

APPLICATION OF ECOC SVMs IN REMOTE SENSING IMAGE CLASSIFICATION

Zhigang YAN^{a, b, *}, Yuanxuan YANG^{a, b}

^a School of Environmental Science and Spatial Informatics, China University of Mining and Technology, Xuzhou, Jiangsu, 221116
P.R.China. - zhg-yan@163.com

^b Jiangsu Key Laboratory of Resources & Environmental Information Engineering, China University of Mining and Technology,
Xuzhou, Jiangsu, 221116 P.R.China - zhg-yan@163.com

Technical Commission II

KEY WORDS: Remote Sensing Classification, ECOC-SVMs, Multi-class Classifier

ABSTRACT:

Image processing has been one of the efficient technologies for GIS data requisition. Support Vector Machines (SVMs) have peculiar advantages in handling problems with small sample sizes, nonlinearity, and high dimensionality. However, SVMs can only solve two-class problems while multi-class decision is impossible. Error correcting output coding (ECOC) SVMs enhance the ability of fault tolerance when solving multi-class classification problems, which makes ECOC SVMs suitable for remote sensing image classification. In this paper, the generalization ability of ECOC SVMs is discussed. ECOC SVMs with optimum coding matrices are selected by experiment, and applied to remote sensing image classification. Experimental results show that, compared with Conventional multi-class classification methods, less SVM sub-classifiers are needed for ECOC SVMs in remote sensing image classification, and the classification accuracy is also improved.

1. INTRODUCTION

Image processing has been one of the efficient technologies for GIS data requisition. Support Vector Machine (SVM) method has been widely and extensively studied recently, and has been successfully applied in remote sensing image classification (Custavo et al., 2004) and (ZHU et al., 2002). However, it was originally proposed for two-class classification problems, which is not suitable for remote sensing image classification with multiple geographical object classes. Dietterich proposed a framework in which the error correcting output coding (ECOC) was used to convert multi-class problem into several two-class problems (Dietterich et al., 1995). The combination of it with SVMs, ECOC SVMs, enhances the fault tolerance of classification models when solving the multi-class classification problems, which makes ECOC SVMs effective in remote sensing image classification. In this study, the generalization ability of ECOC SVMs is analyzed, and new ideas on how to construct ECOC SVMs are suggested and applied in remote sensing image classification.

2. ECOC SVMs

2.1 Principles of error correcting output coding

ECOC is a kind of method which converts a multi-class classification problem into two-class classification problems. In ECOC, the code bit values in each column of the error correcting code matrix are used to reclassify the training samples. Then several unrelated two-class sub-classifiers are constructed. The classification results of the testing samples are decided by specific decoding strategy.

In ECOC, a k -class classification problem is converted into n two-class problems. Firstly, a $k*n$ binary coding matrix is needed, each row of which is called a code word and is used to represent a sample class; each column binary number is called a code bit, the number of bits in a code word is called code

length. An error correcting coding matrix for class number $k=4$ and code length $n=7$ is explained in Table 1, where f_1, f_2, \dots, f_7 represent the different decision functions for dichotomizing the 4 classes into 2 classes signed by the code bit.

class	ECOC						
	f_1	f_2	f_3	f_4	f_5	f_6	f_7
1	0	0	0	0	0	0	0
2	1	1	1	1	0	0	0
3	1	1	0	0	1	1	0
4	1	0	1	0	1	0	1

Table 1. An error correcting coding matrix for class number $k=4$ and code length $n=7$

2.2 Coding and decoding of ECOC

2.2.1 Coding: An available ECOC should satisfy the following conditions:

(1) The rows of a coding matrix should not be correlated, while the columns of a coding matrix should be neither correlated, nor complementary;

(2) A column should be neither all "0", nor all "1";

In order to have available ECOCs satisfy the above conditions, exhaustive enumeration coding method, BCH coding method, column selection method, random coding method (Dietterich et al., 1995), and searching coding method are commonly used (Jiang et al., 2005).

* Corresponding author. Associate professor, research fields is spatio-temporal data mining and knowledge discovering.

The above-mentioned methods have their own characteristics: the ECOCs generated by enumeration coding method have strong error correcting capability, however the code length increases as the class number increases and the learning process is complicated. The numbers of code words in the coding matrix obtained by BCH coding method are always power of 2. Therefore if the class number m is not the power of 2, some heuristic methods are needed for shortening the code length and reducing the code words, which makes this method inconvenient to use. Column selection method and random coding method are non-deterministic methods. Searching coding method is suitable for any number of classes, moreover it can automatically generate alternative codes according to different indices such as class number, minimum Hamming distance, etc. However it cannot overcome the problem of identical columns. At present, there is no universal deterministic coding method which can obtain available ECOCs for any class number K .

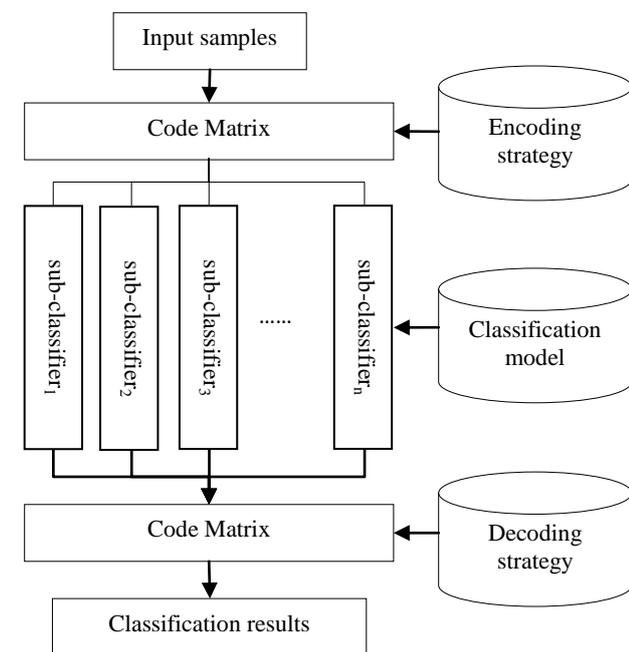


Figure 1. ECOC algorithm framework

The two-class classification results are fused in order to output the multi-class classification result. Classical decoding methods include: Hamming distance decoding and the inverse Hamming distance decoding which is improved version of the former, as well as Euclidean distance decoding, center distance decoding, loss function based decoding, etc. (Escalera et al., 2010).

Among the above methods, the principles of Hamming distance decoding can be described as follows: for a K -class classification problem, combine the binary results of L basic

classifiers in order, obtaining the binary classification code C for each classification object. The new sample's class is represented by the code word which has the minimum Hamming distance with C among the K standard codes.

2.3 Experiments on generalization ability of ECOC SVMs

It is generally believed that, for coding performance evaluation, the performance of ECOCs is related to the mathematical characteristics, such as code length, minimum inter-code Hamming distance, and the distribution order of code words (Xia et al., 2003). To verify the relationship between the generalization ability of ECOC SVMs and the mathematical characteristics of ECOC coding as well as the generalization ability of each sub-SVM classifier, experiments are carried out in this study for assessment.

2.3.1 Tests on UCI data base: The Page Blocks, Landsat Satellite, Zoo, Segment and Optdigits data sets in UCI data base are chosen for testing. The details of each data set are tabulated in Table 2.

Data set	Class number	Code length for test
Page Blocks	5	1-15
Landsat Satellite	6	1-31
Zoo	7	1-63
Segment	7	1-63
Optdigits	10	1-150

Table 2. The explanation on UCI data base used in the tests

The testing process is described here. For given samples in n classes, construct an n by k ECOC coding matrix (k is the minimum code length) using exhaustive enumeration coding method, then separately train SVM sub-classifiers according to each column of the ECOC coding matrix.

ECOC SVMs composed by codes from column 1 to k in each data set are tested in the experiments. In the experiments, all SVM sub-classifiers unitedly use RBF kernel function with identical parameters. The prediction accuracies of the SVM sub-classifiers are used as the index for generalization ability to sort the SVM sub-classifiers. The sub-classifiers are sorted in forward, original (random), and backward sequences. Take the first $m=CL$ SVM sub-classifiers of all the sorted SVM sub-classifiers according to code length. The predicted results after decoding are compared with the real results, and then the prediction accuracies using different sorting methods are calculated and shown in Figure 2.

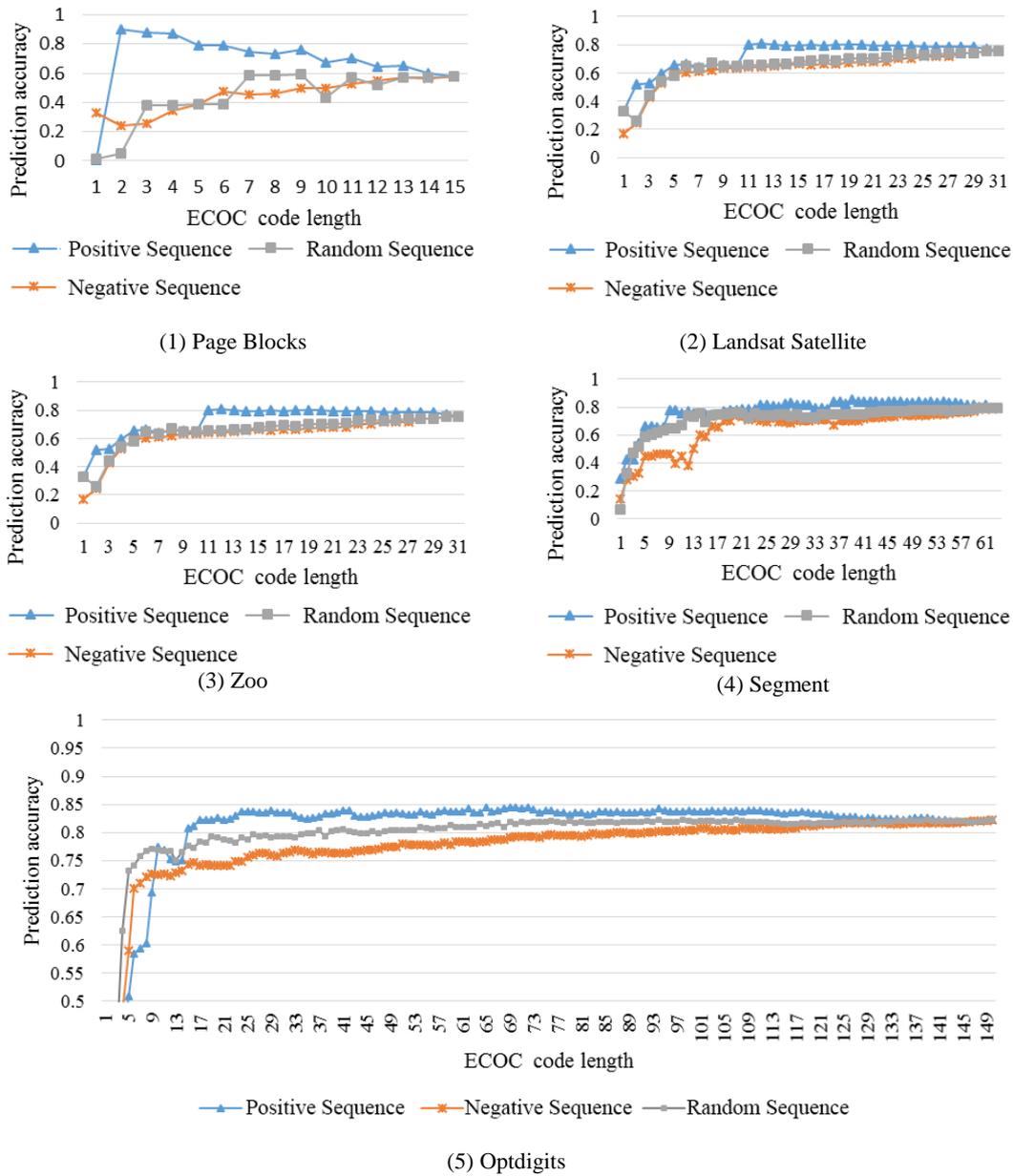


Figure 2. the prediction accuracies according to the different sorting sequences of sub-classifiers

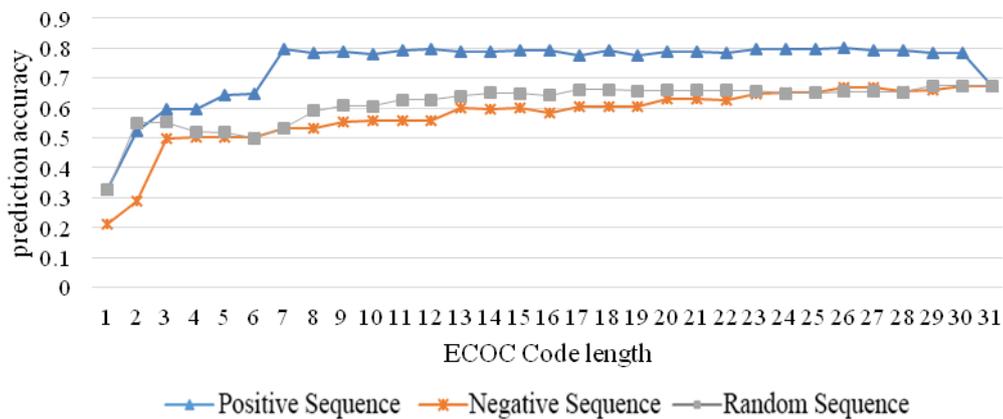


Figure 3. Testing results on Landsat-TM images

2.3.2 Test on remote sensing data: Extract pixels in different geographical object classes with true class labels in Landsat-TM images. Twenty training samples and 200 test samples are selected for each class of pixels. The testing process is the same as the previous. For a 6-class test sample set with code length from 1 to 31, the prediction accuracies using different sorting methods are shown in Figure 3.

2.3.3 A experimental result for the generalization ability of ECOC SVMs: Analyzing the generalization ability results of ECOC SVMs for UCI database and remote sensing image data, it can be concluded that after sorting the generalization ability of SVM sub-classifiers, the prediction accuracy of the forward sequence of ECOCs is higher than that of the backward sequence, whereas the prediction accuracy of the original sequence (which can be viewed as random ECOCs) generally lies between the other two. The more overlapped code words between the forward, backward sequences, the closer the prediction accuracies of them. When the code length increases to a certain extent, the prediction accuracy of forward sequence decreases and finally becomes stable. The random sequence increases with fluctuations and finally becomes smooth, whereas the backward sequence keeps an increasing trend.

The above analysis shows that when the code length increases, if the generalization ability of the ECOC SVMs corresponding to the newly added columns is good, the coding performance is improved, such as forward sequences; if the generalization ability is bad, the coding performance worsens, such as backward sequences; for random sequences (original sequences), the generalization ability of each newly added SVM is at random, therefore the final prediction accuracy fluctuates.

It can be seen that the performance of ECOC SVMs is neither improved with the increasing of the minimum Hamming distance, nor does it deteriorates with the increasing of code length. Instead, it has a complicated relationship with code length and inter-code minimum Hamming distance. ECOC SVMs' performance depends on the performance of the corresponding sub-SVMs. The mathematical characteristics expressed in coding are minor. When the performance of all sub-SVMs is not evaluable or considered to be identical, the code length, inter-code minimum Hamming distance, and the sorting method of code words can be used as evaluation indices of ECOCs.

Therefore, for sub-SVMs with good generalization ability, ECOC coding can be short; however if the performance of sub-SVMs cannot be determined, the code length should be longer and the inter-code Hamming distance should also be longer, but the code length should be appropriate. The longer is by no means the better.

3. APPLICATION OF ECOC SVMs FOR REMOTE SENSING CLASSIFICATION

3.1 Data for experiments and tools

The remote sensing image data chosen in this study is the Landsat-TM image of someplace. The bands selected are TM1~TM7 with image size being 640×400 pixels. Figure 4 shows the combination of bands 7, 4, 2. The software tools used in the experiments are ENVI 4.8 and LibSVM toolbox.

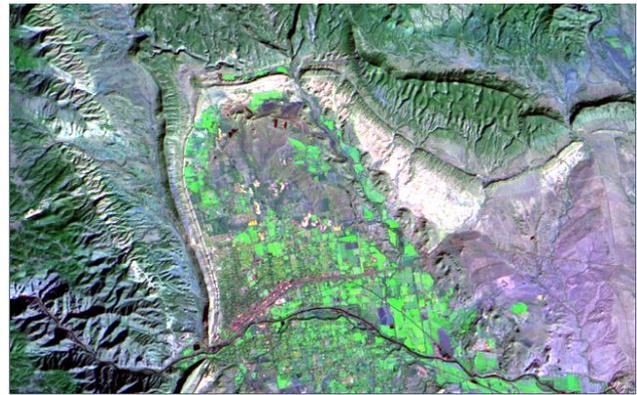


Figure 4. the remote sensing image for test

Feature extraction and selection are performed for the remote sensing image, and the image is classified by using ECOC SVMs combined with original spectral information. The main steps include relevant data preprocessing, extracting attributed feature information, generating ECOC coding matrix, selecting training samples, data normalization, SVM kernel parameter settings, ECOC coding selection, classification by ECOC SVMs and decoding, as well as the accuracy assessment of the final classification results.

Feature extraction and selection are the most significant factors for classification accuracy and reliability, as well as the classifier design and performance. In the experiments, the PCA transform information (Francesco et al., 2004), NDVI (McFeeters, 1996), and texture information (Gong et al., 1992), as well as the original spectral information are extracted and combined for classification.

3.2 Generating ECOC coding

In the experiments, according to visual interpretation and survey, the remote sensing data is classified into 5 classes, namely field, sand, bare land, forest, and shadow. After obtaining the class number K , the exhaustive enumeration coding method can be used to generate available ECOCs satisfying valid conditions, whose code length CL satisfies $\log_2 k < CL \leq 2^{K-1} - 1$. Then the ECOC coding selection operation is performed subsequently.

Using the idea in this study to test the data, we obtain the classification accuracies corresponding to the forward sequences ECOC codes, shown in Figure 5. Finally, the ECOC codes with code length being 7 are selected to construct ECOC SVMs.

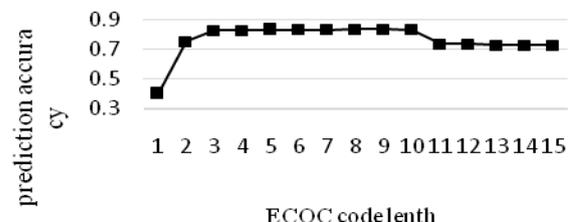


Figure 5. the prediction accuracies corresponding to the forward sequences ECOC codes

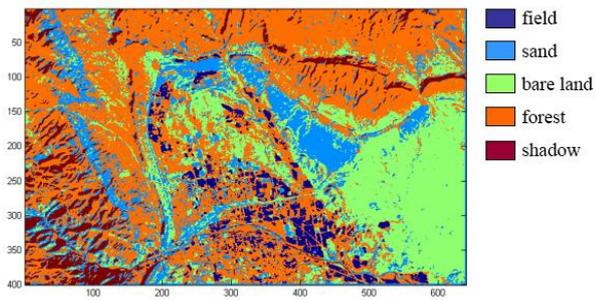


Figure 6. the classification results of the selected remote sensing image

accuracy of ECOC SVMs for the remote sensing image, we obtain the error matrix which is tabulated in Table 2 Using the same training data and test data for traditional 1-V-R and 1-V-1 SVMs for classification, and evaluating the classification accuracy, we obtain the error matrix tabulated in Table 3 and 4 The overall classification accuracies and the kappa coefficients of the three multi-class classification methods are shown in Table 5.

3.3 Result analysis

The experimental result of ECOC SVMs for the TM remote sensing image is shown in Figure 6 Evaluating the classification

Standard data type	Classified data type					Total	Product Accuracy	Omission error
	field	sand	bare land	forest	shadow			
field	354	0	0	57	0	411	354/411	57/411
sand	0	419	3	24	0	446	419/446	27/446
bare land	0	8	544	32	0	584	544/584	40/584
forest	0	51	16	444	11	522	444/522	78/522
shadow	0	3	28	24	286	341	286/341	55/341
Total after classified	354	481	591	581	297			
User Accuracy	354/354	419/481	544/591	444/581	286/297	Total data(2304)		
misclassification error	0/354	62/481	47/591	137/581	11/297			
Overall Accuracy = 0.8885								

Table 2. Error matrix of ECOC SVMs classification

Standard data type	Classified data type				
	field	sand	bare land	forest	shadow
field	354	0	0	57	0
sand	7	413	2	24	0
bare land	11	0	541	32	0
forest	53	0	14	444	11
shadow	4	0	27	24	286

Table 3. Error matrix of 1-V-R SVMs classification

Standard data type	Classified data type				
	field	sand	bare land	forest	shadow
field	323	46	0	42	0
sand	0	419	5	22	0
bare land	0	4	547	33	0
forest	0	53	13	443	13
shadow	0	3	30	21	287

Table 4. Error matrix of 1-V-1 SVMs classification

methods	Overall Accuracy	kappa coefficient
1-V-R SVMs	0.8745	0.8542
1-V-1 SVMs	0.8763	0.8436
ECOC-SVMs	0.8885	0.8590

Table 5. the overall classification accuracies and the kappa coefficients among 3 methods

It can be seen through comparison that for same experimental remote sensing data, the overall classification accuracy and kappa coefficient of ECOC SVMs are both higher than the other two

4. CONCLUSIONS

The influencing factors of ECOC SVMs performance are discussed in this study. It is pointed out that for a given training data set, reasonable ECOCs should be constructed based on SVM sub-classifiers with good generalization ability to improve the generalization performance of ECOC SVMs.

ECOC SVMs improve the fault tolerance of classification models when solving the multi-class classification problem for remote sensing images. It has the advantage of fewer sub-classifiers and high classification accuracy. Experiments show that ECOC SVMs outperforms 1-V-R SVMs and 1-V-1 SVMs.

ACKNOWLEDGEMENTS

This work was supported by a grant from Natural Scientific Fund of China (No. 41271445) and a Project Funded by the

Priority Academic Program Development of Jiangsu Higher Education Institutions.

REFERENCES

- Custavo C V, Luis G C, Javier C M, et al, 2004. Robust support vector method for hyperspectral data classification and knowledge discovery. *IEEE Transactions on Geoscience and Remote Sensing*, 42 (7), pp. 1530-1541.
- ZHU G B, DAN G, Blumberg, Classification using ASTER data and SVM algorithms: the case study of beer sheva, israel. *Remote Sensing of Environment*, 80 (2), pp. 233-240.
- Dietterich T G, Bakiri G, 1995. Solving Multiclass Learning Problems via Error-Correcting Output Codes. *Journal of Artificial Intelligence Research*, 2, pp. 263-286.
- Y. Jiang, Q. Zhao, X. Yang, 2005. A Search Coding Method and Its Application in Supervised Classification. *Journal of Software*, 16(6), pp. 1081-1089. (In Chinese).
- Escalera S, Pujol O, Radeva P, 2010. On the decoding process in ternary error-correcting output codes. *IEEE Trans on Pattern Analysis and Machine Intelligence*, 32(1), pp. 120-134.
- Xia Jiantao, He Mingyi, 2003. Multiclass Classification Using Support Vector Machines (SVMs) Combined with Error-Correcting Codes (ECCs). *Journal of Northwestern Polytechnical University*, 21(4), pp. 443-448. (In Chinese)
- Francesco Masulli, Giorgio Valentini, 2004. An experimental analysis of the dependence among codeword bit errors in ECOC learning machines. *Neurocomputing*, 57, pp. 189-214.
- McFeeters S K, 1996. The use of the Normalized Difference Water Index (NDWI) in the delineation of open water features. *International Journal of Remote Sensing*, 17(7), pp. 1425-1432.
- P. Gong, D. J Marceau, and P. J. Howarth, 1992. A Comparison of Spatial Feature Extraction Algorithms for Land-Use Classification with SPOT HRV Data. *Remote Sensing of Environment*, 40, pp. 137-151.