

# A GENETIC ALGORITHM BASED WRAPPER FEATURE SELECTION METHOD FOR CLASSIFICATION OF HYPERSPECTRAL IMAGES USING SUPPORT VECTOR MACHINE

Li Zhuo<sup>a</sup>, Jing Zheng<sup>b,\*</sup>, Fang Wang<sup>a,c</sup>, Xia Li<sup>a</sup>, Bin Ai<sup>a</sup>, Junping Qian<sup>a</sup>

<sup>a</sup> School of Geography and Planning, Sun Yat-sen University, Guangzhou 510275, China  
– (zhuoli, wangfang, lixia, aibin, qianjunping) @mail.sysu.edu.cn

<sup>b</sup> Guangdong Climate Centre, Guangzhou 510080, China – zhengjing@grmc.gov.cn

<sup>c</sup> School of Geographical sciences, Guangzhou University, Guangzhou 510006, China)

Commission VII, WG VII/3

**KEY WORDS:** Feature Selection, Hyperspectral, Genetic Algorithm, Supported Vector Machine

## ABSTRACT:

The high-dimensional feature vectors of hyper spectral data often impose a high computational cost as well as the risk of “over fitting” when classification is performed. Therefore it is necessary to reduce the dimensionality through ways like feature selection. Currently, there are two kinds of feature selection methods: filter methods and wrapper methods. The form kind requires no feedback from classifiers and estimates the classification performance indirectly. The latter kind evaluates the “goodness” of selected feature subset directly based on the classification accuracy. Many experimental results have proved that the wrapper methods can yield better performance, although they have the disadvantage of high computational cost. In this paper, we present a Genetic Algorithm (GA) based wrapper method for classification of hyper spectral data using Support Vector Machine (SVM), a state-of-art classifier that has found success in a variety of areas. The genetic algorithm (GA), which seek to solve optimization problems using the methods of evolution, specifically survival of the fittest, was used to optimize both the feature subset, i.e. band subset, of hyper spectral data and SVM kernel parameters simultaneously. A special strategy was adopted to reduce computation cost caused by the high-dimensional feature vectors of hyper spectral data when the feature subset part of chromosome was designed. The GA-SVM method was realized using the ENVI/IDL language, and was then tested by applying to a HYPERION hyper spectral image. Comparison of the optimized results and the un-optimized results showed that the GA-SVM method could significantly reduce the computation cost while improving the classification accuracy. The number of bands used for classification was reduced from 198 to 13, while the classification accuracy increased from 88.81% to 92.51%. The optimized values of the two SVM kernel parameters were 95.0297 and 0.2021, respectively, which were different from the default values as used in the ENVI software. In conclusion, the proposed wrapper feature selection method GA-SVM can optimize feature subsets and SVM kernel parameters at the same time, therefore can be applied in feature selection of the hyper spectral data.

## 1. INTRODUCTION

The rich and detailed spectral information provided by hyperspectral images can be used to identify and quantify a large range of surface materials which cannot be identified by multispectral images. However, the high-dimensional feature vectors of hyperspectral data also impose high computational cost as well as the risk of “over fitting” when classification is performed (Chen et al., 2006). Therefore it is necessary to reduce the dimensionality of hyperspectral data, which can be done by feature selection (Yang et al., 2006; Wang et al., 2006). Currently, there are two kinds of feature selection methods: filter methods and wrapper methods (Liu & Zheng, 2006). The filter methods require no feedback from classifiers and estimate the classification performance by some indirect assessments, such as distance measures which reflect how well the classes separate from each other. The wrapper methods, on the contrary, are classifier-dependent. The “goodness” of the selected feature subset are evaluated directly based on the classification accuracy. The main advantage of the filter methods is that the computation cost is low, because only a small number of features are used in classification. However, a small number of

features, even the “best” ones, do not necessarily guarantee high classification accuracy (Cover, 1974; Elashoff et al., 1967; Toussaint, 1971). The wrapper methods can yield better performance, which, as a matter of fact, has been proved by many experimental results (Kohavi & John, 1997; Huang & Wang, 2006; Mao, 2004; Yu & Cho, 2006; Bi et al., 2003; Sikora & Piramuthu, 2007). But the high computational complexity involved have limited their applications.

Currently, feature selection for hyperspectral data primarily use the filter methods, which, however, can not meet the need of high classification accuracy. Meanwhile, fast and efficient computation methods have been developing, therefore more and more researches are focusing on the wrapper methods.

As a state-of-art classifier that has found success in a variety of areas, Support Vector Machines (SVM) were frequently used as the classifier in the wrapper feature selection method (Vapnik, 1995; Cortes & Vapnik, 1995; Bradley et al., 1998; Pontil & Verri, 1998; Guyon et al., 2002; Mao, 2004). Among the many wrapper algorithms used, the Genetic Algorithm (GA), which solves optimization problems using the methods of evolution, specifically “survival of the fittest”, has proved as a promising one thanks to its prominent capability in solving global

---

\* Corresponding author: zhengjing@grmc.gov.cn

optimization problems (Raymer et al., 2000; Yang & Li, 2007; Yu & Cho, 2003). However, existing GA-based wrapper methods were primarily developed for non spatial datasets. There is few wrapper feature selection method for spatial data, especially hyperspectral data. And most of the existing studies focused on feature subset only. But the efficiency and accuracy of the SVM classifier are not only affected by the feature subset, but also the kernel function, or the kernel parameters if the function has been specified (Weston et al., 2001). Moreover, feature subset and kernel parameters should be optimized simultaneously to get the optimal results (Huang & Wang, 2006). Existing methods for optimizing the kernel parameters, like the leave-one-out method, the Step-size Gradient-descent Algorithm etc., are not ideal solutions for reasons of large computation cost, localized optimal results, or requirement of proper initial values (Lee & Lin, 2000; Bengio, 2000; Keerthi, 2002).

In this paper, we present a new wrapper feature selection method for hyperspectral data, which integrates the Genetic Algorithm and the SVM classifier through properly designed chromosome and fitness function. The purpose is to optimize both the feature subset, i.e. band subset, of hyperspectral data and SVM kernel parameters simultaneously and finally achieve higher classification accuracy.

## 2. GENETIC ALGORITHMS FOR FEATURE SELECTION USING SVM

### 2.1 Brief introduction to SVM

Support Vector Machines (SVM) is a classification system derived from statistical learning theory. It has been applied successfully in fields such as text categorisation, hand-written character recognition, image classification, biosequences analysis, etc.

The SVM separates the classes with a decision surface that maximizes the margin between the classes. The surface is often called the optimal hyperplane, and the data points closest to the hyperplane are called support vectors. The support vectors are the critical elements of the training set. The SVM can be adapted to become a nonlinear classifier through the use of nonlinear kernels. While SVM is a binary classifier in its simplest form, it can function as a multiclass classifier by combining several binary SVM classifiers (creating a binary classifier for each possible pair of classes). For multiclass classification, the pairwise classification strategy is often used. The output of SVM classification is the decision values of each pixel for each class, which are used for probability estimates. The probability values represent "true" probability in the sense that each probability falls in the range of 0 to 1, and the sum of these values for each pixel equals 1. Classification is then performed by selecting the highest probability.

SVM includes a penalty parameter that allows a certain degree of misclassification, which is particularly important for non-separable training sets. The penalty parameter controls the trade-off between allowing training errors and forcing rigid margins. It creates a soft margin that permits some misclassifications, such as it allows some training points on the wrong side of the hyperplane. Increasing the value of the penalty parameter increases the cost of misclassifying points and forces the creation of a more accurate model that may not generalize well.

There are different types of SVM classifier kernel functions, such as linear, polynomial, radial basis function (RBF), and sigmoid. In this study, we chose the radial basis function kernel because it works well in most cases and has only two parameters, C and  $\gamma$  (Hsu, Chang, & Lin, 2003).

### 2.2 Brief introduction to GA

Genetic algorithms (GA) is a general adaptive optimization search methodology based on a direct analogy to Darwinian natural selection and genetics in biological systems. It has been proved to be a promising alternative to conventional heuristic methods. Based on the Darwinian principle of 'survival of the fittest', GA works with a set of candidate solutions called a population and obtains the optimal solution after a series of iterative computations.

GA evaluates each individual's fitness, i.e. quality of the solution, through a fitness function. The fitter chromosomes have higher probability to be kept in the next generation or be selected into the recombination pool using the tournament selection methods. If the fittest individual or chromosome in a population can not meet the requirement, successive populations will be reproduced to provide more alternate solutions. The crossover and mutation functions are the main operators that randomly transform the chromosomes and finally impact their fitness value. The evolution will not stop until acceptable results are obtained. Associated with the characteristics of exploitation and exploration search, GA can deal with large search spaces efficiently, and hence has less chance to get local optimal solution than other algorithms.

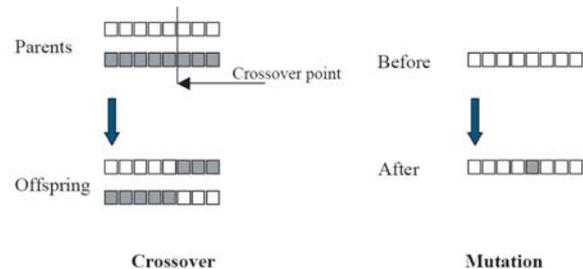


Figure. 1. Illustration of the crossover and mutation operators (Huang & Wang, 2006).

Figure. 1 illustrates the genetic operators of crossover and mutation. Crossover, the critical genetic operator that allows new solution regions in the search space to be explored, is a random mechanism for exchanging genes between two chromosomes using the one point crossover, two point crossover, or homologue crossover. In mutation the genes may occasionally be altered, for example, changing the gene value from 0 to 1 or vice versa in a binary code chromosome. (Goldberg, 1989; Davis, 1991).

### 2.3 GA-based feature selection and parameter optimization

Figure. 2 depicts the system architectures of the proposed GA-based feature selection and parameters optimization for Support Vector Machines. The fundamental steps of the GA-SVM model include: i) designing a chromosome that consists of the feature subset and the kernel parameters; ii) designing a proper

fitness function that combines the goals of high classification accuracy and low computation cost into one.

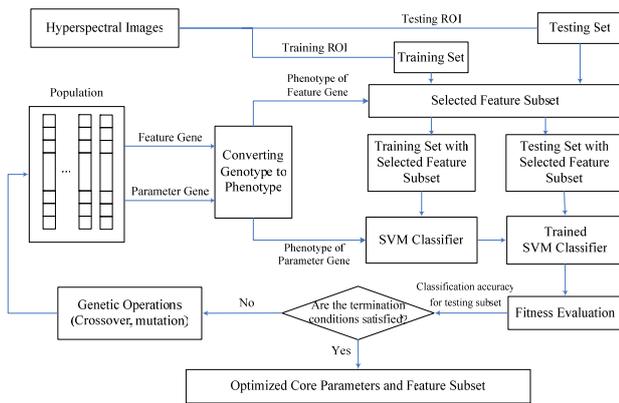


Figure.2. Flowchart of the GA-SVM wrapper approach

### 2.3.1 Design of chromosome

To optimize the kernel parameters and feature subset simultaneously, the chromosome was designed to comprise three parts,  $C$ ,  $\gamma$ , and the features mask. The binary coding system was used to represent the chromosome. Fig. 3 shows the design of the binary chromosome.

$C$			$\gamma$			$f$		
$C_1$	... $C_j$ ...	$C_{nc}$	$\gamma_1$	... $\gamma_j$ ...	$\gamma_{n\gamma}$	$f_1$	... $f_k$ ...	$f_{nf}$

Figure.3. Coding of chromosome

In Figure. 3,  $C_i$  represents the  $i$ th bit's value of bit string that represents parameter  $C$ , and  $nc$  is the number of bits representing parameter  $C$ ;  $\gamma_j$  represents the  $j$ th bit's value of bit string that represents  $\gamma$ , and  $n_\gamma$  is the number of bits representing parameter  $\gamma$ ;  $f_k$  represents the mask value of  $k$ th feature, and  $nf$  is the number of bits representing the selected features.  $nc$ ,  $n_\gamma$  and  $nf$  can be modified according to the calculation precision and/or efficiency required.

The bit strings representing the genotype of parameter  $C$  and should be transformed into phenotype by Eq. (1). Note that the precision of representing parameter depends on the length of the bit string ( $nc$  and  $n_\gamma$ ), and the minimum and maximum value of the parameter is determined by the user. For chromosome representing the feature mask, the bit with value '1' represents the feature is selected, and '0' indicates feature is not selected.

$$\Gamma(R, l) = \min_R + \frac{\max_R - \min_R}{2^l - 1} \times d \quad (1)$$

$\Gamma(R, l)$  represents phenotype of the chromosome (or part of it) that has  $l$  bits, i.e. genes.  $\min_R$  and  $\max_R$  represents the minimum and maximum value of parameter  $R$ ;  $d$  represents the decimal value corresponding to the bit string; the number of bits representing the parameter  $l$  can be modified according to the calculation precision required.

Because the number of features of hyperspectral data, i.e., the number of bands, is so large, traditional binary coding method will produce an enormously large solution space, which makes it impossible to find the optimal feature subset. Therefore a novel binary coding strategy was adopted, which set the length of chromosome different to the number of features. Suppose that the length of chromosome for the feature subset is  $nf$ ,  $nf$  features were randomly selected from all the feature subset and sorted according to their identity number. A  $nf$  bits bit string was then generated and used as a mask for the selected  $nf$  features. Thus the number of solutions decreased from  $2^n$  to  $2^{nf} \times C_n^{nf}$ . When  $nf$  is much less than  $n$ , computation efficiency will be improved greatly.

### 2.3.2 Design of fitness function

A fitness function is needed in the Genetic Algorithm to evaluate whether an individual is "fit" to survive. In the GA-SVM model, we used two criteria, namely classification accuracy and the number of selected features, to design the fitness function. The principle is that individuals with high classification accuracy and small number of features has a high fitness value, and thus high probability to be pass its genes to the next generation. A single objective fitness function that combines the two goals into one was designed to solve the multiple criteria problem. The formula is as below.

$$fitness = w_a \times accuracy + \frac{w_f}{\sum_{i=1}^{n_f} f_i} \quad (2)$$

Where  $w_a$  represents the weight value for classification accuracy,  $w_f$  for the number of features;  $f_i$  is the mask value of the  $i$ th feature, '1' represents that feature  $i$  is selected; '0' represents that feature  $i$  is not selected;  $w_a$  can be adjusted to 100% if accuracy is the most important. Generally,  $w_a$  can be set from 75 to 100% according to user's requirements. In our study, we set  $w_a$  to 80% and  $w_f$  to 20%. It can be inferred that high fitness value is determined by high classification and small feature number.

### 2.3.3 Basic Steps of the GA-SVM method

The basic steps of the GA-SVM method are as below:

- 1) Create a initial population of certain size, i.e. a group individuals with different chromosomes. Each individual's chromosome consists of three parts, namely parameter  $C$ , parameter  $\gamma$  and band subset of the hyperspectral data. The chromosomes of the initial population were randomly created. The size of the initial population should be determined properly by user to include as many possible solutions as possible.
- 2) Calculate the fitness value of each individual in the initial population using Eq. (2) and rank them according to their fitness. To calculate the fitness value of an individual, or a chromosome, the genotypes are firstly converted to phenotypes, i.e. converting the binary codes to the parameter  $C$  and  $\gamma$ , and the identities of the selected features; These values, together with the training sets of the hyperspectral image, are then used as input to the SVM classifier to perform classification; After that, the classification accuracy is evaluated based on the testing sets; finally, fitness value of the individual is calculated using Eq. (2) based on the classification accuracy and number of selected features.

- 3) Select a certain number of individuals with high fitness value as “elitism” of the population and retain them in the next generation. It is in this way that the individuals with high fitness value are retained in the population and the principle of “survival of the fittest” of the Genetic Algorithm is conveyed.
- 4) Check whether the termination conditions are satisfied. If so, the evolution stops and the optimal result represented by the best individual is returned. Otherwise, the evolution continues and the next generation is produced. The termination conditions can be either a predefined fitness threshold or number of generation evolved.
- 5) If the population continue to evolve, the next generation are produced following procedure below. First, a certain number of individuals are selected randomly to compete the mating right. Two individuals of the highest fitness values are selected as a pair of parents. Crossover is operated on their chromosomes to produce two children individuals. Location of the crossover point on the chromosomes is also randomly determined; Second, a floating-point number in the range of 0.0 to 1.0 is generated randomly. If it is less than the predefined mutation possibility, mutation is operated for the two children individuals; Repeat the two steps before to produce all the children individuals (except the “elitism” individuals ) in the new generation.
- 6) Repeat operations from step 2) to step 4).

#### 2.3.4 Realization of the GA-SVM

The GA-SVM method was realized in the ENVI/IDL language. IDL (Interactive Data Language) is an array-oriented, interpreted, structured programming language that provides powerful data analysis and visualization capabilities. ENVI (Environment for Visualizing Images) is a software for the visualization, analysis, and presentation of all types of digital imagery. It is written in IDL, so most of its functionalities are provided to users in the form of functions that can be easily called in programming.

In our research, we realized functionalities like opening and closing of the hyperspectral images, extraction data in the training ROI (Regions of Interest) and testing ROI, performing classification using the SVM, as well as evaluation of classification accuracy all through calling existing function provided by ENVI. The Genetic Algorithm was realized using the Object-Oriented programming technique of IDL. Two object classes, i.e. population and chromosome, were created with their properties defined respectively. Methods were then defined for these two object classes, respectively. Methods for the population class include population initialization, tournament selection, chromosome crossover and mutation etc. Methods for the chromosome class include coding and decoding of the chromosomes, setting and getting of chromosome properties etc. Finally, the GA-SVM method was programmed according to the steps described in section 2.3.3.

### 3. EXPERIMENTS

#### 3.1 Data Sets

The hyperspectral data used in this study is a cloudless Hyperion image taken on Dec. 18, 2005. The image centers at 113° 20' 48" E, 23° 5' 36" N and covers part of the Guangzhou city, China.

The Hyperion system on board the EO-1 platform provides a high resolution hyperspectral imager capable of resolving 220 spectral bands (from 0.4 to 2.5 μm) with a 30-meter resolution. The instrument can image a 7.5 km by 100 km land area per image, and provide detailed spectral mapping across all 220 channels with high radiometric accuracy. The Hyperion images has wide ranging applications in mining, geology, forestry, agriculture, and environmental management. Detailed classification of land assets through the Hyperion will enable more accurate remote mineral exploration, better predictions of crop yield and assessments, and better containment mapping.

In this study, we selected part of a Hyperion image that covers the Haizhu district of Guangzhou city to test the GA-SVM method. Atmospheric correction and geometric correction of the image were first performed in order to eliminate atmospheric effects and compare to the ground truth data.

#### 3.2 Feature selection for the Hyperion image using the GA-SVM method

The steps of feature selection for the Hyperion image using the GA-SVM method include:

- 1) Create the training sets and testing sets needed by the GA-SVM method using the ENVI software. Land uses of the study area were categorized into six classes, namely built-up area, water body, grassland, forest and unused land. Training sets and testing sets were created for each land use class.
- 2) Set parameters of the GA-SVM method. These parameters include: range of kernel parameter C and  $\gamma$ , bit lengths for the three chromosome parts, initial population size, population size of each generation, elitism size, number of generation to evolve, tournament size, crossover rate and mutation rate. These parameters were set as follow:  $min\_C = 90.0$ ,  $max\_C = 110.0$ ,  $min\_gamma = 0.9/bit\_length$ ,  $max\_gamma = 0.25$ ,  $initPopulationSize = 100$ ,  $populationSize = 50$ ,  $numGenerations = 40$ ,  $offspringPerGen = 46$ ,  $tournamentSize = 6$ ,  $crossOverRate = 0.98$ ,  $elitism = 4$ ,  $mutationRate = 0.02$ .
- 3) Run the GA-SVM model with the inputs of Hyperion image, training sets and testing sets.

	c	γ	Selected Bands ( with their central wavelength, unit: mm)	Classification Accuracy	Kappa Coefficient
No Feature Selection	100.000	0.000	All 196 bands	88.81%	0.8619
Feature Selection using the GA-SVM Method	95.0297	0.2021	4(457.34), 12(538.74), 25(671.02), 33(752.43), 61(1033.50), 67(1094.09), 75(1174.77), 88(1305.96), 106(1487.53), 109(1517.83), 117(1598.51), 127(1699.40), 132(1749.79)	92.51%	0.9261

Table 1. Performance of the GA-SVM Feature Selection

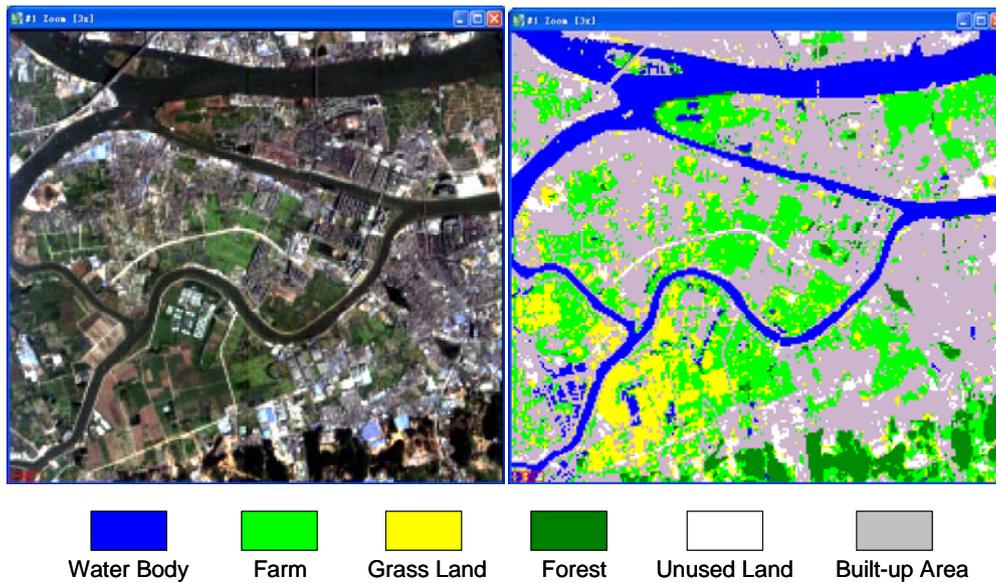


Figure 4. the input image and classification result of the GA-SVM

### 3.3 Results

Outputs of the GA-SVM model include optimized kernel parameters and feature subset, classification result, evaluation of the classification accuracy, as shown in Table 1. The classification result is shown in Figure 4. The evolution of fitness value with number of generation is shown in Figure 5.

It can be seen in Table 1 that the number of features, i.e. bands, selected for the use of classification was only 13, much smaller than the total number of bands (198). Classification accuracy was 92.51% when using the optimized kernel parameters and feature subset, comparing to 88.81% when no feature selection was performed. The optimized values of the two SVM kernel parameters were 95.0297 and 0.2021, respectively, different from the default values of 100 and 0.005 in ENVI, which again proved the necessity of optimizing the kernel parameters.

As can be seen in Fig. 5, the fitness value increases with number of generation. There are sharp increase at the 3rd and 30th generation, although the fitness value may remain unchanged during several generations. The acceptable fitness level was reached at the 30th generation.

Although the maximum number of generation evolved was only 40 in our study due to limited computer memory. The GA-SVM method has proved to be able to return acceptable result under the limited memory resource. It may be a safe assumption that provided with larger computer memory, better results may be gained.

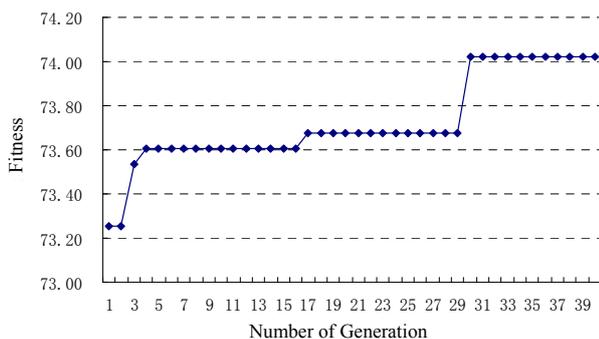


Figure 5. Change of fitness with population evolution

### 4. CONCLUSION AND DISCUSSION

In this paper we proposed a GA based wrapper feature selection method GA-SVM for hyperspectral data. The Genetic Algorithm was used to optimize the kernel parameters and feature subsets simultaneously. Realized using the ENVI/IDL programming language, the GA-SVM method was tested using the HYPERION hyperspectral image.

The results showed that the GA-SVM method could significantly reduce the computation cost while improving the classification accuracy. The number of bands used for classification was reduced from 198 to 13, while the classification accuracy increased from 88.81% to 92.51%. The optimized values of the two SVM kernel parameters were 95.0297 and 0.2021, respectively, which were different from the default values as used in the ENVI software.

In conclusion, the proposed wrapper feature selection method GA-SVM can optimize feature subsets and SVM kernel parameters at the same time, therefore can be applied in feature selection of the hyper spectral data.

### REFERENCES

Bengio Y., 2000. Gradient-based Optimization of Hyper-parameters. *Neural Computation*, 12(8), pp. 1889-1900.

Bi J., Bennett K.P., Embrechts M., Breneman C.M., Song M., 2003. Dimensionality reduction via sparse support vector machines. *J. Mach. Learn. Res.* 3, pp. 1229-1243.

Bradley P. S., Mangasarian O. L., Street W. N., 1998. Feature selection via mathematical programming. *INFORMS Journal on Computing*, 10, pp. 209-217.

Chen M., Yi Y. H., Liu Zh. G., Li D. R., QIN Q. Q., 2006. Study of Band Selection Methods of Hyperspectral Image

- Based on Grouping Property. *Bulletin of Surveying and Mapping*, 3, pp. 10-13. (In Chinese)
- Cortes C., Vapnik V.N., 1995. Support vector networks. *Mach. Learn.*, 20(3), pp. 273-297.
- Cover T.M., 1974. The best two independent measurements are not the two best. *IEEE Transactions on Systems, Man, and Cybernetics*, 1, pp. 116-117.
- Elashoff, J.D., Elashoff, R.M., Goldman, G.E., 1967. On the choice of variables In classification problems with dichotomous variables. *Biometrika*, 54, pp. 668-670.
- Guyon I., Weston J., Barnhill S., Bapnik V., 2002. Gene selection for cancer classification using support vector machines. *Machine Learning*, 46(1-3), pp. 389-422.
- Hsu C. W., Chang C. C., Lin C. J., 2003. A practical guide to support vector classification, <http://www.csie.ntu.edu.tw/~cjlin/papers/guide/guide.pdf>.
- Huang Ch. L., Wang Ch. J., 2006. A GA-based feature selection and parameters optimization for support vector machines. *Expert Systems with Applications*, 31, pp. 231-240
- Keerthi S. S., 2002. Efficient Tuning of SVM Hyperparameters Using Radius/Margin Bound and Iterative Algorithms. *IEEE Transactions on Neural Networks*, 13(5), pp. 1225-1229.
- Kohavi R., John G. H., 1997. Wrappers for feature subset selection. *Artif. Intell.*, 97, pp. 273-324.
- Lee J. H., Lin C. J., 2000. *Automatic Model Selection for Support Vector Machines*. Taiwan University, Taipei.
- Liu Y., Zheng Y. F., 2006. FS\_SFS: A novel feature selection method for support vector machines. *Pattern Recognition*, 39, pp. 1333-1345.
- Mao K. Z., 2004. Feature subset selection for support vector machines through discriminative function pruning analysis. *IEEE Transactions on Systems, Man, and Cybernetics*, 34(1), pp. 60-67.
- Pontil M., Verri A., 1998. Support vector machines for 3D object recognition. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 20(6), pp. 637-646.
- Raymer M.L., Punch W.F., 2000. Goodman E.D., Kuhn L.A., Jain A.K. Dimensionality reduction using genetic algorithms. *IEEE Trans. Evol. Comput.*, 4(2), pp. 164-171
- Sikora R., Piramuthu S., 2007. Framework for efficient feature selection in genetic algorithm based data mining. *European Journal of Operational Research*, 180(2), pp. 723-737.
- Toussaint, G.T., 1971. Note on optimal selection of independent binary-valued features for pattern recognition. *IEEE Transactions on Information Theory IT*, 17, pp. 618.
- Vapnik, V.N., 1995. *The Nature of Statistical Learning Theory*. Springer-Verlag, New York.
- Wang C. Y., Liu Zh. J., Yan Ch. Y., 2006. A Experimental Study on Imaging Spectrometer Data Feature Selection and Wheat Type Identification. *Journal of Remote Sensing*. 10(2), pp. 249-255.
- Weston J., Mukherjee S., Chapelle O., Pontil M., Poggio T., Vapnik V., 2001. *Feature selection for SVMs. Advances in Neural Information Processing Systems*, MIT Press, Cambridge, MA.
- Yang Q. Sh., Li X., 2007. Calibrating urban cellular automata using genetic algorithms. *Geographical Research*, 26(2), pp. 226-231. (In Chinese)
- Yang Zh. H., Zhang Y. Zh., Gong D. P., Li Zh. X., Han J. F., 2006. Feature Selection in Hyperspectral Classification Based on Tabu Search Algorithm. *Hydrographic Surveying and Charting*, 26(4), pp. 11-14. (In Chinese)
- Yu, E., Cho, S., 2003. GA-SVM wrapper approach for feature selection in keystroke dynamics identity verification. In: *Proceedings of 2003 INNS-IEEE International Joint Conference on Neural Networks*, pp. 2253-2257.
- Yu E., Cho S., 2006. Ensemble based on GA wrapper feature selection. *Computers & Industrial Engineering*, 51(1), pp. 111-116.