ISSN : 0974 - 7435 Volume 8 Issue 3

BioTechnology
An Indian Journal

FULL PAPER

BTAIJ, 8(3), 2013 [421-427]

Nasopharynx segmentation in MR images based on one-class immune feature weighted support vector machines

Lei Guo[1]*, Ming Hu[1], Lei Zhao[2, 3], Ying Li[1], Guizhi Xu[1]

[1]Province-Ministry Joint Key Laboratory of Electromagnetic Field and Electrical Apparatus Reliability, Hebei University of Technology, Tianjin 300130, (CHINA)

[2]Department of Radiology, Harvard Medical School and Brigham & Women's Hospital, Boston, MA 02115, (USA)

[3]XinAoMDT Technology Co., Ltd., Langfang, Hebei 065001, (CHINA)

E-mail : guoshengrui@163.com

ABSTRACT

In the brain Magnetic Resonance (MR) images, the nasopharynx part is highly irregular. It is difficult to accurately segment this part. Owing to its powerful capacity in solving non-linearity problems, One-class Support Vector Machine (SVM) method has been widely used as a segmentation tool. However, the conventional one-class SVMs assume that each feature of the samples has the same importance degree for the segmentation result, which is not necessarily true in real applications. In addition, one-class SVM parameters also affect the segmentation result. In this study, Immune Algorithm (IA) was introduced in searching for the optimal feature weights and the parameters simultaneously. An Immune Feature Weighted SVM (IFWSVM) method was used to segment the nasopharynx in MR images. Theoretical analysis and experimental results showed that the IFWSVM had better performance than the conventional methods.

© 2013 Trade Science Inc. - INDIA

KEYWORDS

Image segmentation;
Feature weight;
Immune algorithm;
Support vector machine.

INTRODUCTION

Medical image segmentation is applied not only in clinical diagnosis, but it is also the groundwork of image-guided navigation and interventional therapy etc.[1]. The improvement of segmentation accuracy is very important in practical applications[2]. The encephalic tissues including the nasopharynx are difficult to be accurately segmented due to their highly irregular boundaries. Owing to its ability of learning the non-linear distribution of the real data without using any prior knowledge, one-class Support Vector Machines (SVMs) have been applied in tissue segmentation. The aim of the one-class classification is to decide whether a data is in target class or not. Abnormality detection tasks, such as machine fault or medical diagnosis, all belong to the one-class classification problem[3]. Each feature of a sample is supposed to have its different importance degree to the segmentation result[4]. Conventional one-class SVMs, however, do not take feature weights into account[5]. Moreover, the parameters of one-class SVM directly affect the segmentation result[6]. Reference[7] reported a Receiver Operating Characteristics-based weighting feature method for two-class SVM. But for

FULL PAPER

the one-class SVM, the specificity can not be obtained. Immune Algorithm (IA) has the abilities of learning, memorizing and self-adaptive adjusting[8]. In this study, an IA-based method is introduced in searching for the optimal feature weights and parameters simultaneously. Combining one-class SVM with IA and feature weights, one-class Immune Feature Weighted SVM (IFWSVM) is used to segment the nasopharynx in MR images.

ONE-CLASS FEATURE WEIGHTED SVM

Conventional one-class SVM puts all the features of a sample as equally important to the final segmentation result. But in practice, all factors affecting one thing usually have the different importance. Different factors should be given different weights to reflect the different importance of each factor[9]. Similarly, different feature should be given its corresponding weight, which is the put forward concept of feature weighting in SVM.

In the conventional one-class SVM, let $D = \{x_i\}_{i=1}^{l}$, $x_i \in R^d$ be a dataset of target samples[5]. In one-class FWSVM, D is transformed to $D' = \{\lambda_m x_{im}\}$ by the feature weights λ_m. Let $X_i = (\lambda_m x_{im})$. A certain projection Φ is introduced to transform the original dataset into a high dimensional feature space. A decision function $f(X) = W \bullet \Phi(X) - \rho$ should be found to accept the target samples and reject the non-target samples. The primal constrained optimization problem should be solved by the following formula:

$$\begin{aligned} &\min_{W,\xi_i,b} : \tfrac{1}{2}\|W\|^2 + \tfrac{1}{\nu l}\sum_{i=1}^{l}\xi_i - \rho \\ &\text{s.t.: } W \bullet \Phi(X_i) \geq \rho - \xi_i \end{aligned} \quad (1)$$

where W is the normal vector of the hyperplane which represents the decision boundary; $\rho \geq 0$ is the bias of $f(X)$; $\xi_i \geq 0$ is the slack variable; the regularization term ν is a parameter which controls the trade-off and indicates the fraction of samples that should be accepted by the description. Two Lagrangian multipliers $\alpha_i \geq 0$ and $\beta_i \geq 0$ are introduced. The Lagrangian form is constructed as follows:

$$L(W,\xi_i,\rho,\alpha_i,\beta_i) = \tfrac{1}{2}\|W\|^2 + \tfrac{1}{\nu l}\sum_{i=1}^{l}\xi_i - \sum_{i=1}^{l}\beta_i\xi_i - \sum_{i=1}^{l}\alpha_i(W \bullet \Phi(X_i) - \rho + \xi_i) \quad (2)$$

For Eq. 2, setting the partial derivatives to zero, new constraints are obtained. Using the kernel function $K(X_i, X_j) = \Phi(X_i) \bullet \Phi(X_j)$, which is a symmetric function and satisfies the Mercer condition, Eq. 1 can be converted to the following dual problem:

$$\begin{aligned} &\min_{\alpha} : \textstyle\sum_{i,j=1}^{l}\alpha_i\alpha_j K(X_i, X_j) \\ &\text{s.t.: } \textstyle\sum_{i=1}^{l}\alpha_i = 1 \text{ and } 0 \leq \alpha_i \leq \frac{1}{\nu l} \end{aligned} \quad (3)$$

The dual problem in Eq. 3 presents in a quadratic form, and its minimization can be solved by the Quadratic Programming (QP) optimization method. Training one-class SVM on the target dataset is the process of solving the QP problem to obtain Support Vectors (SVs) which support the optimal hyper-plane. The following is the decision function:

$$\begin{aligned} &f(X) = \operatorname{sgn}\left(\textstyle\sum_{i=1}^{l}\alpha_i K(X_i, X) - \rho\right) \\ &\rho = \textstyle\sum_{i=1}^{l}\alpha_i (K(X_i, X_p)) \end{aligned} \quad (4)$$

Where X_p is anyone of the SVs. In this paper, Radial Basis Function (RBF) is chosen as the kernel function:

$$K(X_i, X) = \exp\left\{-\frac{\|X_i, X\|^2}{2\sigma^2}\right\} \quad (5)$$

If the feature weights λ_m can truly reflect the different importance of each feature, the one-class IFWSVM is bound to get better segmentation accuracy. It is a complicated problem to determine the weight of each feature. Immune Algorithm is introduced in intelligently optimizing the feature weights.

ONE-CLASS IMMUNE FEATURE WEIGHTED SVM

Biological Immune System (BIS) is a highly evolved intelligent system which has parallel and self-organized features. Artificial Immune System (AIS) simulates BIS to construct mathematical model and strategy by using mathematical method and computer technology. Immune Algorithm based on AIS is developed to solve certain problems of engineering applications [10]. In IA, a goal function of the optimization is regarded as an Antigen (Ag); the optimal solution is regarded as an

Antibody (Ab). The matching degree between the Ag and the Ab is described as the affinity: $\|Ab - Ag\|$, which reflects the closeness between the goal and the potential solution. Resemblance among antibodies is described as the similarity which also reflects the antibody diversification[11].Biological Immune System (BIS) is a highly evolved intelligent system which has parallel and self-organized features. Artificial Immune System (AIS) simulates BIS to construct mathematical model and strategy by using mathematical method and computer technology. Immune Algorithm based on AIS is developed to solve certain problems of engineering applications [10]. In IA, a goal function of the optimization is regarded as an Antigen (Ag); the optimal solution is regarded as an Antibody (Ab). The matching degree between the Ag and the Ab is described as the affinity: $\|Ab - Ag\|$, which reflects the closeness between the goal and the potential solution. Resemblance among antibodies is described as the similarity which also reflects the antibody diversification[11]. TP=True Positive indicates the number of points correctly classified; FP=False Positive indicates the number of points wrong classified. $MP = TP/(TP + FP)$ (Match Percentage) is the criterion to estimate the performance of the classifier. The goal function (Ag) is to maximize MP. Abs are λ_m, σ and v, reflects the importance degree of each feature andis the width of the RBF controlling the effective range of the kernel function. Different, andaffect. IA is introduced to optimize λ_m, σ and v to get the highest. MP

N potential solutions from the solution space are randomly chosen as the initial antibody generation. For theiteration, there areantibody sets as potential solutions:

$$\begin{aligned} &\mathbf{set1} = (\lambda_1^1, \lambda_2^1, ...\lambda_m^1, \mathbf{v}^1, \sigma^1)_t \\ &... \\ &\mathbf{setN} = (\lambda_1^N, \lambda_2^N, ...\lambda_m^N, \mathbf{v}^N, \sigma^N)_t \end{aligned} \quad (6)$$

The antibody diversification should be maintained to avoid the degeneration and the immaturity of the algorithm. The similarity among antibodies can be described as follows:

$$s_{ij} = 1 - \frac{\|\mathbf{Ab_i} - \mathbf{Ab_j}\|}{\max_{1 \le i, j \le N} \|\mathbf{Ab_i} - \mathbf{Ab_j}\|}, \ \mathbf{i, j = 1,..., N}, \ \mathbf{i \neq j} \quad (7)$$

According to the similarity, Abs with high similarities will be suppressed. Assume that the most similar antibody sets areand, one of them is removed and the other is kept. In this study, let $N = 20$, the 5 most similar antibody sets are removed and the rest 15 antibody sets are kept.

For the clone selection phase, the affinity of the antigen-antibody is computed. According to the affinity, clone selection is executed with antibody removal for antibodies with low affinity and antibody clone for antibodies with high affinity. In this way, the convergence speed of the algorithm is expedited. In this study, each antibody set is put into the one-class FWSVM system respectively, obtaining eachfor each antibody set. Comparing these of the 15 remaining antibody sets, the 5 antibody sets which correspond to the 5 lowest recognition accuracies are removed; the 5 antibody sets which correspond to the 5 highest recognition accuracies are cloned; and the rest 5 antibody sets which correspond to the 5 medium recognition accuracies are kept as they are. After this phase, the number of antibody sets is still 15.

For the antibody mutation phase, the next generation of 15 antibodies is produced by the antibody mutation formula:

$$\mathbf{Ab_i^*} = \mathbf{Ab_i} - (1 - \mathbf{e}^{-\|\mathbf{Ab_i} - \mathbf{Ag}\|}) \|\mathbf{Ab_i} - \mathbf{Ag}\| \quad (11)$$

In our framework, $\|Ab_i - Ag\|$ is defined as $(1 - MP)$ ratio. It means that an antibody is more suitable for an antigen if this ratio becomes smaller. Then 15 mutated antibody sets are put into the one-class FWSVM respectively to obtain their MPs. 10 antibody sets of the 15 mutated antibody sets which correspond to the 10 highest MPs are stored into the immune memory matrix. These 10 antibody sets are called excellent antibodies for the current iteration. A part of the initial antibody generation can be obtained from this matrix for the next iteration in order to enhance the searching ability. After several iterations, the best recognition result can be obtained if the MP can no longer be improved or the termination condition is reached. The feature weights and the parameters of the one-class SVM which correspond to the highestare the optimal, and.

EXPERIMENT AND DISCUSSION

In this study, a dataset of T2-weighted MR images

FULL PAPER

withmatrix size and 1mm slice thickness are acquired using a Centauri 0.3T MRI system (XinAoMDT Technology Co., Ltd., China) under approved IRB. The range of MR images which contains the nasopharynx is 42-116. The experiments are implemented in MATLAB on a PC with a 2.16GHz CPU and 2GB of RAM.

The generalization ability of the one-class SVM is heavily dependent on feature extraction. In this study, 9 statistical features of gray include pixel gray value, average gray values and standard deviations of template andtemplate, average gray value oftemplate, statistical moments based on histogram for 2, 3 and 4 order moments; 48 local texture measures include 12 texture measures for each of 4 orientations () from gray co-occurrence matrix oftemplate including angular 2 order moment, inertia, inverse difference moment, entropy, sum entropy, difference entropy, correlation, sum average, difference average, variance, sum variance, difference variance. Totally, 57 features are extracted from MR images[12].

The structure of nasopharynx is divided into 3 parts (P1: the maxillary sinus, frontal sinus sphenoid sinus and nasopharyngeal; P2: the ethmoid sinus; P3: the nasal passages) for training and generalization respectively due to the different morphological characteristics of different parts. In the training phase, slice 72 including 3 parts of nasopharynx is chosen as the training slice to construct 3 one-class IFWSVM classifiers. The optimal parameters of the 2 one-class SVM algorithms for 3 parts are shown in TABLE 1. The optimal feature weights of 3 parts are shown in Figure 1-3.

TABLE 1 : The Optimal Parameters of The 2 SVM Algorithms for 3 Parts

	P 1	P 2	P 3
IFWSVM	v*=0.37 σ*=1.097	v*=0.43 σ*=0.259	v*=0.76 σ*=0.108
ISVM	v*=0.45 σ*=2.13	v*=0.29 σ*=3.05	v*=0.57 σ*=4.24

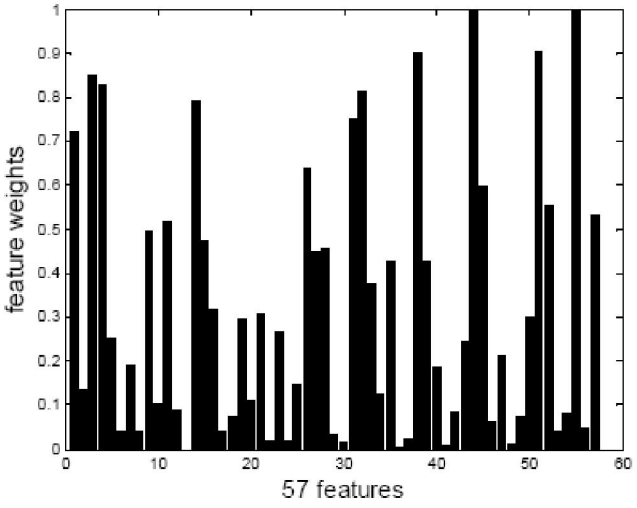


Figure 1 : The feature weight distribution for P1

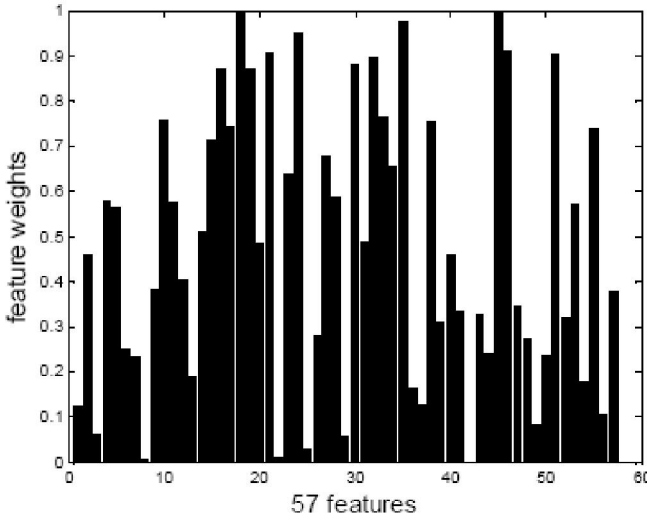


Figure 2 : The feature weight distribution for P2

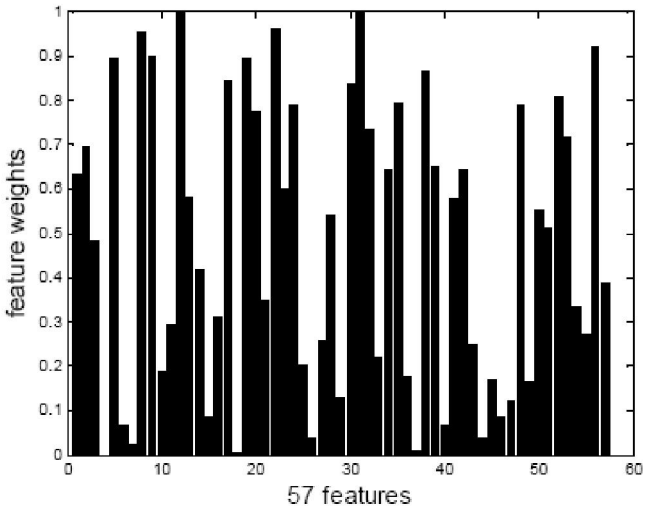


Figure 3 : The feature weight distribution for P3

As shown in Figure 1-3, the feature weights are different. For the example of the P2, the smallest feature weight isand the biggest feature weight is. This means thefeature is the weakest relative toand the feature is the strongest relative to. In this way, the proposed method can obtain superiorby different feature weights.

As a comparison, one-class Immune SVM (ISVM) is also implemented to segment nasopharynx in the same MR images. In the one-class ISVM, onlyandare optimized andis not taken into account. For the generalization phase, the structures of nasopharynx in MR images are segmented by 2 kinds of one-class SVM classifiers. For the example of 5 slices, the comparative performances of these 2 kinds of one-class SVM classifiers are shown in TABLE 2-6

TABLE 2 : The segmentation Results of slice 65

	Algorithm	TP	FP	MP	Time(s)
P1	IFWSVM	682	42	94.2%	0.500
	ISVM	671	53	92.7%	0.484
P2	IFWSVM	491	6	98.8%	0.395
	ISVM	483	14	97.2%	0.265

TABLE 3 : The segmentation results of slice 72

	Algorithm	TP	FP	MP	Time(s)
P1	IFWSVM	1305	4	99.7%	0.496
	ISVM	1294	15	98.9%	0.484
P2	IFWSVM	736	1	99.9%	0.484
	ISVM	728	9	98.8%	0.250
P3	IFWSVM	377	2	99.5%	0.282
	ISVM	375	4	98.9%	0.218

TABLE 4 : The segmentation Results of slice 79

	Algorithm	TP	FP	MP	Time(s)
P1	IFWSVM	1013	37	96.5%	0.485
	ISVM	964	86	91.2%	0.484

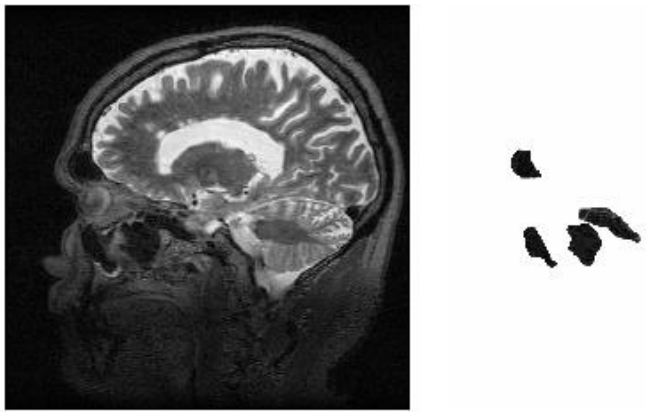
TABLE 5 : The segmentation Results of slice 86

	Algorithm	TP	FP	MP	Time(s)
P1	IFWSVM	1442	53	96.5%	0.500
	ISVM	1367	128	91.5%	0.485
P2	IFWSVM	478	10	98.0%	0.375
	ISVM	474	14	97.1%	0.266
P3	IFWSVM	260	12	95.6%	0.328
	ISVM	254	18	93.4%	0.234

TABLE 6 : The segmentation Results of slice 103

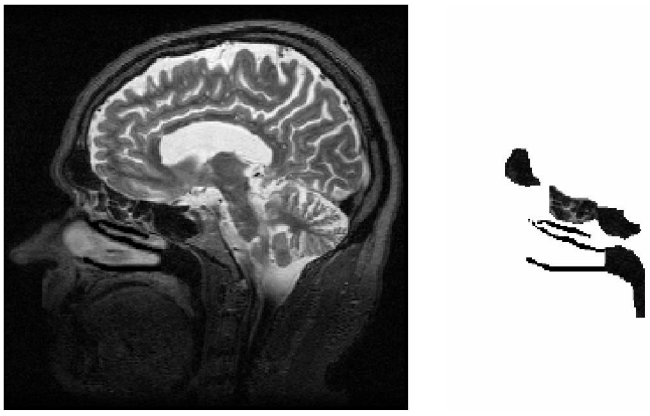
	Algorithm	TP	FP	MP	Time(s)
P1	IFWSVM	995	7	99.3%	0.484
	ISVM	955	47	95.3%	0.365

As shown in TABLE 2-6, the of one-class IFWSVM are higher than that of one-class ISVM due to the searching for optimal weights of different features. It means that optimal feature weights have a positive effect on. Secondly, theof one-class IFWSVM are smaller than that of one-class ISVM, which means that the positive samples are recognized more in the target object by using the one-class IFWSVM. Although the time complexity of one-class IFWSVM is a little bit



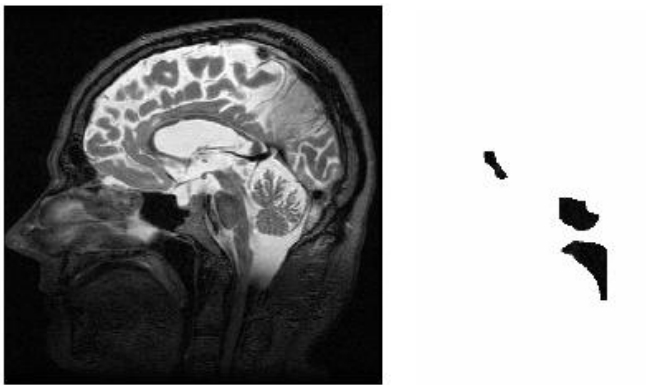
(a) Slice 65 **(b) Segmentation result**

Figure 4 : The segmentation result of slice 65



(a) Slice 72 **(b) Segmentation result**

Figure 5 : The segmentation result of slice 72

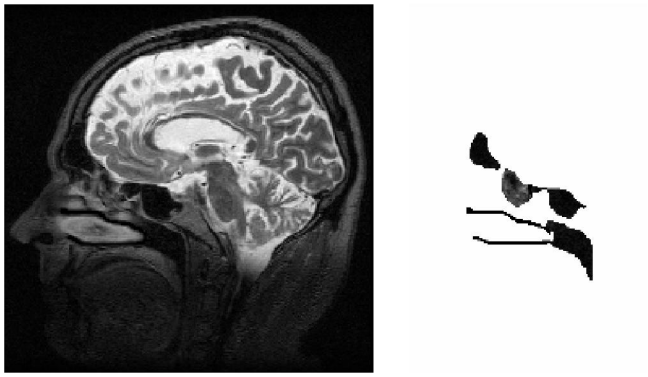


(a) Slice 79 **(b) Segmentation result**

Figure 6 : The segmentation result of slice 79

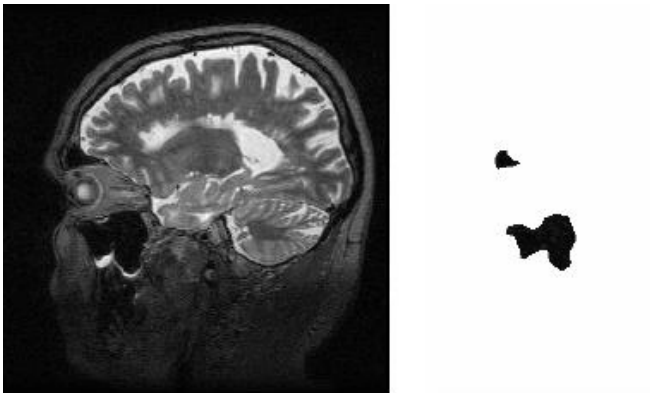
FULL PAPER

higher, it is still practical and tolerable. In addition, the improvement of, even modest, is very important for surgical planning or image-guided therapy applications. The satisfactory visual results of one-class IFWSVM are shown in Figure 4-8.



(a) Slice 86 (b) Segmentation result

Figure 7 : The segmentation result of slice 86



(a) Slice 103 (b) Segmentation result

Figure 8 : The segmentation result of slice 103

CONCLUSION

The conventional one-class SVM assumes that all the features of a sample have the same importance degree for the final segmentation result. The introduction of the feature weight concept can generate higherby suppressing features that are weakly related to the segmentation result and strengthening features that are strongly related to the segmentation result. In this study, combining one-class SVM with IA and feature weights, one-class IFWSVM is successfully implemented to segment the structure of nasopharynx in MR images. Theoretical analysis and experimental results clearly show that the one-class IFWSVM, with the optimal feature weights, has superior performance.

ACKNOWLEDGMENT

This work was supported in part by the National Natural Science Foundation of China (No. 50937005 and No. 51077040) and the Natural Science Foundation of Hebei Province, China (No. H2012202035).

REFERENCES

[1] J.Stoitsis, L.Valavanis, S.G.Mougiakakou, et al.; Computer aided diagnosis based on medical image processing and artificial intelligence methods, Nuclear Instruments and Methods in Physics Research, **569(2)**, 591-595, **(2006)**.

[2] Daw Tung Lin, De-Cheng Pan; Integrating a mixed-feature model and multi-class support vector machine for facial expression recognition, Integrated Computer-Aided Engineering, **6**, 61-74 **(2009)**.

[3] J.Zhou, K.L.Chan, V.F.Chong, et al.; Extraction of brain tumor from MR images using one-class support vector machine, The Conference on IEEE Eng Med.Biol.Soc., 6411-6414, **(2005)**.

[4] S.Dube, S.El-Saden; Content based image retrieval for MR image studies of brain tumors, Conference Proceedings. Annual International Conference of the IEEE Engineering in Medicine and Biology Society (IEEE Cat.No.06CH37748), **12(10)**, 23-25 **(2006)**.

[5] Ding Sheng, Li Shunxin; Colonel selection algorithm for feature selection and parameters optimization of support vector machines [J].2009 2nd International Symposium on Knowledge Acquisition and Modeling.KAM 2009, **2**, 17-20 **(2009)**.

[6] Xiyan He; Feature subspaces selection via one-class SVM: application to textured image segmentation [J].2010 2nd International Conference on Image Processing Theory, Tools and Applications, **5**, 12-15 **(2010)**.

[7] S.Y.Zhang, M.M.Hossain, M.R.Hassan et al.; Feature weighted SVMs using receiver operating characteristics, In.Proc.of SDM, 497-508, **(2009)**.

[8] V.Cutello, G.Niscosia, M.Pavone, J.Timmis; An immune algorithm for protein structure prediction on lattice models, IEEE Trans.Evol.Comput., **11(1)**, 101-117 **(2007)**.

[9] Yinshan Jia, Chuanying Jia; A weighted support vector machine (SVM) classification algorithm [J].

Computer engineering, **10(15)**, 23-25 **(2005)**.

[10] Tieying Tang; Study on artificial immune algorithm and its application in electric power system [D].Zhejiang: Zhejiang University, **(2007)**.

[11] Hongwei Mo; Principle and application of artificial immune system [M].Harbin: Harbin industrial university press, **(2002)**.

[12] Lei Guo; Research on 3D reconstruction for real head model based on SVM, Ph.D.dissertation, Hebei University of Technology, Tianjin, China, **(2007)**.