Twin Support Vector Machines Based on Particle Swarm Optimization

Shifei Ding[1,2], Junzhao Yu[1], Huajuan Huang[1], Han Zhao[1]
1. School of Computer Science and Technology, China University of Mining and Technology, Xuzhou, China
Email: junzhao1989@163.com
2. Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, Chinese Academy of Sciences, Beijing, 100190 China
Email: dingsf@cumt.edu.cn

***Abstract*—Twin support vector machines (TWSVM) is similar in spirit to proximal SVM based on generalized eigenvalues (GEPSVM), which constructs two nonparallel planes by solving two related SVM-type problems, so that its computing cost in the training phase is only 1/4 of standard SVM. In addition to keeping the advantages of GEPSVM, the classification performance of TWSVM is also significantly better than that of GEPSVM. However, there are also many deficiencies in TWSVM, difficult to specify the parameters is one of them, in order to overcome this deficiency, in this paper, we propose the twin support vector machines based on particle swarm optimization (PSO-TWSVM). This algorithm use PSO to find the parameters for TWSVM, so that blindly parameters selection is avoided. The experimental results show that this algorithm is able to find the suitable parameters, and has higher classification accuracy compared with some other algorithms.**

***Index Terms*—Twin Support Vector Machines; Particle Swarm Optimization; Pattern classification; Parameter optimization**

I. INTRODUCTION

Support vector machines (SVM) [1-2] is a new machine learning method based on statistical learning theory and structural risk minimization [3-4], and it has become a hot research topic in the field of machine learning because of its excellent performance. In order to reduce the computational cost of SVM, Fung et al. [5] proposed proximal support vector machines (PSVM) in 2001, does binary classification by obtaining two parallel hyperplanes on the premise of guaranteeing the maximum interval. In 2006, Mangasarian and Wild [6] proposed proximal SVM based on generalized eigenvalues (GEPSVM), which successfully overcomes the existing shortcomings of PSVM. This algorithm abandons the constraint of PSVM that hyperplanes must be parallel. The optimization target of it is that each hyperplane should be as close as possible to the samples for its own class and as far as possible from the samples for the other class at the same time. Jayadeva et al. [7] proposed twin support vector machines (TWSVM) [8] in 2007, as a variant of GEPSVM, attempts to improve the generalization of GEPSVM, its thought is to solve two dual quadratic programming problems(QPPs) of smaller size rather than solving one dual quadratic programming problem with large number of parameters in standard SVM. Compared with SVM, one of the main advantages of TWSVM is that it is four times faster. The classification performance of TWSVM also is better than GEPSVM, and it is very powerful to deal with large-scale datasets, while the standard SVM is not suitable for a large number of samples. However, it is inevitable for TWSVM to solve two QPPs that lead to rather high computational complexity.

Although TWSVM is proposed only recently, it has become a hot research topic because of its solid theoretical and practical foundation. Many scholars devote themselves to the study of TWSVM [9-10] and propose some improved algorithms. For example, Jing Chen and Ji Guangrong [11] proposed weighted least squares twin support vector machines (WLSTWSVM)， in order to eliminate the impact of noise and obtain better classification performance[12-13], different weights are put on the error variables. Ye Qiaolin et al. [14] proposed weighted twin support vector machines with local information (WLTWSVM) which is a new nonparallel plane classifier. It can mine as much correlation between data points with the same labels that may be important for classification performance as possible. WLTSVM can not only get better classification accuracy, but also reduce the computational cost. Qi Zhiquan et al. [15] proposed a new robust twin support vector machines (RTWSVM) via second order cone programming formulations for classification. This algorithm can deal with data with measurement noise efficiently.

TWSVM is widely used in various fields and get impressive experimental results because of its high classification accuracy and speed. For example, Ganesh R. Naik [16] and SP Arjunan [17] applied TWSVM to the gesture classification based on sEMG, and the result shows that it is eminently suited to such applications.

Manuscript received November 1, 2012; revised December 1, 2012; accepted January 1, 2012.
Corresponding author: Shifei Ding, Email: dingsf@cumt.edu.cn

doi:10.4304/jcp.8.9.2296-2303

Hanhan Cong et al. [18] applied TWSVM with Gaussian Mixture Mode (GMMs) to the text independent speaker recognition system, and TWSVM presents better performance than the standard SVM because of the ability of processing large-scale dataset. Zhang Xinsheng et al. [19] applied TWSVM to the detection of clustered microcalcifications (MCs), then improved the TWSVM by Boosting algorithm, proposed Boosting-TWSVM [20] and applied it to the microcalcifications clusters detection. The experiments show that this method improves the detection accuracy and rate to some extent. After that, they proposed Bagging and Boosting-TWSVM [21] by combining the Bagging algorithm with Boosting algorithm together. Compared with TWSVM, Bagging and Boosting-TWSVM can solve the unstable problem of TWSVM while keeping the detection accuracy in a noisy environment. After that, they made further improvement by using TWSVM and subspace learning algorithms to detect MCs[22].

Although in recent years the study of TWSVM [23-26] has made great progress in the algorithm improvement and its application, there are still some deficiencies, for example, multiple parameters in TWSVM [27-29] need to be specified by rule of thumb, but doing so is difficult to find the most suitable parameters, and has an adverse impact on the final classification results. So in this paper, we propose the twin support vector machines based on particle swarm optimization (PSO-TWSVM). Firstly, using PSO to find the most suitable parameters, and then taking these parameters into TWSVM to further improve its classification accuracy.

II. SUPPORT VECTOR MACHINES

The principle of SVM is to find an optimal classification hyperplane to maximize the blank area on its both sides and ensure high classification accuracy at the same time. In theory, support vector machines can achieve the optimal classification performance on the linearly separable data.

Consider a binary classification problem, suppose $(\mathbf{x}_i, y_i), i=1,2,\cdots,l, \mathbf{x}\in R^n, y\in\{\pm 1\}$ is the training dataset, the hyperplane is denoted as $(\mathbf{w}\cdot\mathbf{x})+b=0$, in order to get high accuracy and maximum classification margin, it should satisfy the following constraints: $y_i[(\mathbf{w}\cdot\mathbf{x}_i)-b]\geq 1, i=1,2,\cdots,l$, so the margin is $2/\|\mathrm{w}\|$. The problem transforms to find the find the optimal solution of:

$$\min \Phi(\mathbf{w})=\frac{1}{2}\|\mathbf{w}\|^2=\frac{1}{2}(\mathbf{w}\cdot\mathbf{w}) \qquad (1)$$

To solve this problem, introduce the Lagrange function:

$$\mathrm{L}(\mathbf{w},b,\mathrm{a})=\frac{1}{2}\|\mathbf{w}\|^2-\sum_{i=1}^{l}\mathrm{a}_i(\mathrm{y}_i((\mathbf{w}\cdot\mathbf{x}_i)+b)-1) \quad (2)$$

Where $\mathrm{a}_i>0$ is the multiplier, and the solution is determined by the saddle point of Lagrange function. Then the QP problem is transformed into a dual problem:

$$\begin{aligned} \max \quad & Q(\mathrm{a})=\sum_{j=1}^{l}\mathrm{a}_j-\frac{1}{2}\sum_{i=1}^{l}\sum_{j=1}^{l}\mathrm{a}_i\mathrm{a}_j y_i y_j(\mathbf{x}_i\cdot\mathbf{x}_j) \\ s.t. \quad & \sum_{j=1}^{l}\mathrm{a}_j y_j=0,\ j=1,2,\cdots,l \\ & \mathrm{a}_j\geq 0,\ j=1,2,\cdots,l \end{aligned} \qquad (3)$$

The optimal solution is $\mathrm{a}^*=(\mathrm{a}_1^*,\cdots,\mathrm{a}_l^*)^T$ 。

Calculate the optimal weight vector

$$\mathbf{w}^*=\sum_{j=1}^{l}\mathrm{a}_j^* y_j\mathbf{x}_j \qquad (4)$$

and the optimal partial derivative

$$b^*=y_i-\sum_{j=1}^{l}y_j\mathrm{a}_j^*(\mathbf{x}_j\cdot\mathbf{x}_i) \qquad (5)$$

Where $i\in\{i\,|\,\mathrm{a}_i^*>0\}$ 。Then we get the optimal hyperplane $(w^*\cdot\mathbf{x})+b^*=0$, and the optimal classification function is:

$$f(\mathbf{x})=\mathrm{sgn}\{(\mathbf{w}^*\cdot\mathbf{x})+b^*\}=\mathrm{sgn}\{(\sum_{j=1}^{l}\mathrm{a}_j^* y_j(\mathbf{x}_j\cdot\mathbf{x}_i))+b^*\},\mathbf{x}\in R^n$$

The geometrical interpretation of SVM is depicted in Figure.1 for a toy example, where two lines represent two hyperplanes, red and blue dots represent the training points belonging to category 1 and category -1.

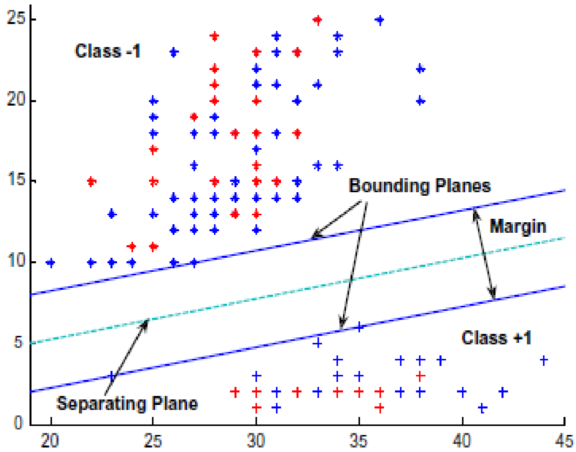

Figure 1. SVM

For the further improvement of SVM, Jayadeva et al. proposed twin support vector machines (TWSVM) in 2007.

III .TWIN SUPPORT VECTOR MACHINES

For a binary classification problem, the time complexity of the standard SVM is $O(m^3)$, where m is the number of samples. Assuming that the number of samples of each class is m/2, the time complexity of solving two optimization problems is $O(2*(m/2)^3)$, so the time complexity of the TWSVM 1/4 of SVM. In addition, in order to overcome the problem that SVM is not suitable to deal with the unbalanced datasets, when the samples number of one class is much larger than the other one, TWSVM can set different penalty parameters for two classes of misclassification samples.

A. Algorithm Thought of TWSVM

Different from standard SVM which constructs two parallel hyperplanes, TWSVM [30-32] constructs a positive hyperplane and a negative hyperplane without the restriction of parallel. Similar to the idea of maximum interval, TWSVM needs the hyperplane as far as possible from one of the two classes of samples, while the difference is that TWSVM also needs the hyperplane as close as possible to the other class of samples.

All in all, the thought of TWSVM [33-35] is to construct two hyperplanes for each class of samples and each hyperplane should be as close as possible to one class of samples, and as far as possible from the other class of samples at the same time. The new sample will be assigned to one of the classes depending on its proximity to the hyperplane.

The geometrical interpretation of TSVM is depicted in Figure.2.

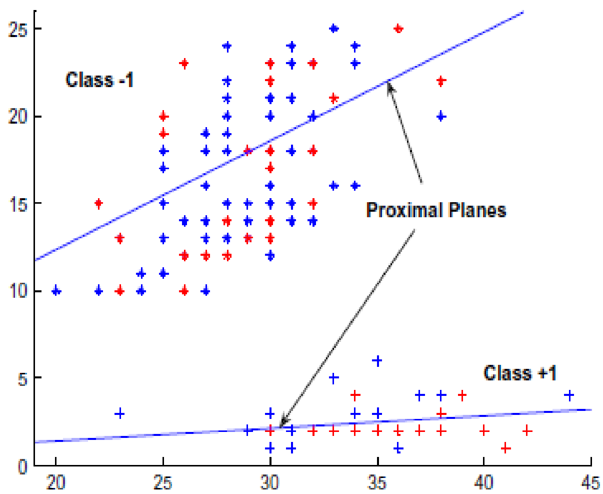

Figure 2. TWSVM

B. Model of TWSVM

Consider a binary classification problem, m_1 training points belonging to category +1 and m_2 training points belonging to category -1 in the n-dimensional real space R^n. Let matrix A in $R^{m_1 \times n}$ represents the training points of category +1 and matrix B in $R^{m_2 \times n}$ represents the training points of category -1.

The central thought of TWSVM is to construct two nonparallel hyperplanes in n-dimension input space: $x^T w_1 + b_1 = 0, x^T w_2 + b_2 = 0,$ and each hyperplane should be as close as possible to the samples of its own class and at the same time as far as possible from the samples of the other class.

The formulation of TWSVM can be expressed as following:

$$\text{(TWSVM1)}\ \min \frac{1}{2}(Aw_1+e_1b_1)^T(Aw_1+e_1b_1)+C_1e_2^T\xi$$

$$\text{s.t.}\ -(Bw_1+e_2b_1)+\xi \ge e_2, \xi \ge 0 \qquad (6)$$

$$\text{(TWSVM2)}\ \min \frac{1}{2}(Bw_2+e_2b_2)^T(Bw_2+e_2b_2)+C_2e_1^T\xi$$

$$\text{s.t.}\ -(Aw_2+e_1b_2)+\xi \ge e_1, \xi \ge 0 \qquad (7)$$

Where C_1 and C_2 are the penalty parameters, e_1 and e_2 are column vectors of ones of appropriate dimensions, superscript T denotes transposition, and $w_1 \in R^n$, $w_2 \in R^n$, $b_1 \in R$ and $b_2 \in R$, ξ is the slack variable.

The objective functions seek the distance from the sample to the hyperplane by the square distances, and minimize the distance to ensure the hyperplane is as close as possible to the samples for its own class. The inequality constraint ensures that the distance from the sample to hyperplane is at least 1.

IV. Twin Support Vector Machines Based on Particle Swarm Optimization

The research of the application and theory of twin support vector machines is still in its infancy, but because TWSVM has a solid theoretical and practical basis that it is based on support vector machines and proximal SVM based on generalized eigenvalues, in essence it is based on statistical learning theory. In recent years, many scholars have flung themselves into this field and proposed many improved twin support vector machines algorithms.

The most common way to select parameters of TWSVM is choosing randomly according to the experiences in a certain scope. However, this method has several limitations, such as arbitrariness and blindness. The inappropriate way in selecting TWSVM parameter will result in an inaccurate classification substantially. The PSO algorithm is a better approach comparing with the previous method. It has a speedy convergence, high solving quality, and robust result in the area of multidimensional space function optimization, dynamic goal seeking optimization and so on. Moreover, PSO algorithm is a relatively simple approach, it has small amount of calculation, practical and easier to realize programming. Therefore, we endeavor to find out more accurate parameters of TWSVM by the use of PSO to improve the final classification performance.

A. The Background of PSO

According to the observation and investigation on biological groups, the swarm intelligence generated by individual cooperation and competition in the complex sexual behavior within the groups of organisms, can frequently provide efficient solutions for certain problems. Kennedy et al. [36] inspired by a flock of birds feeding behavior, proposed Particle Swarm Optimization (PSO) in 1995. Compared with evolutionary algorithm, PSO retains global search strategy based on the swarm, and its speed-displacement search model is easy to operate which can avoid the complexity in evolutionary operation.

The PSO Algorithm is divided into two kinds, global PSO and local PSO. The study showed that: the convergence rate of global PSO algorithm is fast, but sometimes it is easy to sink into local optima. On the contrary, local PSO can easily avoid local optima, but the convergence rate of it is relatively slow. So, many scholars proposed improved PSO algorithms. Such as, adaptive PSO based on evolutionary state estimation [37], adaptive PSO based on Sigmoid inertia weight [38], multi-objective PSO based on fuzzy-learning sub-swarm [39] and PSO based on stable strategy [40]. What is more, PSO is also successfully used to optimize SVM [41-42] and the radial basis function network [43].

B. PSO-TWSVM Principle

The position of particle in PSO represents a potential solution of the optimization problem in the search space. Every particle has a fitness value that determined by the optimal function and a speed to determine the direction and range of their travel. PSO initializes a group of random particles (random solutions), and then searches the optimal solution in the solution space following the current optimum particle through iteration. The particle tracks two extremes to update itself in every iteration. One is the optimal position for itself, called individual extreme; the other one is the optimal swarm position, called global extreme.

The formulation of PSO can be expressed as following:

$$v_i = wv_i + c_1r_1(p_i - x_i) + c_2r_2(g - x_i) \tag{8}$$

$$x_i = x_i + v_i \tag{9}$$

Where v_i is the speed of the i-th particle, x_i is the position of the i-th particle, and p_i is the optimal position of the i-th particle, g is the optimal position of all the particles, r_1 and r_2 are the random number uniformly distributed in (0,1), w is the flexible coefficient of v_i, it is used to expand the search space to get better solutions. The larger w, the larger search space is, while smaller w can ensure PSO convergent to the optimum position faster. Generally, w is set as 0.8, c_1 and c_2 is the weight of the speed that each bird fly to p_i and g, and if c_1 = 0, it means the birds do not have the cognitive ability, while c_2 = 0 means the birds do not share the swarm information. Generally, we set $c_1 = c_2 = 2.0$.

The thought of PSO-TWSVM is to initialize a particle swarm in a two-dimensional search space, in which the position of the particle is expressed as a two-dimensional vector $X_i = (x_{i1}, x_{i2})$, x_{i1} represets the penalty factor C_1 in TWSVM, and x_{i2} represents the penalty factor C_2. The best position of the i-th particle is denoted as p_i, and the best positon of the particle swarm is denoted as g; the flight speed of the i-th particle is denoted as v_i. PSO-TWSVM finds the optimal parameters through initialization and iterative optimization, and then brings them into TWSVM to classify.

C. Algorithmic Flow

The specific flow of PSO-TWSVM is shown in the following figure:

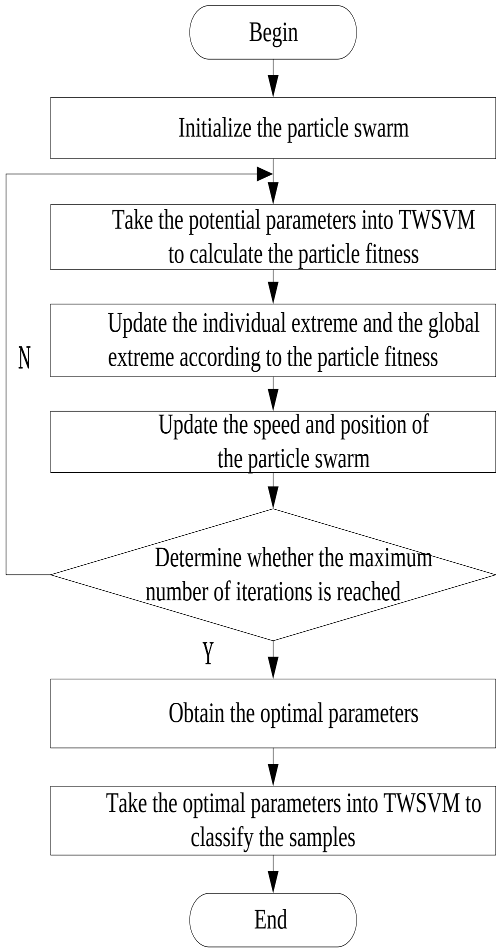

Figure 3. Flow chart

The steps of using PSO to optimize the parameters in TWSVM are as follows:

Step1 Set the swarm size as N and the maximum number of iterations as K. Then initialize the particle swarm.

Step 2 Take the particles obtained from the initialization into TWSVM to classify the training datasets. Then take the classification accuracy as the fitness.

Step 3 Search the optimal solution through iteration. According to the following formulation: $v_i = wv_i + c_1 r_1 (p_i - x_i) + c_2 r_2 (g - x_i)$ and $x_i = x_i + v_i$, continually update the speed and position of the particles and calculate the fitness. If the fitness is better than the fitness of the particle's best position, update the individual extreme. If the fitness of all the particles' best position is better than the fitness of the current global best position, update the global extreme.

Step 4 Determine whether the maximum number of iterations is reached. If done, stop the iteration. Otherwise add 1 to the number of iteration, repeat **Step 3** and record the individual extreme and the global extreme.

Step 5 Finally, get the optimal vector $g = (x_1, x_2)$, where x_1 represents the optimal penalty parameter C_1 of TWSVM, x_2 represents the penalty parameter C_2. Then take them into TWSVM and constitute the PSO-TWSVM model.

Step 6 Stop the operation.

V. Experimental Results and Analysis

In order to verify the effectiveness of twin support vector machines based on particle swarm optimization, we use the Australian dataset, Sonar dataset and the Pima-Indian dataset from UCI machine learning database (http://archive.ics.uci.edu/ml/) for validation. Our algorithm takes the classification accuracy of TWSVM on different datasets as the fitness value. The experiment is completed MATLAB environment in the computer with 2G memory and 6.4G hard disk. Two positive coefficients $c_1 = c_2 = 2.0$, the number of particle swarm is 30, and the flexible coefficient w = 0.8, the maximum number of iterations is 200.

The description of three datasets is shown in Table I:

TABLE I.

DESCRIPTION OF THREE DATASETS

Data Sets	Samples	Attributes
Australian	690	14
Sonar	208	60
Pima-Indian	768	8

Using the optimal parameters obtained on different datasets by PSO to classify corresponding dataset and then comparing the classification accuracy with GEPSVM and PSVM, the results can be seen from Table II:

TABLE II.

CLASSIFICATION ACCURACY COMPARISON

Data Sets	PSO-TWSVM	GEPSVM	PSVM
Australian	87.77	80.00	85.43
Sonar	76.74	72.62	74.51
Pima-Indian	80.52	76.66	77.86

From the experimental results, we can clearly see that compared with the traditional classification algorithms, the proposed algorithm has better performance in testing accuracy because of the use of PSO to find the optimal parameters for TWSVM.

The convergence curves of PSO-TWSVM on three datasets are shown in Figure 4-6.

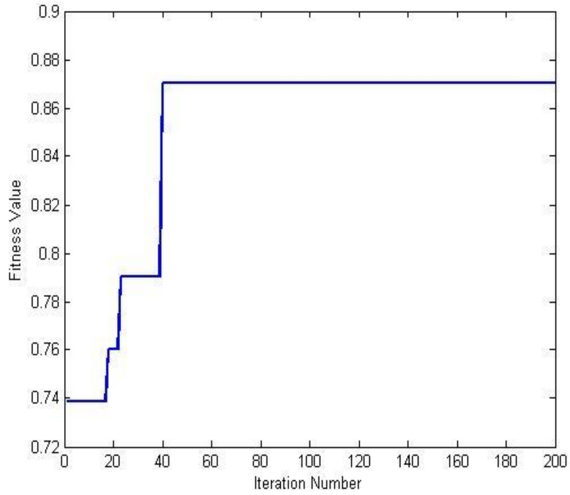

Figure 4. The convergence curve of Australian dataset

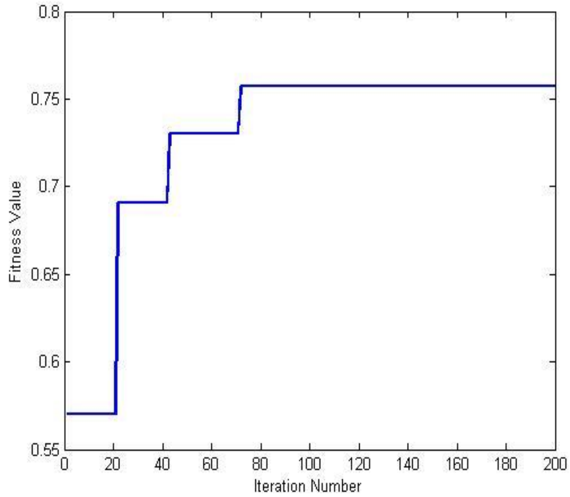

Figure 5. The convergence curve of Sonar dataset

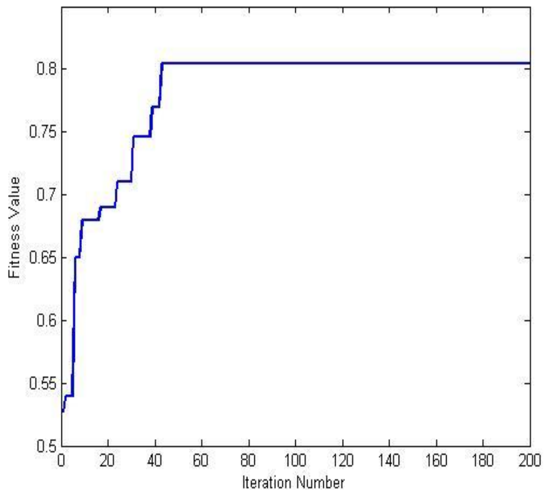

Figure 6. The convergence curve of Pima-Indian dataset

Figure 7 shows the classification accuracy comparison of three classification algorithms more intuitively:

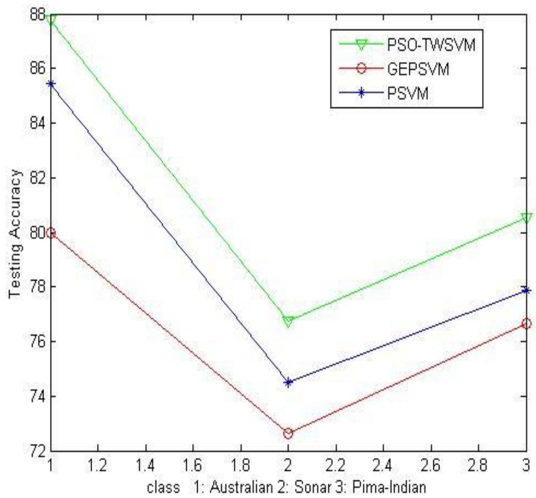

Figure 7. Comparison of three algorithms

The abscissa represents three different datasets used in the experiment, while the ordinate represents the classification accuracy and the three polylines represent three classification algorithms: Twin SVM, proximal SVM based on generalized eigenvalues and proximal SVM.

VI. CONCLUSIONS

Compared with the traditional classification algorithms, such as SVM, PSVM and GEPSVM, twin support vector machines is an effective method for solving large datasets and unbalanced datasets classification problem, and its calculation accuracy and training speed are also far superior to those of other algorithms, but at the same time there are also some shortcomings, such as it is difficult for TWSVM to set parameters. In this paper, in order to overcome this disadvantage, we efficiently take advantage of the fast convergence rate and strong optimization capability of PSO, and propose twin support vector machines based on particle swarm optimization (PSO-TWSVM), it uses PSO to optimize the parameters of TWSVM to avoid the blindness of parameters selection. The experimental results show that this algorithm is able to find suitable algorithm parameters, and has higher classification accuracy compared with GEPSVM and PSVM. The actual application field for twin support vector machines based on particle swarm optimization is limited at present. So how to apply PSO-TWSVM to the daily life effectively is one of the important aspects in future research. What is more, how to combine other parameter optimization methods with TWSVM, and compare their advantages and disadvantages is also the next work we should do.

ACKNOWLEDGMENT

This work is supported by the National Key Basic Research Program of China (No. 2013CB329502), the National Natural Science Foundation of China (No.41074003), the Opening Foundation of the Key Laboratory of Intelligent Information Processing of Chinese Academy of Sciences (IIP2010-1).

REFERENCES

[1] Cristianini N, Taylor JS. An Introduction to Support Vector Machines and Other Kernel-based Learning Methods. Beijing: Publishing House of Electronics Industry. 2004.
[2] Ding Shifei, Qi Bingjuan., Tan HongYan. “An Overview on Theory and Algorithm of Support Vector Machines”. Journal of University of Electronic Science and Technology of China, vol. 40, no.1, pp. 2-10, 2011.
[3] Vapnik VN. The Nature of Statistical Learning Theory. Beijing: Tsinghua University Press. 2000
[4] Vapnik VN. Statistical Learning Theory. Beijing: Publishing House of Electronics Industry. 2004.
[5] Fung G, Mangasarian O L. “Proximal Support Vector Machine Classifiers”. In: Proc 7th ACMSIFKDD Intl Conf on Knowledge Discovery and Data Mining, pp.77-86, 2011.
[6] Mangasarian Olvi L, Wild Edward W. “Multisurface Proximal Support Vector Machine Classification via Generalized Eigenvalues”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.28, no.1, pp. 69-74, 2006.
[7] Jayadeva, Khemchandni Reshma, Suresh Chandra. “Twin Support Vector Machines for Pattern Classification”. IEEE Transactions on Pattern Analysis and Machine Intelligence, vol.29, no.5, pp.905-910, 2007.
[8] Ding Shifei, Yu Junzhao, Qi Bingjuan. “An Overview on Twin Support Vector Machines”. Artificial Intelligence Review (DOI:10.1007/s10462-012-9336-0) , 2012.
[9] Huajuan Huang, Shifei Ding, Fengxiang Jin, Junzhao Yu, Youzhen Han. “A Novel Granular Support Vector Machine Based on Mixed Kernel Function ”. International Journal of Digital Content Technology and its Applications, vol.6, no.20, pp.484-492, 2012.
[10] Junzhao Yu, Shifei Ding, Fengxiang Jin, Huajuan Huang, Youzhen Han. “Twin Support Vector Machines Based on

Rough Sets ". International Journal of Digital Content Technology and its Applications, vol.6, no.20, pp .493-500. 2012.

[11] Chen Jing, Ji Guangrong. "Weighted Least Squares Twin Support Vector Machines for Pattern Classification ". 2010 The 2nd International Conference on Computer and Automation Engineering, Singapore: [s.n.], vol.2, pp.243-246, 2010.

[12] Arun Kumar M., Gopal M. "Least Squares Twin Support Vector Machines for Pattern Classification." Expert Systems with Applications, vol. 36, no.4, pp.7535-7543, 2009.

[13] M. Arun Kumar, Reshma Khemchandani, M. Gopal, Suresh Chandra, "Knowledge based Least Squares Twin support vector machines". Information Sciences, vol.180, no.23, pp.4606–4618, 2010.

[14] Ye Qiaolin, Zhao Chunxia, Gao Shangbing. "Weighted Twin Support Vector Machines with Local Information and its application". Neural Networks. Vol.35, pp.35:31-39, 2012.

[15] Qi Zhiquan, Tian Yingjie, Shi Yong. "Robust Twin Support Vector Machine for Pattern Classification". Pattern Recognition. Vol.46, no.1, pp. 305-316,2013.

[16] Ganesh R N, Dinesh K K, Jayadeva. "Twin SVM for Gesture Classification using The Surface Electromyogram". IEEE Transactions on Information Technology in Biomedicine, vol.14, no.2, pp.301-308, 2010.

[17] Arjunan S P, Kumar D K, Naik G R. "A Machine Learning based Method for Classification of Fractal Features of Forearm sEMG using Twin Support Vector Machines". Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp.4821-4824, 2010.

[18] Cong Hanhan, Yang Chengfu, Pu Xiaorong. "Efficient Speaker Recognition based on Multi-class Twin Support Vector Machines and GMMs". 2008 IEEE Conference on Robotics, Automation and Mechatronics, pp. 348-352, 2008.

[19] Zhang Xinsheng, Gao Xinbo, Wang Ying. "Twin Support Vector Machine For MCs Detection". Journal of Electronics (China), vol.26, no.3, pp.318-325, 2009.

[20] Zhang Xinsheng. "Boosting Twin Support Vector Machine approach for MCs Detection". Asia-Pacific Conference on Information Processing, vol.46, pp.149-152, 2009.

[21] Zhang Xinsheng, Chen Yongfeng, Luo Zhengshan, "Microcalcification Clusters Detection with Bagging and Boosting based Twin Support Vector Machine". Information-an International Interdisciplinary Journal, vol.15, no.1, pp.67-74, 2012.

[22] Zhang Xinsheng, Gao Xinbo. "Twin Support Vector Machines and Subspace Learning Methods for Microcalcification Clusters Detection". Engineering Applications of Artificial Intelligence. Vol.25, no.5, pp.1062-1072, 2012.

[23] Wang Di, Ye Qiaolin, Ye Ning. "Localized Multi-plane TWSVM Classifier via Manifold Regularization". 2010 2nd International Conference on Intelligent Human-Machine Systems and Cybernetics (IHMSC), vol.2, pp.70-73, 2010.

[24] Gao SB, Ye QL, Ye N. "1-Norm Least Squares Twin Support Vector Machines". Neurocomputing. Vol.74, no.17, pp.3590-3597,2010.

[25] Ye Qiaolin, Zhao Chunxia, Ye Ning. "Least Squares Twin sSupport Vector Machine Classification via Maximum One-class within Class Variance". Optimization Methods & Software. vol.27, no.1, pp.53-69, 2012.

[26] M. Arun Kumar, M. Gopal. "Application of Smoothing Technique on Twin Support Vector Machines". Pattern Recognition Letters , vol.29, no.13, pp. 1842-1848, 2008.

[27] Lee Y J, Mangasarian O L. "A Smooth Support Vector Machine for Classification". SSVM:Comput. Optinm. Appl, vol.20, no.1, pp.5-22, 2001.

[28] Chen C, Mangasarian O L. "Smoothing Methods for Convex inEqualitiesand Linear Complementarity Problems". Math Program, vol.71, no.1, pp. 51-69, 1995.

[29] Santanu Ghorai, Shaikh Jahangir Hossian, Anirban Mukherjee, Pranab K. Dutta. "Unity Norm Twin Support Vector Machine Classifier". 2010 Annual IEEE India Conference, pp. 1-4, 2010.

[30] Shao YH, Deng NY. "A Coordinate Descent Margin based-Twin Support Vector Machine for Classification". Neural Networks. Vol.25, pp. 114-121, 2012.

[31] Ye Qiaolin, Zhao Chunxia, Chen Xiaobo. "A Feature Selection Method for TWSVM via a Regularization Technique". Journal of Computer Research and Development, vol.48, no.6, pp. 1029-1037, 2011.

[32] Shao Yuanhai, Zhang Chunhua, Wang Xiao-Bo, and Nai-Yang Deng. "Improvements on Twin Support Vector Machines". IEEE Transactions on Neural Networks, vol.22, no.6, pp.962-968, 2011.

[33] Xie JY, Zhang BQ, Wang WZ. "A Partial Binary Tree Algorithm for Multiclass Classification based on Twin Support Vector Machines". Journal of Nanjing University (Natural Sciences), vol.47, no.4, pp. 354-363,2011.

[34] Peng Xinjun. "A *v*-twin Support Vector Machine (*v*-TWSVM) classifier and its geometric algorithms". Information Sciences, vol.180, no.20, pp. 3863-3875, 2010.

[35] S P Arjunan, D K Kumar, G R Naik. "A machine learning based method for classification of fractal features of forearm sEMG using Twin Support Vector Machines". Engineering in Medicine and Biology Society (EMBC), 2010 Annual International Conference of the IEEE, pp.4821-4824, 2010.

[36] Kennedy J, Eberhart R C. "Particle Swarm Optimization ". Institute of Electrical and Electronics Engineers, vol.1, no.27, pp.1942-1948, 1995.

[37] Zhan ZH, Zhang J, Li Y, et al, "Adaptive Particle Swam Optimization". IEEE Transaction Systems, Man and Cybernetic. Vol.39, no.6, pp.1362-1391, 2009.

[38] Huang Li, Du Weiwei, Ding Lixin. "Adaptive Particle Swarm Optimization Algorithm based on Sigmoid Inertia Weight". Application Research of Computers. Vol.29, no.1, pp. 38-40, 2012.

[39] Jiang Xunlin, Guo Jianyi, Tang Jian, Ling Haifeng. "Multi-objective Particle Swarm Algorithm based on Fuzzy-learning Sub-swarm". Application Research of Computers. Vol.28, no.12, pp.98-100, 2011.

[40] Wei Bo, Li Yuanxiang, Xu Xing, Shen Dingcai. "Particle Swarm Optimization Algorithm based on Stable Strategy ". Computer Science. vol.38, no.12, pp.221-223, 2011.

[41] Ren Honge, Huo Mandong. "Support Vector Machine Optimized by Particle Swarm Optimization Algorithm for Holding Nail Force Forecasting". Application Research of Computers. Vol.26, no.3, pp.867-869,2009.

[42] Li Ming, Zhang Yong, Li Junquan, Zhang Yafen. "Application of Improved PSO-SVM Approach in Speaker Recognition". Journal of University of Electronic Science and Technology of China. vol.36, no.6, pp.1345-1349, 2007.

[43] Song Liwei, Peng Minfang, Tian Chenglai, Shen Meie. "Analog Circuit Diagnosis based on Particle Swarm Optimization Radial Basis Function Network". Application Research of Computers. vol.29, no.1, pp.72-74, 2012.

Shifei Ding is a professor and Ph.D. supervisor at China University of Mining and Technology. His research interests include intelligent information processing, pattern recognition, machine learning, data mining, and granular computing et al. He has published 3 books, and more than 100 research papers in journals and international conferences. He received his B.Sc., M.Sc. degree in mathematics and computer science in 1987, 1998 respectively from Qufu Normal University, received Ph.D. degree in computer science from Shandong University of Science and Technology in 2004, and received post Ph.D. degree in computer science from Key Laboratory of Intelligent Information Processing, Institute of Computing Technology, and Chinese Academy of Sciences in 2006.

In addition, prof. Ding is a senior member of China Computer Federation (CCF), and China Association for Artificial Intelligence (CAAI). He is a member of professional committee of Artificial Intelligence and Pattern Recognition, CCF, professional committee of distributed intelligence and knowledge engineering, CAAI, professional committee of machine learning, CAAI, and professional committee of rough set and soft computing, CAAI. He is an associate Editor-in-Chief for International Journal of Digital Content Technology and its Applications (JDCTA), acts as an editor for Journal of Convergence Information Technology (JCIT), International Journal of Digital Content Technology and its Applications (JDCTA) et al.

Junzhao Yu is currently a graduate student now studying at School of Computer Science and Technology, China University of Mining and Technology, and his supervisor is Prof. Shifei Ding. He received his B.Sc. degree in computer science from China University of Mining and Technology in 2011. His research interests include pattern recognition, machine learning, and twin support vector machines et al.

Huajuan Huang, born in 1984. Receive d her B.Sc.degree and M.Sc. degree in ap plied computer Technology from Guang xi University for Nationalities, Guangxi, China, in 2006 and 2009 respectively. Si nce 2011, she has been a Ph.D. degree ca ndidate in applied computer Technology from the China University of Mining and Technology, Xuzhou, China. Her curren t research interests include data mining, parrern recognition and computational intelligence.

Han Zhao is currently a graduate student now studying in School of Computer Science and Technology, China University of Mining and Technology, and her supervisor is Prof. Shifei Ding. He received his B.Sc. degree in computer science from China University of Mining and Technology in 2012. His research interests include support vector machines, pattern recognition, machine learning et al.