonstrate there is better result in our algorithm. The fitness function can be written as Equation (9).

$$fitness = \frac{True\ Positive\ Rate}{False\ Positive\ Rate} = \frac{\frac{TP}{TP + FN}}{\frac{FP}{FP + TN}} \qquad (9)$$

5. Experiment results and analysis

Most of our experiment data can be find in the share library of Chih-Jen Lin, National Taiwan University. Other data source can be found in the patient library in Medical Multimedia Information Tech. Laboratory, Northeastern University, China.

Take the CT images of liver cancer and liver cancer as original input, and get the lesion. The result of image segmentation and lesion extraction is shown in Figure 1.



a. Normal liver sample b. ROI of normal liver

c. Liver cancer sample d. ROI of liver cancer

e. Liver cyst sample f. ROI of liver cyst

Figure 1. Images Segmentation Results

After we obtain the interest region of the normal liver and the lesion of the samples of liver cancer and liver cyst, we extract the feature vectors. We get the region of every kind of disease and we have find texture changes in all of these pictures. The feature vectors are: gray value, Fourier energy spectrum, gradient matrix, texture energy, gray difference, and some second-order statistical features etc.

As variables have certain correlations between each other, the dimension should be reduced. Here we use principal component analysis. The principle of this method is simple. We get the correlation matrix of the features and calculate the eigenvalues, then the key elements can form a principal component. Then, we can get the new feature vectors according to the principal components. Finally, we get 5 groups of data in different diseases and divide them randomly into training samples and testing samples. The selection of training and testing samples is shown in Table 3. And the dimension of their feature vectors is 28, 8, 13, 6 respectively.

Group of samples	Training samples	Testing samples
Liver cyst-normal	53	24
Diabetes-normal	574	194
Heart disease-normal	170	100
Liver cancer-normal	92	41

Table 3. No.of training and testing samples

Then, we use our method to optimize the process of SVM train, and compare our method with traditional PSO method. In most experiment, we choose the radial-basis function as the kernel function. The parameter C for every disease is shown in Table 4, and the parameter ó for every disease is shown in Table 5. For every kind of disease, we train different parameters for the identification model.

Group of samples	PSO-SVM	Improved PSO-SVM
Liver cyst-normal	0.78	34.4337
Diabetes-normal	45.8236	99.9583
Heart disease-normal	45.7734	54.7654
Liver cancer-normal	48.4783	32.9941

Table 4. Parameter *C* for Different Diseases

Group of samples	PSO-SVM	Improved PSO-SVM
Liver cyst-normal	0.01	0.01
Diabetes-normal	0.01	0.01
Heart disease-normal	0.2543	0.01
Liver cancer-normal	0.2605	0.01

Table 5. Parameter Σ for Different Diseases

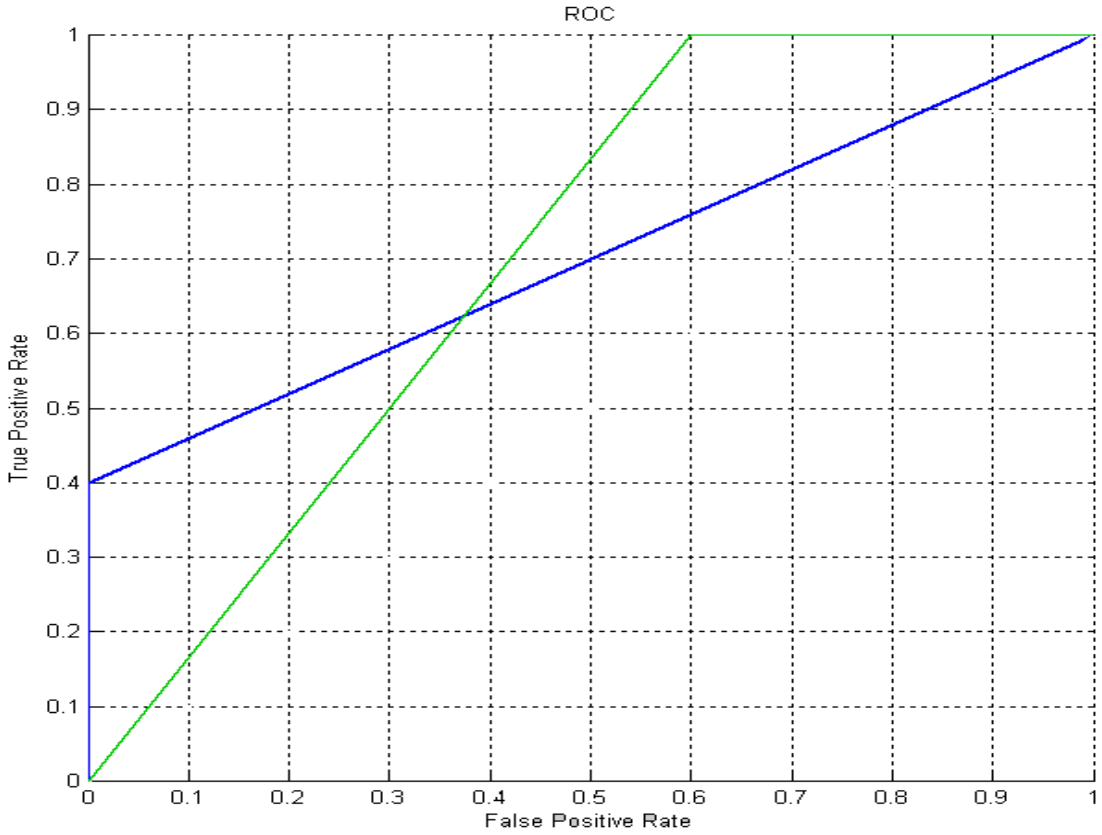
After parameters selection using and our proposed improved method, we classify the samples into two classes.
The classification result of liver cyst, diabetes and liver cancer is in the form of ROC curves as follows:

In Figure.2, comparing *a* with *b*, we can find that the area under the curve in figure a is obviously larger that in figure b. It indicates that our proposed method has better performance than traditional PSO-SVM method. There are same conclusion in the results of liver cancer, diabetes and heart disease.

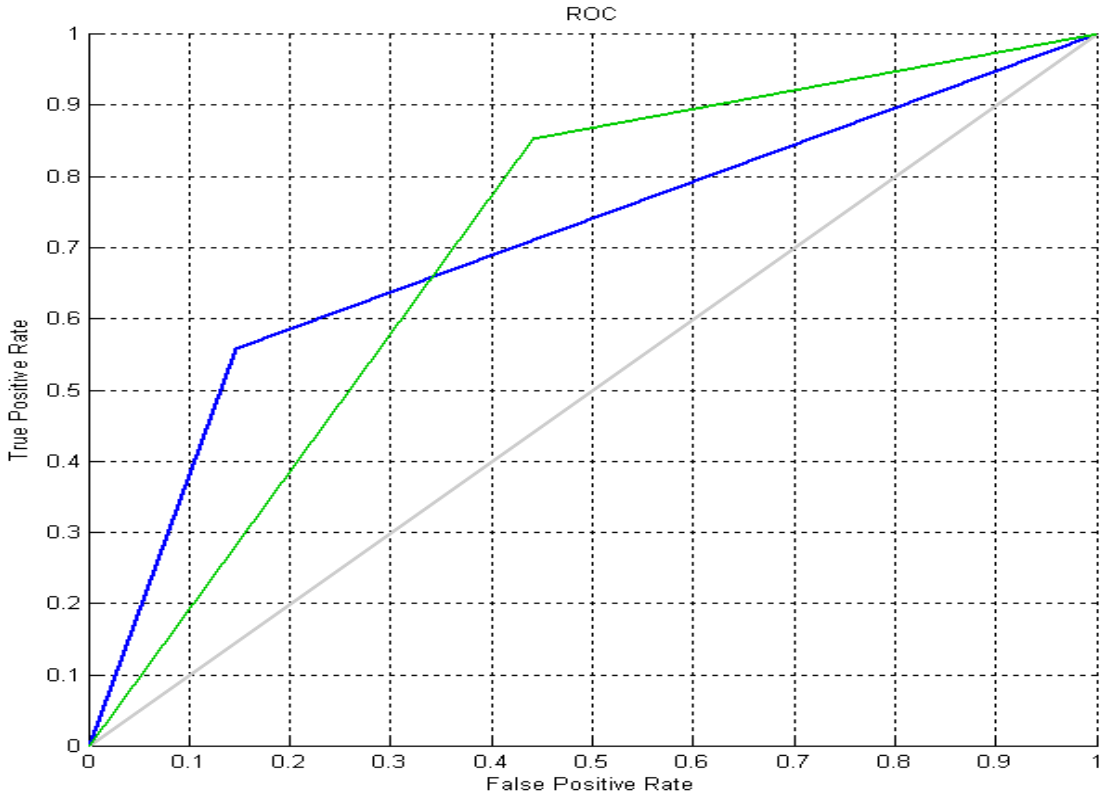
For accurately measurement, we calculate the areas under the curve for each picture above and list them in the table below:

Group of samples	PSO-SVM	Improved PSO-SVM
Liver cyst-normal	63.7	65.0
Diabetes-normal	65.0	66.0
Liver cancer-normal	61.0	63.5

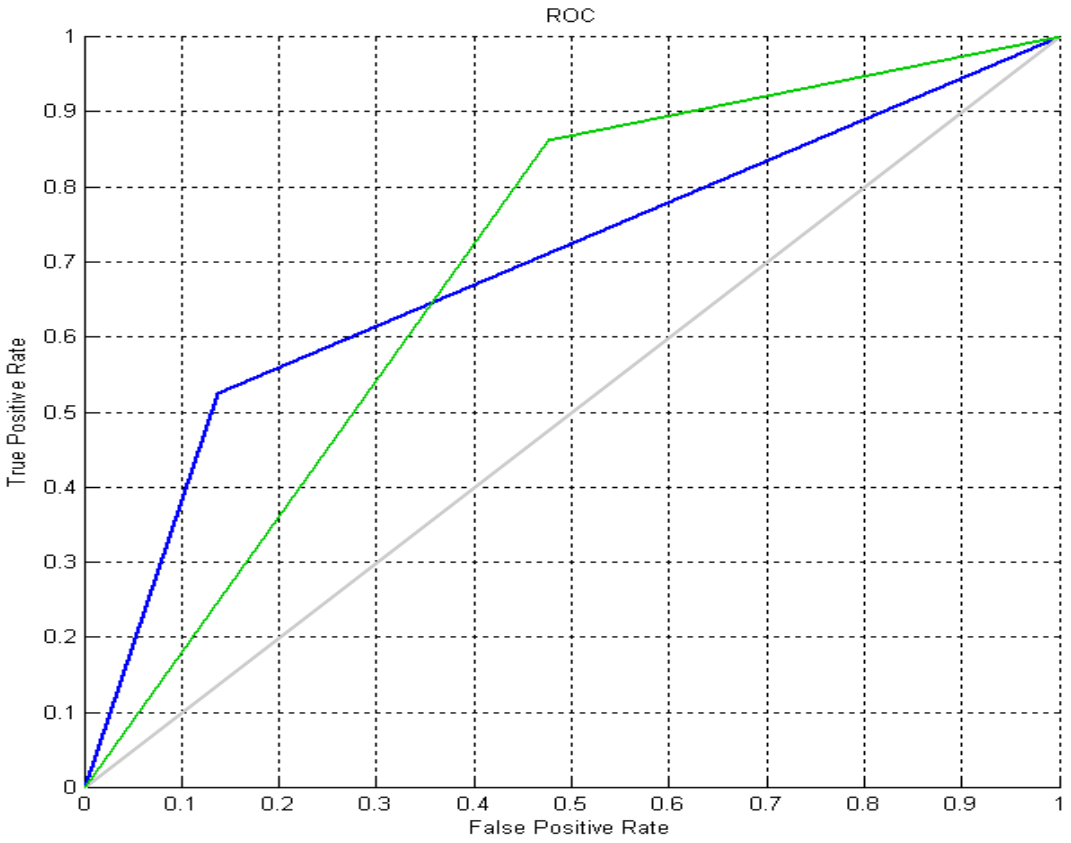
Table 6. Area under the curve for samples(grid units)



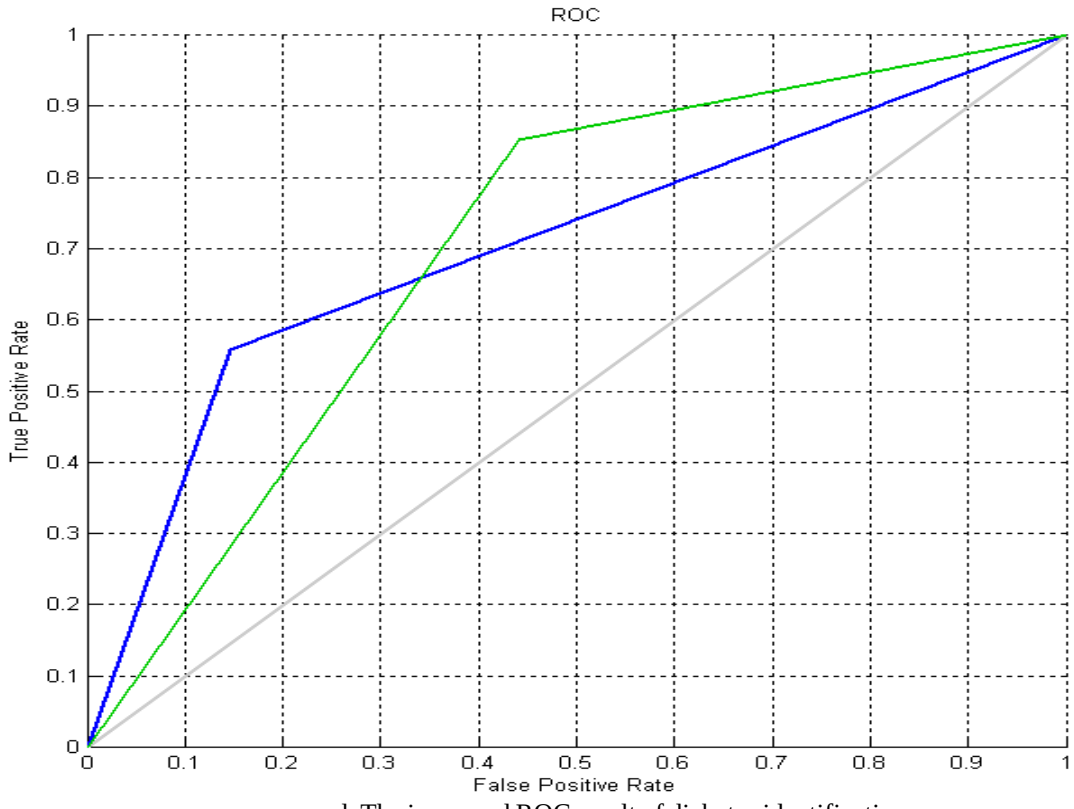
a. The ROC result of liver cyst identification



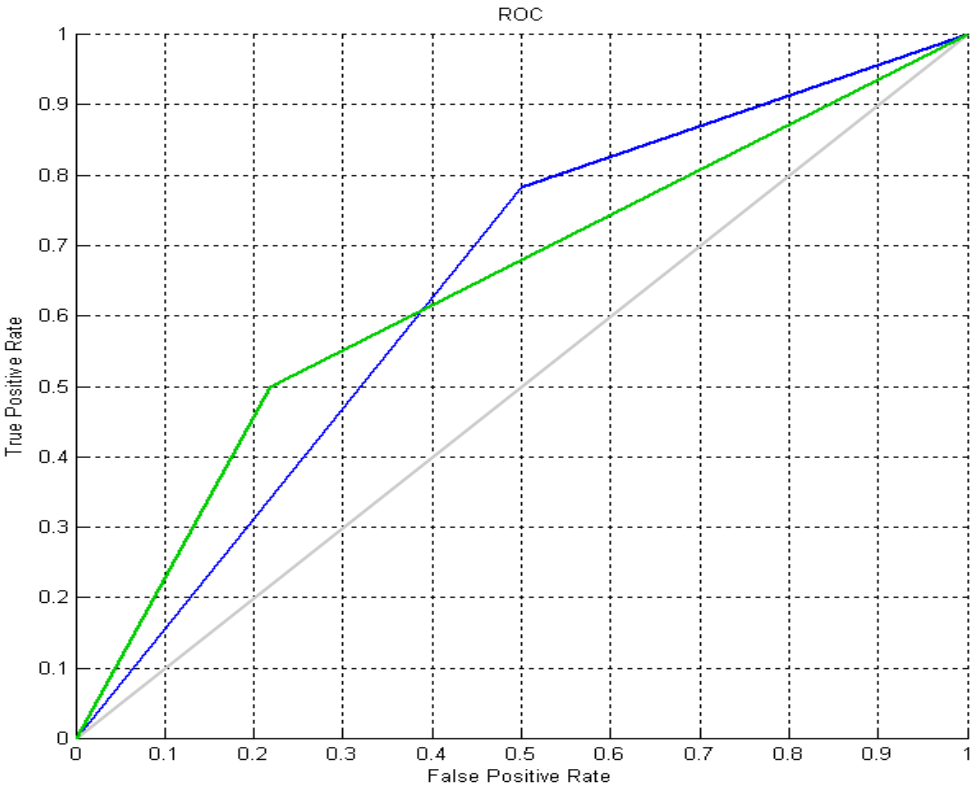
b. The improved ROC result of liver cyst identification



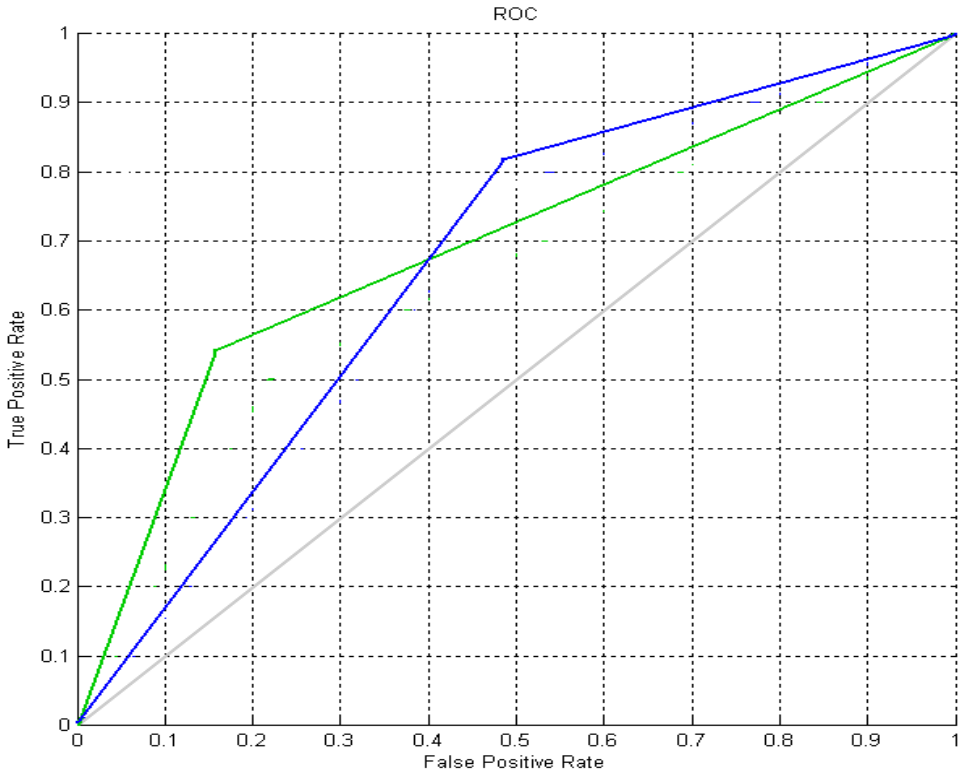
c. The ROC result of diabetes identification



d. The improved ROC result of diabetes identification



e. The ROC result of liver cancer identification



f. The improved ROC result of liver cancer identification

Figure 2. ROC Results of Different Samples Using PSO and Improved PSO Method

In Table 6, we find that the areas under the curve for each disease using our method are bigger than traditional PSO-SVM classification, that is to say, our proposed method has better performance and it is more suitable for clinical use.

The classification accuracy of our improved PSO-SVM model is shown in Table 7.

Group of samples	PSO-SVM	Improved PSO-SVM
Liver cyst-normal	69.2308	68.7466
Diabetes-normal	73.4375	74.2188
Heart disease-normal	77.2368	78.4453
Liver cancer-normal	64.4444	68.7539

Table 7. Classification accuacy comparation (%)

We can draw the conclusion from Table 6 that our proposed method's diagnose accuracy is a little higher than PSO-SVM method, but there exists an exception in the identification of liver cyst.

The evolution process chart and average fitness line is shown in Figure 3.

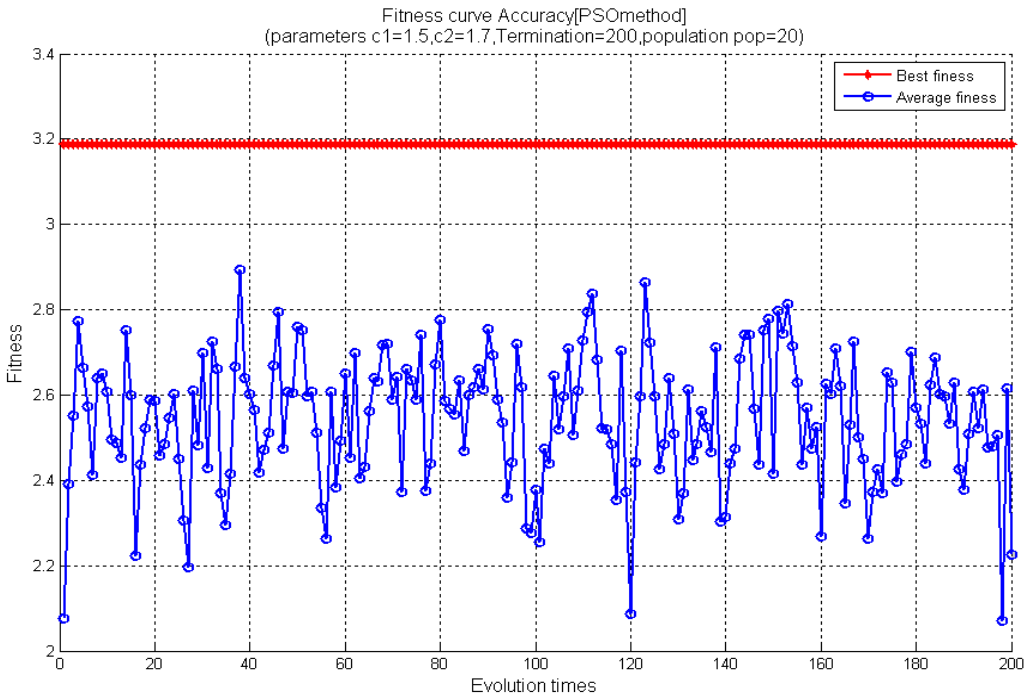


Figure 3. Evolution process chart of diabetes

The elapsed time of PSO-SVM and Improved PSO-SVM process is shown in Table 8.

Group of samples	PSO-SVM	Improved PSO-SVM
Liver cyst-normal	9.9	9.4
Diabetes-normal	350.8	225.7
Heart disease-normal	170.5	122.3
Liver cancer-normal	211.7	183.7

Table 8. Elapsed time for training (Second)

Table 8 shows that the improved PSO-SVM method has lower time complexity than traditional PSO optimizing algorithm.

6. Conclusion

As discussed above, we proposed an improved PSO-SVM method in medical images processing. This model of SVM has a better performance in ROC result curves although there is lower accuracy result in an individual example. Moreover, our method has a lower time complexity than traditional PSO optimizing algorithm.

However, how to make the algorithm faster is our research content in the future.

7. Acknowledgment

This research is supported by the National Science Foundation of China (No: 60973071) and the Liaoning province Natural Science Foundation (No: 20092004).

References

[1] Zhang, R., Ma, J. (2008). An improved SVM method P-SVM for classification of remotely sensed data, *Remote Sensing*, 29 (20) 6029-6036.

[2] Gletsos, M., Mougiakakou, S. G., Matsopoulos, G. K., Nikita, K. S., Nikita, A. S., Kelekis, D. (2003). A computer-aided diagnostic system to characterize CT focal liver lesions, Design and optimization of a neural network classifier,. *IEEE Transactions on,Information Technology in Biomedicine*, 7 (3) 153-162.

[3] Xian, G. M. An identification method of malignant and benign liver tumors from ultrasonography based on GLCM texture features and fuzzy SVM,. EXPERT SYSTEMS WITH APPLICATIONS, 2010, 37 (10) 6737-6741

[4] Vapnik, V. (1995). The nature of statistical learning theory, Springer-Verlag, New York.

[5] Liu, S., Yingjia, C. etc. (2005). A new weighted support vector machine with GA-based parameter selection,. *In:* Proceedings of the Fourth International Conference on Machine Learning and Cybernetic s, Guangzhou, 18-21 August.

[6] Li, Z., Zhou, W. D., Jiao, L. C. (2004). Wavelet Support Vector Machine, *IEEE Transactions on Systems, Man and Cybernetics*, 34 (1) 4-39.

[7] Liu, C., Zhang, J., Ren, G. (2009). Fault Diagnosis of Analogue Circuits Using Hyper-sphere SVM and AdaBoost, *In:* 2009 Second International Symposium on Computational Intelligence and Design

[8] Huiyan Jiang, Fengzhen Tang, Xiangying Liu, Wenhan Wang, Xihe Gao, (2011). Liver Cancer Identification Based on ACO-AdaBoost-SVM model IFMIA in Okinawa, Japan.

[9] Xiangying Liu, Huiyan Jiang, Fengzhen Tang. Parameters Optimization in SVM Based-on Ant Colony Optimization Algorithm, *In:* 2010 IITA International Conference on Nanotechnology and Computer Engineering(CNCE 2010), 9-10, July (2010), Qingdao, China.

[10] Gao, H. H., Yang, H. H., Wang, X. Y. (2005). Ant Colony Optimization Based Network Intrusion Feature Selection and Detection, *In:* Proc. of the Fourth International Conference on Machine Learning and Cybernetics, 3871-3875.

[11] Zhang, X. Y., Guo, Y. L. (2009). Optimization of SVM Parameters Based on PSO Algorithm, *In*:. ICNC'09, Fifth International Conference on, (1) 536-539.

[12] Wang, S. Z., Meng, B. (2010). PSO Algorithm for Support Vector Machine, *In:* Third International Symposium on Electronic Commerce and Security, 377-380.

[13] Huiyan Jiang, Xiangying Liu, Lingbo Zhou, Hiroshi Fujita, Xiangrong Zhou, (2011). Morlet-RBF SVM Model for Medical Images Classification [J]. Lecture Notes in Computer Science, 6676, 121-12.