



Journal of
**Software
Engineering**

ISSN 1819-4311



Academic
Journals Inc.

www.academicjournals.com



Research Article

Classification of High-resolution Remotely Sensed Images Based on Random Forests

Lianjun Chen and Xinwen Cheng

Faculty of Information Engineering, China University of Geosciences (Wuhan), 430074 Wuhan City, China

Abstract

Background: Using the high resolution image to establish the remote sensing classification model and extracting the urban land information, can provide the information support for the urban land use planning and management. The accuracy of traditional classification models is unsatisfactory and there is the problem of over-fitting. With the development of the model algorithm, the Random Forest (RF) ensemble-learning algorithm has the potential to solve these problems. This study attempts to establish a high resolution remote sensing image classification model based on the random forest method and study its performance. **Materials and Methods:** This study presents a comparison between the classification results obtained with RFs and the Support Vector Machine (SVM) by using Hymap high-resolution remotely sensed image of Berlin. Before the establishment of the model we first band to pick out the right band and texture data. After that, use data selection, optimization parameter model, then RF and SVM model is established for the high resolution image classification, finally the validation process using the total classification accuracy and Kappa coefficient. **Results:** The study results show that for the RF algorithm the overall classification accuracy reaches 92.6%, the Kappa coefficient is 0.9024, for the SVM algorithm the overall classification accuracy is 91.2%, the Kappa coefficient is 0.8840. **Conclusion:** Using the random forest method for high resolution remote sensing image classification is feasible, the main performance is better than the SVM, compared with SVM, the random forest algorithm has fewer parameters, more easily parameter optimization, higher classification accuracy, easier to serve the production practice.

Key words: Random forests, support vector machine, high-resolution remotely sensed image, classification accuracy

Received: April 21, 2016

Accepted: May 13, 2016

Published: September 15, 2016

Citation: Lianjun Chen and Xinwen Cheng, 2016. Classification of high-resolution remotely sensed images based on random forests. *J. Software Eng.*, 10: 318-327.

Corresponding Author: Lianjun Chen, College of Information Engineering, China University of Geosciences, Road HongShan, 430074 Wuhan City, China
Tel: +8615871384876

Copyright: © 2016 Lianjun Chen and Xinwen Cheng. This is an open access article distributed under the terms of the creative commons attribution License, which permits unrestricted use, distribution and reproduction in any medium, provided the original author and source are credited.

Competing Interest: The authors have declared that no competing interest exists.

Data Availability: All relevant data are within the paper and its supporting information files.

INTRODUCTION

Remote sensing image classification, an important means of remote sensing information, is one of the most common aspects of remote sensing study¹. Traditional remote sensing image classification methods are the supervised classification method (such as maximum likelihood and minimum distance) and unsupervised classification method (ISODATA)². The classification method has some obvious shortcomings: Reliance on statistical model to determine the spectral images are numerical points that have to be categorized, the parameters of the statistical model can approximate out of the related data, easy fitting³. The supervised classification method used for the training data set needs to satisfy a normal distribution. Secondary data can be used to improve the accuracy of the classification, but this kind of data fusion into the traditional supervised classification method is difficult⁴. In recent years, with the development of the machine-learning algorithm, its application in remote sensing has also been increasing. Machine-learning algorithms are mainly based on the Support Vector Machine (SVM), neural networks, etc. The machine-learning algorithm does not depend on statistical models; rather, it depends on the relationship between the response data and learning classifier⁵. The machine-learning algorithm can use different kinds of data for classification (such as DEM, climate data, etc.). A Random Forest (RF) is a kind of superior performance of machine-learning algorithm of classifier⁶, which uses a multiple tree classifier. In reality, it is difficult to find a suitable classifier for a variety of application requirements and single classifier classification accuracy is not high and easily leads to fitting problems. However, if a multiple classifier combination is more accurate than a single classifier up to the collective decision, it is more able to resist noise⁷. In this study, a comparison is carried out between classification results retrieved by RF and SVM over a study area of part of Berlin, Germany, based on HYMAP Data. By combining a ground-truth imaging study area, analysis and evaluation of its classification performance, their application is promoted for remote sensing.

MATERIALS AND METHODS

Random forests: A random forest is an ensemble-learning algorithm on the basis of a classification tree and was proposed by Breiman. This algorithm involves Breiman's bagging theory⁸ and Ho's stochastic subspace identification⁹. The algorithm has the advantages of fast computation speed,

easy parameter quantization, over fitting does not easily occur and high classification accuracy¹⁰. The random forest algorithm is mainly realized by three steps: (1) N training subsets were selected from the original data set D through bootstrap random with replacement. About 36.8% of samples in the original data sets did not appear in selected samples of subsets (namely, each training subset contained $2/3$ of the original data). In addition, about $1/3$ of data, which were not selected, were considered as OOB (out of bag) data¹¹. An OOB error could be obtained to evaluate the internal error of the random forest method. The RF algorithm requires two important parameters, N and m , to be defined, where N is the number of bootstrap resampling and m is the number of variables on each node of the decision tree split. With step (1) to extract the N bootstrap sample sets, using each bootstrap sample set as the training set to generate a CART (classification and regression tree), finally, there are N CART trees. During tree generation, the attribute subset m was randomly selected from the whole attribute characteristic M when splitting each node in CART trees; then, one optimal attribute was selected to perform node splitting¹². Breiman thought the m value, which was, could produce the best result. Finally, N trained decision trees were combined together to enable voting. After that, the results of most trees were selected to make a prediction for a new sample. The result of the regression problem was the average number of the results of all decision trees. The main flow chart is shown in Fig. 1.

Random forests were obtained from different training samples by using the bootstrap sampling strategy and random selection of attribute characteristics, which increases the difference between the different classifiers of random forest and reduces their relevance, is the prerequisite for the effective integration of learning strategies. When dealing with high-dimensional data classification, it can be adjusted and optimized by random selection of a feature subset, which limits the number of variables; thus, it is highly suited to processing high-dimensional data¹³.

At present, the main random forest classification modeling tools are the Waikato environment for knowledge analysis (Weka), FORTRAN, R language, MATLAB and EnMAPBox. Weka is a free tool based in the environment of Java open-source machine learning and data mining software¹⁴. The original RF algorithm is Fortran language. The R language enables a statistical analysis of the language environment, free source code with open software. The R language has two software packages that can run random forests, namely random forest and 4.6-7 party¹⁵. The random

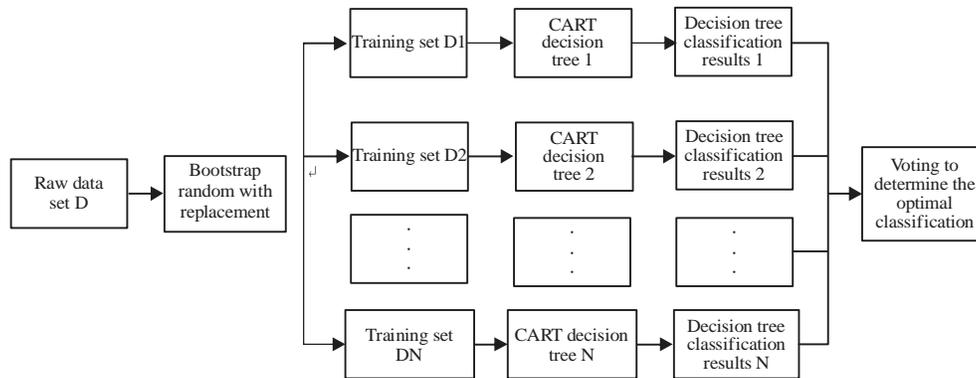


Fig. 1: Random forest flow chart

forest program package based on MATLAB platform is RF_MexStandalone-v0.02. EnMAP-Box software, which is based on IDL. The package can run on ENVI and the software is embedded in the random forest algorithm¹⁶. In this study, the research platform is 2012b MATLAB, using the RF_MexStandalone-v0.02 package for image classification study.

Support vector machine: Support Vector Machine (SVM) is a kind of machine learning algorithm based on statistical learning theory and the structural risk minimization principle of VC¹⁷. It was first proposed by Cortes and Vapnik¹⁸. The SVM can solve problems using small sample training, a high-dimensional feature space, problems of a nonlinear nature, regression problems, pattern recognition (discriminant analysis, classification)¹⁹ and so on. The principle is to find the optimal classification hyper plane, which makes the sample points along the boundary between the two types of adjacent samples the largest, with the edge of the sample points as the support vector and the intermediate section of the edge the optimal classification hyper plane. When faced with nonlinear problems, it can be transformed into a high-dimensional space and a linear discriminant function is constructed to realize the spatial nonlinear discriminant by setting the penalty factor to process the category attribution of outliers²⁰. The SVM performance depends on the choice of kernel function. There are four kinds of common kernel functions: linear, polynomial, RBF kernel function and the S-type. The Radial Basis Function (RBF) kernel is widely used in remote sensing classification²¹.

The SVM can only deal with two kinds of classification problems²², i.e., the realization of multi-class classification mainly follows two approaches: One is to decompose the multi-class classification problem into two multiple classification problems; training a certain amount of SVM and then merged to make decisions, the classification criteria are

obtained. The other is to change the original optimization problem in SVM, such that it can calculate all multi-class classification decision functions at the same time to obtain the classification criterion²³. The second strategy is too large to be calculated. Hence, it is difficult to solve the problem and the application is not widespread. The most commonly used SVM software package is libsvm, which was developed by Lin Zhiren of the National Taiwan University²⁴. The software package used in this paper is the Faruto Ultimate 3.0 version, which is an improvement of libsvm based on Faruto.

RESULTS

Experimental data: The experimental data are derived from EnMAP-Box of the sample data: Berlin-A image and Berlin-A classification ground truth. Hymap Berlin-A image data is for part of the city of Berlin, Berlin-A classification ground truth is for the Berlin-A image classification of the ground truth image. Hymap is an airborne imaging spectrum system, which was developed by the Australian HyVista company. It has 128 bands, a bandwidth of 3-10, a ground resolution of 15-16 nm and data encoding is 16 bit. Hymap data in the study area have a size of 300 lines, 300 rows, a total of 114 bands, a resolution of 3.5 m and UTM-WGS84 projection as shown in Fig. 2.

Data processing: Hymap data of the research area have 114 bands. Reduction of the data redundancy requires the data band to be selected, which is generally consists of two aspects; band information and band correlation. Band information content and standard deviation were positively correlated and the greater the standard deviation, the greater the degree of band dispersion, the more abundant the information. The absolute value of the correlation coefficient between the bands is smaller, which indicates that the independence is stronger and

Table 1: Band information statistical results for the study area

Channel groups	Bands	Wavelength value	Minimum value	Maximum value	Mean value	Standard deviation
Blue channel group	1	0.4520	22	9181	780.46	741.94
	2	0.4644	23	9176	823.58	776.09
	3	0.4804	23	9166	816.80	779.22
	4	0.4952	24	9154	862.91	810.22
	5	0.5101	44	9152	900.50	822.76
	6	0.5259	45	9166	907.90	814.69
Green channel group	7	0.5415	45	9173	957.48	835.50
	8	0.5568	45	9172	1001.17	859.28
	9	0.5719	44	9165	1024.80	886.33
	10	0.5871	23	9161	1028.74	900.24
	11	0.6025	23	9158	1055.10	916.02
Red channel group	12	0.6180	23	9154	1059.17	923.67
	13	0.6335	23	9152	1072.46	936.58
	14	0.6487	23	9146	1081.54	952.96
	15	0.6638	23	9139	1084.17	961.34
	16	0.6790	23	9137	1099.34	968.06
	17	0.6944	25	9180	1236.54	918.40
	18	0.7098	0	9219	1373.51	937.71
	19	0.7249	0	9254	1507.80	1019.86
	20	0.7400	0	9286	1642.45	1150.74
	21	0.7552	0	9292	1700.06	1207.21
Near infrared channel group	22	0.7703	0	9298	1757.28	1268.11
	23	0.7852	0	9304	1813.73	1332.27
	24	0.8003	0	9308	1820.70	1343.63
	25	0.8156	0	9311	1827.79	1355.40
	26	0.8307	0	9314	1834.82	1367.25
	27	0.8460	0	9318	1841.96	1379.46
	28	0.8613	0	9319	1860.68	1396.53
	29	0.8765	0	9320	1887.34	1418.44
	30	0.8886	0	9324	1908.95	1438.86
	31	0.8948	0	9313	1905.33	1426.75
	32	0.9112	0	9312	1918.69	1432.61



Fig. 2: False color image of Hymap (R: Band 30, G: Band 16, Band 11)

the information redundancy is smaller²⁵. Therefore, the smaller the correlation coefficient, the larger the standard deviation bands can be retained. The first 32 bands were

selected in this study, which are divided into four groups: The red, green, blue and near infrared channels, such that the data could be balanced. The standard deviations of the bands are listed in Table 1.

The results in Table 1 show that the standard deviation of bands 4, 5, 6, 9, 10, 11, 14, 15, 16, 30, 31 and 32 is larger, with the maximum standard deviation of the bands being band 30 and the minimum being band 1. The bands that were selected to carry out the correlation coefficient statistics are listed in Table 2.

The values in Table 2 show that the correlation coefficient of B5 is the largest information band in the blue channel group and B11, which belongs to the green channel group, is small. The correlation coefficient of B5 and B16, which belong to the red channel group, is small, whereas B11 and B16 are both the largest information bands in their respective groups. The correlation coefficient of B5 and B32, which belong to the near-infrared channel group is also small, but B30 is the largest information band in the near-infrared channel group. Based on these results, we selected data from bands 30, 16, 11 and 5 to participate in image classification.

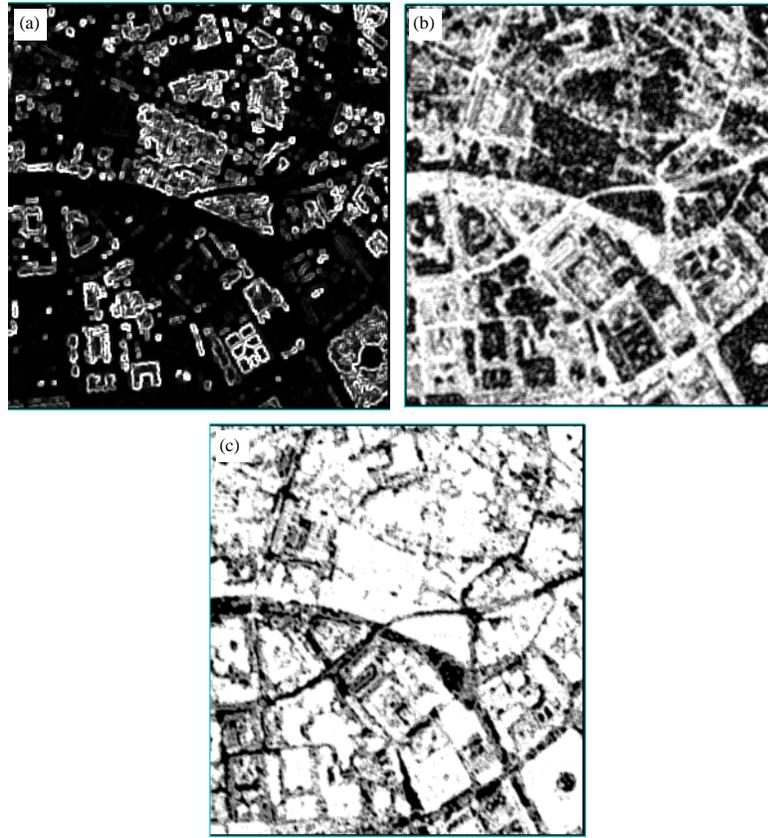


Fig. 3(a-c): Main texture image map of the study are (a) Variance, (b) Homogeneity and (c) Entropy

Table 2: Correlation coefficient matrix of each band in the study area

Correlation	4	5	6	9	10	11	14	15	16	30	31	32
4	1.000											
5	0.999	1.000										
6	0.995	0.998	1.000									
9	0.982	0.987	0.992	1.000								
10	0.975	0.980	0.984	0.997	1.000							
11	0.965	0.970	0.973	0.992	0.998	1.000						
14	0.945	0.950	0.952	0.977	0.989	0.996	1.000					
15	0.940	0.945	0.946	0.972	0.986	0.994	0.999	1.000				
16	0.934	0.939	0.941	0.969	0.983	0.992	0.998	0.999	1.000			
30	0.342	0.363	0.415	0.448	0.429	0.422	0.398	0.389	0.390	1.000		
31	0.336	0.357	0.408	0.442	0.423	0.417	0.393	0.383	0.385	0.999	1.000	
32	0.337	0.358	0.410	0.444	0.425	0.419	0.396	0.387	0.388	0.998	0.999	1.000

The Normalized Differential Vegetation Index (NDVI) and texture information of the image in the study area was estimated by choosing the band. Firstly, principal component analysis was carried out on the image before calculating texture information. The first principal component, which contained most of the information of the image, was used to calculate the mean, variance, contrast, homogeneity, dissimilarity, entropy second moment and correlation, a total of eight kinds of texture information. Then, choose the

variance, homogeneity, entropy and correlation, i.e., four kinds of texture band, to participate in image classification (Fig. 3).

Classification and accuracy assessment: The study area was divided into five categories: Vegetation, construction, impervious layer, soil and water. According to the distribution characteristics of the image features combined with the real classification value of the study area. Visual interpretation was used by combining images and to select samples of all of

these types. In this way we extracted the B5, B11, B16, B30, NDVI, variance, homogeneity, entropy and correlation of nine attributes ($M = 9$), including 2000 training samples and 500 testing samples.

Optimal model parameter selection: In this study, image classification was performed by using the RF classification algorithm and SVM to classify the images based on the respective training samples. The random forest program package based on MATLAB platform is RF_MexStandalone-v0.02.

Model formula: $\text{Model} = \text{classRF_train}(X, Y, \text{ntree}, \text{mtry})$, where X is the attribute matrix, Y is the class label, ntree is the number of trees and mtry is the number of attributes of random selection that is m .

As the accuracy of the random forest depends on the number of trees N and mtry , in order to obtain the best model parameters, the following selections of combinations of several parameters were tested, through the comparison of the OOB error rate (smaller the error rate, the higher the accuracy rate) to select the best values of N and m . Several groups of parameters were selected for $N = 500$ and m 2, 3 and 4, respectively, as shown in Fig. 4.

Figure 4 shows that when the value of m is 4, the OOB error rate is the smallest, which means that the correct rate is the best, the number of trees is about 450, the OOB error rate tends to be stable and the model has good effect. Thus, the random forest model parameter N is 450 and m is 4.

The support vector machine was based on the Faruto Ultimate 3.0 version of the MATLAB platform. Model

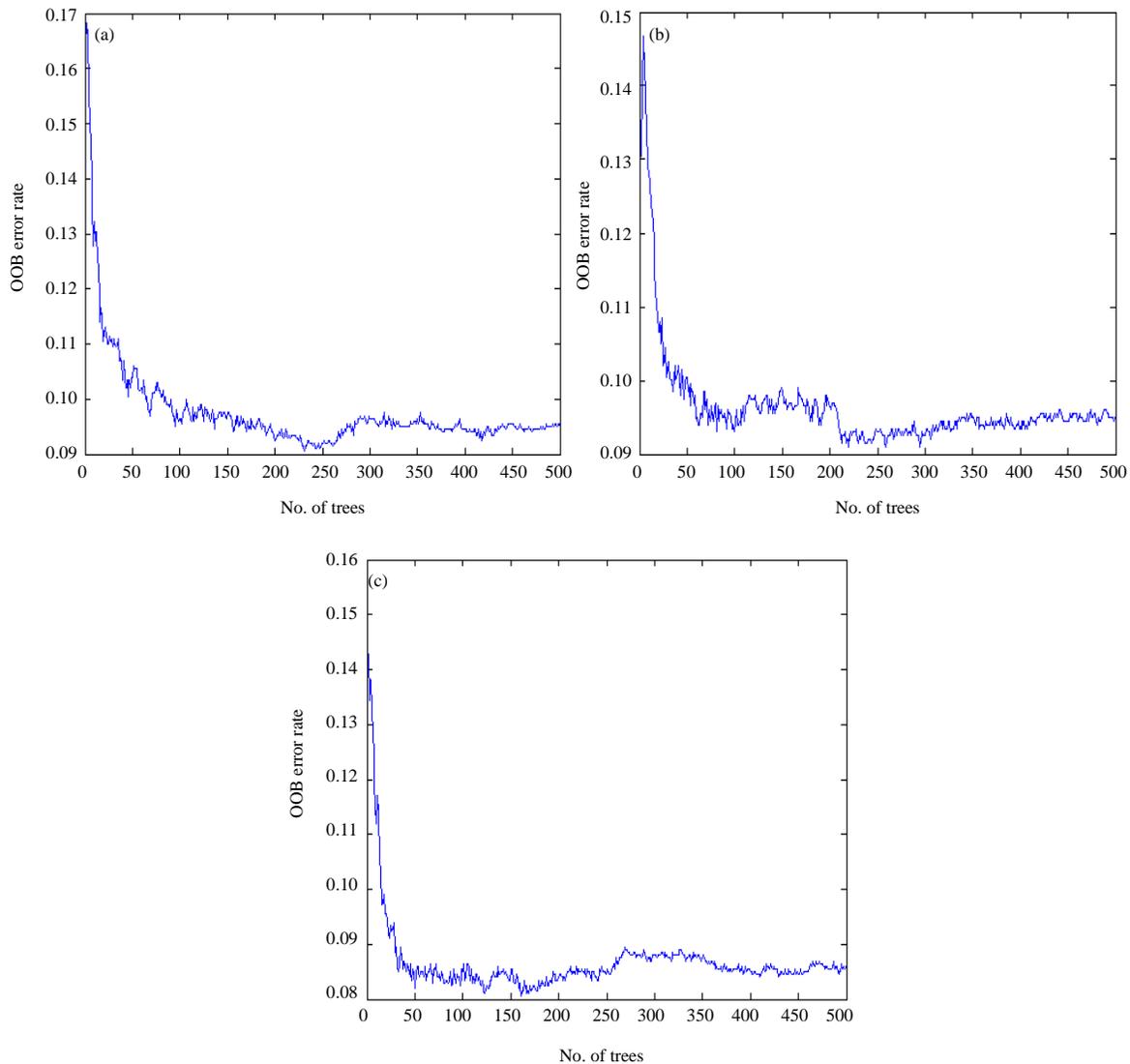


Fig. 4(a-c): Random split variables (a) $m = 2$, (b) $m = 3$ and (c) $m = 4$ on OOB error rate

formula: Model = SVM train (Y, X'-s, -t, -c, -g'), where X is the attribute matrix vector, Y is a class label, -s is the SVM type with a default of 0, -t is the type of kernel function, with the default option value of 2, i.e., an RBF kernel function and -c and -g represent the penalty parameter and gamma (γ) value, respectively. Using the heuristic genetic algorithm function gaSVMcgForClass to carry out parameter optimization, the optimal gamma (γ) value and the penalty parameter were determined to be 50 and 46, respectively.

Classification and accuracy verification: Based on the 2000 training samples, the RF and SVM methods were used to classify the images and the classification results were obtained. The D-map is the RF classification result, the E-map is the SVM classification result and the F-map is the ground truth classification map (Fig. 5).

Figure 5 shows that the different algorithm classification results are very close. Although, the main features can be classified, there are some differences in classification, confusion and even error classification between some categories. The comparison of D and F in Fig. 5 showed that range of soil classification of the D-map is greater than that of F and some buildings are divided into soil categories. There is also some confusion between the construction category and the impervious layer. In contrast, vegetation and water body classification are good. Comparison between E and F showed range of soil classification of the E-map is greater than that of F and is also greater than the range of soil classification of the D-map. There is also some confusion between the soil category and the impervious layer. Misclassification phenomena appear between the impervious layer and the water body, although vegetation classification is more accurate.

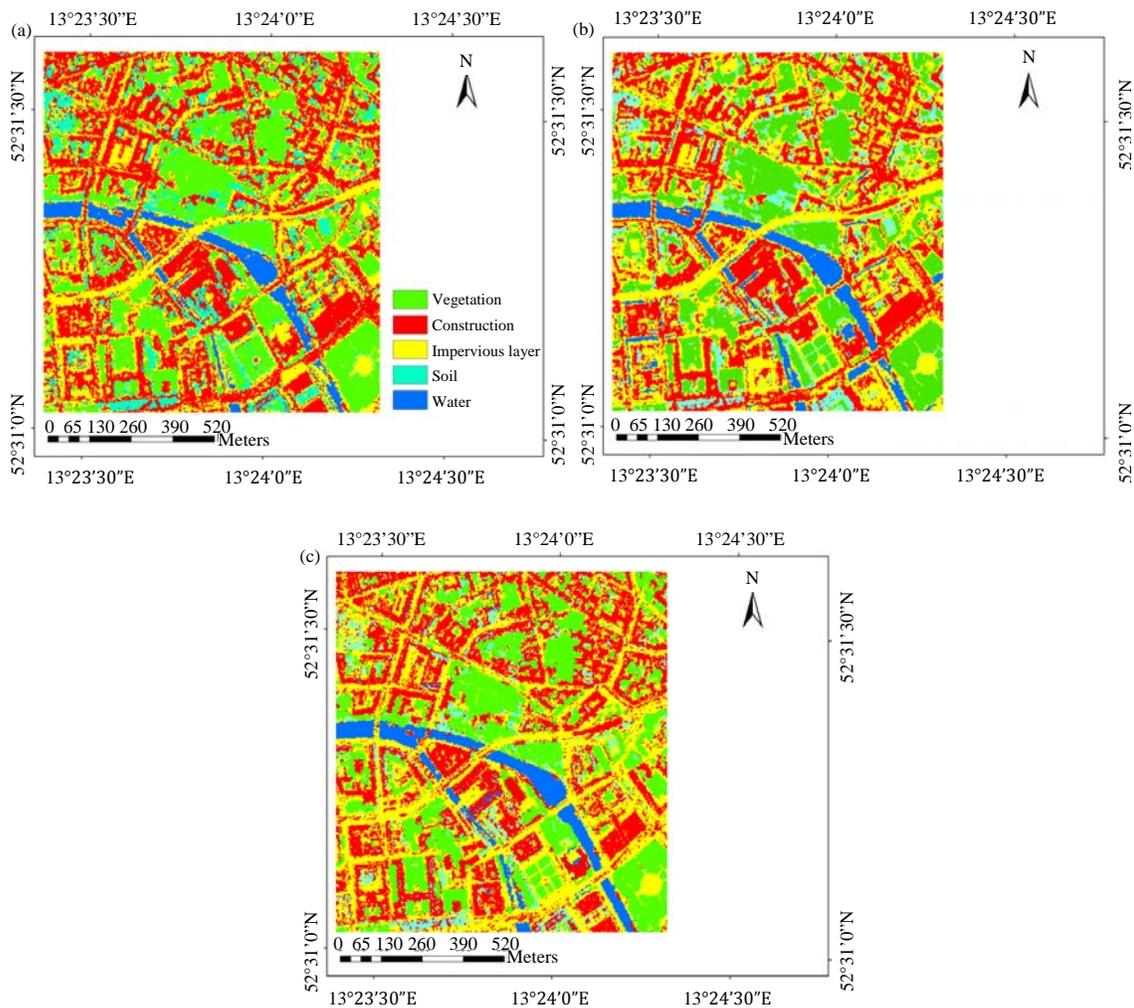


Fig. 5(a-c): (a) RF, (b) SVM and (c) Ground-truth classification map

The main factors influencing the classification accuracy evaluation is the confusion matrix, the overall classification accuracy, the Kappa coefficient and so on. The confusion matrix contains the true surface pixel classification results, with which the image classification results are compared. The total number of each column is the prediction for the number of category A, the value in each column indicates that the real data is predicted for the number of category A. The overall classification accuracy is the ratio of the correct number of pixels to the total number of pixels. The mapping accuracy is the ratio of the correct number of category A pixels (diagonal numerical) to the number of real category A pixels. The user accuracy is the ratio of the correct number of category A pixels (diagonal numerical) to the number of pixels in category A as classified by the classifier.

The ground truth classification image sample points were selected by using a total of 500 test samples to test the RF and SVM classification results, respectively, to create the confusion matrix (Table 3 and 4).

The results in Table 3 show that the mapping precision of vegetation, soil and water is high, whereas that of the impervious layer is minimal. The user accuracy for vegetation and water is high and for the soil category it is low. The results in Table 4 show that the mapping precision of vegetation, soil and water is high, but that the user accuracy for the soil category is low. A comparison of the two tables shows that for each category the mapping accuracy of the RF algorithm is higher than that of the SVM algorithm. The user accuracy of the RF algorithm is higher than that of the SVM algorithm for all categories, except for the soil category.

The overall classification accuracies and Kappa coefficients of RF and SVM were calculated by using the confusion matrices of Table 3 and 4.

The results in Table 5 show that the classification results of the two algorithms are both acceptable. The overall classification accuracy and Kappa coefficient of RF are higher than those of SVM. For the RF algorithm the overall classification accuracy reaches 92.6%, the Kappa coefficient is 0.9024, the SVM algorithm the overall classification accuracy reaches 91.2%, the Kappa coefficient is 0.8840.

DISCUSSION

This study based on hysmap high resolution image using RF, SVM two methods to study the classification of Berlin in Germany, classification map (Fig. 5) shows that between the support vector machine classifier different classification categories, the probability of the classification of the error and confusion is greater than random forest. The reason may be due to the multi classifier combination voting mechanism of random forest plays an important role. single classifier classification is easily leads to fitting problems, the combination of multiple classifiers for collective decision-making can have a better resist noise capability and more accurate than a single classifier and the support vector machine does not have this mechanism, so RF can deal with classification problems of unbalanced and multi-class.

The accuracy of the classification results of the two algorithms show that the overall accuracy and kappa coefficient of the RF algorithm is improved by 1.4% and 0.018 compared with SVM. Researchers have studied the random forest algorithm and SVM algorithm. The application of remote sensing classification, but most of the research is based on the low and medium resolution image. Pal²⁶ used Landsat-7 Enhanced Thematic Mapper (ETM+) data (19 June, 2000) of an agricultural area near Littleport, Cambridge shire, UK to divide

Table 3: Accuracy assessment of RF classification

Category	Vegetation	Construction	Impervious layer	Soil	Water	Total	User accuracy (%)
Vegetation	137	0	0	0	0	137	100.00
Construction	0	145	12	1	0	158	91.77
Impervious layer	0	11	92	2	0	105	87.62
Soil	1	4	5	37	0	47	78.72
Water	0	0	1	0	52	53	98.11
Total	138	160	110	40	52	500	
Mapping accuracy (%)	99.28	90.63	83.64	92.50	100.00		

Table 4: Accuracy assessment on SVM classification

Category	Vegetation	Construction	Impervious layer	Soil	Water	Total	User accuracy (%)
Vegetation	137	0	0	0	0	137	100.00
Construction	0	140	16	1	0	157	89.17
Impervious layer	0	13	91	2	1	107	85.05
Soil	1	7	1	37	0	46	80.43
Water	0	0	2	0	51	53	96.23
Total	138	160	110	40	52	500	
Mapping accuracy (%)	99.28	87.50	82.73	92.50	98.08		

Table 5: Overall accuracy and Kappa coefficient comparison of RF and SVM

Classification method	RF	SVM
Overall accuracy (%)	92.600	91.200
Kappa	0.9024	0.8840

into 7 categories (wheat, potato, sugar beet, onion, peas, lettuce and beans) based on the two methods of SVM and RF. The RF algorithm the overall classification accuracy reaches 88.37%, the Kappa coefficient is 0.864, the SVM algorithm the overall classification accuracy reaches 87.9%, the Kappa coefficient is 0.86. Du *et al.*²⁷ used SVM and RF algorithms to divide the AVIRIS image into nine categories (corn-notill, corn-min, grass-pasture, grass-trees, hay-windrowed, soybeans-notill, soybeans-min, soybeans-clean, woods). Results shows that the overall classification accuracy, Kappa coefficients of RF and SVM algorithms are 95.1%, 94.3% and 0.94, 0.93. Li *et al.*²⁸ tested two unsupervised and 13 supervised classification algorithms including SVM and RF algorithms with the Landsat Thematic Mapper data set over Guangzhou City, China. The final results show that the best classification accuracy for RF algorithm using objected-oriented approach is 0.917, the SVM is 0.891. Qin *et al.*²⁹ used the Maximum Likelihood Classification (MLC), Support Vector Machines (SVM) and random forest algorithms for Zhangye city land cover classification based on Lidar data. The results suggest that RF classifier has the most accurate result with overall classification accuracy of 91.82% and the kappa coefficient of 0.88, SVM is 88.48% and 0.83.

Through the above comparison, the authors speculate that with the improvement of the resolution of the image data level, RF and SVM classification accuracy is also gradually improve, in the same resolution imaging conditions, RF classification accuracy is substantially better than the SVM classifier. Of course there will be other factors have an impact on the accuracy of classification, such as reasonable parameter selection.

From the two algorithm model, the parameter optimization selection can show that the number of parameters of random forest algorithm is small (only two), the optimization process is simple and the number of parameters of SVM are relatively more, the optimization difficulty is higher than the random forest.

CONCLUSION

The study presented in this study aimed to use the random forest algorithm to classify high spectral and high-resolution images. The results show that the classification results of the two methods (RF and SVM) are both good, but that RF has a higher classification accuracy and faster

classification speed with improved stability compared to SVM. The study of high resolution image RF classification method and its application, can quickly respond to the complex classification of the situation. In addition, the RF method is easier to use in production practice. It is a very good method and can be widely used in various fields.

In terms of a future research direction, the hmap image contains a large number of bands, but this study only used the information contained in the first 32 bands and the useful information in the other bands (such as mid-infrared) is not used for the classification of useful information. Band selection is simply based on calculation of the Standard Deviation (SD) and correlation coefficient and some of the methods are not used for high spectral band selection, such as the band index calculation. More attribute factors can be introduced to participate in classification such as the normalized difference water index and the normalized difference built-up index. Combining the random forest method and SVM method for remote sensing classification, in order to comprehensively use the advantages of the two methods, may be is a good strategy, these will be evaluated in a future study.

ACKNOWLEDGMENTS

Financial support was provided by Commission of Science, Technology and Industry for National Defence of China (07-Y30A05-9001-12/13).

REFERENCES

1. Jia, K., Q.Z. Li, Y.C. Tian and B.F. Wu, 2011. A review of classification methods of remote sensing imagery. *Spectrosc. Spectral Anal.*, 31: 2618-2623.
2. Hayes, M.M., S.N. Miller and M.A. Murphy, 2014. High-resolution landcover classification using random forest. *Remote Sens. Lett.*, 5: 112-121.
3. Elith, J., J.R. Leathwick and T. Hastie, 2008. A working guide to boosted regression trees. *J. Anim. Ecol.*, 77: 802-813.
4. Lu, D. and Q. Weng, 2007. A survey of image classification methods and techniques for improving classification performance. *Int. J. Remote Sens.*, 28: 823-870.
5. Breiman, L., 2001. Random forests. *Mach. Learn.*, 45: 5-32.
6. Huang, Y. and W.X. Zha, 2012. Comparison on classification performance between random forests and support vector machine. *Software*, 33: 107-110.
7. Dietterich, T.G., 2000. An experimental comparison of three methods for constructing ensembles of decision trees: Bagging, boosting and randomization. *Mach. Learn.*, 40: 139-157.

8. Breiman, L., 1996. Bagging predictors. *Mach. Learn.*, 24: 123-140.
9. Ho, T.K., 1998. The random subspace method for constructing decision forests. *IEEE Trans. Pattern Anal. Mach. Intell.*, 20: 832-844.
10. Liu, Y., P.J. Du, H. Zheng, J.S. Xia and S.C. Liu, 2012. Classification of China small satellite remote sensing image based on random forests. *Sci. Surv. Mapping*, 37: 194-196.
11. Fang, K.N., J.B. Wu, J.P. Zhu and B.C. Shia, 2011. A review of technologies on random forests. *Stat. Inform. Forum*, 26: 32-38.
12. Shishi, D. and H. Zhexue, 2013. A brief theoretical overview of random forests. *J. Integrat. Technol.*, 2: 1-7.
13. Verikas, A., A. Gelzinis and M. Bacauskiene, 2011. Mining data with random forests: A survey and results of new tests. *Pattern Recognit.*, 44: 330-349.
14. Witten, I.H., E. Frank and M.A. Hall, 2011. *Data Mining: Practical Machine Learning Tools and Techniques*. 3rd Edn., Morgan Kaufmann, USA., ISBN-13: 978-0123748560, Pages: 664.
15. Li, X.H., 2013. Using random forest for classification and regression. *Chin. J. Applied Entomol.*, 50: 1190-1197.
16. Griffin, S., J. Rogan and D.M. Runfola, 2011. Application of spectral and environmental variables to map the Kissimmee prairie ecosystem using classification trees. *GISci. Remote Sens.*, 48: 299-323.
17. Wang, D., C.R. Yue, C.Z. Tian, H.G. Fan and Y.H. Wang, 2014. Classification of TM remote sensing image based on random forests of Dayao county. *For. Inventory Plan.*, 2: 1-5.
18. Cortes, C. and V. Vapnik, 1995. Support-vector networks. *Mach. Learn.*, 20: 273-297.
19. Ding, S.F., B.J. Qi and H.Y. Tan, 2011. An overview on theory and algorithm of support vector machines. *J. Univ. Electron. Sci. Technol. China*, 40: 2-10.
20. Zhang, C., S.Y. Zang, Z. Jin and Y.H. Zhang, 2011. Remote sensing classification for Zhalong wetlands based on support vector machine. *Wetland Sci.*, 9: 263-269.
21. Melgani, F. and L. Bruzzone, 2004. Classification of hyperspectral remote sensing images with support vector machines. *IEEE Trans. Geosci. Remote Sens.*, 42: 1778-1790.
22. Zhang, R. and J. Ma, 2009. State of the art on remotely sensed data classification based on support vector machines. *Adv. Earth Sci.*, 24: 555-562.
23. Gou, B. and X. Huang, 2006. SVM multi-class classification. *J. Data Acquisit. Process.*, 21: 334-339.
24. Chang, C.C. and C.J. Lin, 2011. LIBSVM: A library for support vector machines. *ACM Trans. Intell. Syst. Technol.*, 2: 1-39.
25. Chen, G.H., L.L. Tang and X.G. Jiang, 2006. Feature selection and extraction of hyperspectral data-Based on HyMap data of Barrax in Spanish. *Arid Land Geogr.*, 29: 143-149.
26. Pal, M., 2005. Random forest classifier for remote sensing classification. *Int. J. Remote Sens.*, 26: 217-222.
27. Du, P., J. Xia, J. Chanussot and X. He, 2012. Hyperspectral remote sensing image classification based on the integration of support vector machine and random forest. *Proceedings of the IEEE International Geoscience and Remote Sensing Symposium*, July 22-27, 2012, Munich, Germany, pp: 174-177.
28. Li, C., J. Wang, L. Wang, L. Hu and P. Gong, 2014. Comparison of classification algorithms and training sample sizes in urban land classification with Landsat thematic mapper imagery. *Remote Sens.*, 6: 964-983.
29. Qin, Y., S. Li, T.T. Vu, Z. Niu and Y. Ban, 2015. Synergistic application of geometric and radiometric features of LiDAR data for urban land cover mapping. *Optics Express*, 23: 13761-13775.