ISSN 2616-5775 Vol. 4, Issue 6: 67-71, DOI: 10.25236/AJCIS.2021.040611

Hyperspectral remote sensing image classification based on random forest

Luo Zhikun[1], Tan Kaiyao[2]

[1]Hunan University of Science and Technology, School of Resource & Environment and Safety Engineering, Xiangtan 411201, China
[2]Guizhou University, School of Computer Science and Technology, Guizhou 550025, China

Abstract: *With the development of remote sensing technology and machine learning, the research on hyperspectral remote sensing image classification has also progressed rapidly. In this paper, based on the random forest model, a new classification model of hyperspectral remote sensing images is proposed, which can effectively classify the spectral information and texture information of hyperspectral images. First, the texture features are extracted from the hyperspectral remote sensing image data and superimposed on the original spectral domain data, thus forming a new spectral-space domain data. Then, the model forms an optimal combination of parameters by selecting parameters such as the number of decision trees, maximum depth and minimum number of leaves, and this model has a higher classification accuracy for the dataset than the original random forest model. Experimental results on Indian, KSC and Salinas hyperspectral images show that the method proposed in this paper can solve the problems of hyperspectral data nonlinearity and information redundancy due to the expansion of data and the selection of parameter combinations, thus improving the classification effect of the original random forest and the average accuracy of the classification.*

Keywords: *hyperspectral remote sensing, texture features, spectral-space domain, random forest model, optimal combination of parameters*

1. Introduction

Remote sensing is an important earth observation technology. Remote sensing data is the two-dimensional image data produced in the process of remote sensing technology. By acquiring two-dimensional geometric space information and one-dimensional spectral information of the target area, the hyperspectral data with three-dimensional "image cube" can be formed. This quality of hyperspectral remote sensing technology has made it excellent for applications in the fields of geology and minerals [2], precision agriculture [3], atmospheric environment [4].

Xu [6] and Guo [6] show that hyperspectral remote sensing leads to the nonlinear characteristics of hyperspectral data structure due to the high-dimensional characteristics of the signal, uncertainty, and information redundancy, making it difficult for some classification models to directly classify the raw hyperspectral data [5][6]. And the standard classifier is based on one classification method decision, which classification accuracy is poor when facing the hyperspectral data. Wang [7] shows that random forest, as a multi-classifier system, can be well applied to highly redundant datasets, for example, the overall classification accuracy on the Indian dataset can reach 79.8%. However, there is still much room for improving its classification accuracy. Therefore, a large number of related studies have been done by numerous scholars in order to improve the image element classification accuracy [8][9][10].

In this paper, a BRF method is introduced, which, on the one hand, improves the original random forest model RF from the perspective of adjusting parameter combinations, and on the other hand, this paper also enhances the classification effect by superimposing space domain data with texture features on the spectral domain data to obtain the spectral-space domain data. Finally, several classical hyperspectral datasets: Indian, KSC and Salinas are used in this paper to compare the classification effects using the average accuracy and Kappa coefficients.

2. Random forest model

A random forest is an integrated learning method consisting of several decision trees. A random forest

ISSN 2616-5775 Vol. 4, Issue 6: 67-71, DOI: 10.25236/AJCIS.2021.040611

classifier consists of a series of decision tree classifiers, where each decision tree classifier can be viewed as a random vector that exists independently and each decision tree classifier has a vote for the class to which the input samples belong. Moreover, each decision tree requires only a fraction of the original training samples for training. This approach can solve the performance bottleneck problem of decision trees mentioned in the previous paragraph, and it also has good parallelism and scalability for high-dimensional data classification problems, with good tolerance to noise and outliers.

3. BRF based on random forest model

In the original random forest model, the number of decision trees is set to 100 by default, the maximum number of features in the decision tree is the square root of *N*, the maximum depth of the decision tree is not limited, and the minimum number of samples contained in the leaf nodes is 1 by default. However, the importance of the features in each dataset varies from dataset to dataset due to the differences in dimensionality and data volume, etc. Therefore, choosing the right combination of model parameters is the problem that the model needs to be improved.

Figure 1 shows the learning curve of the number of trees. We can see that the accuracy reaches the highest when the number of trees is around 200, and then it starts to fluctuate, so the most suitable number of decision trees for the Indian dataset is around 200. After subsequent further optimization search, the amount of the optimal number of decision trees is finally found to be 217. In addition to the number of decision trees, there are also aspects such as the maximum depth of the tree, the minimum leaf tree of the tree, etc. that affect the accuracy and model complexity of the model. We need to ensure that the model complexity is not high, but improve the accuracy of the model so that the generalization error of the model is as small as possible.

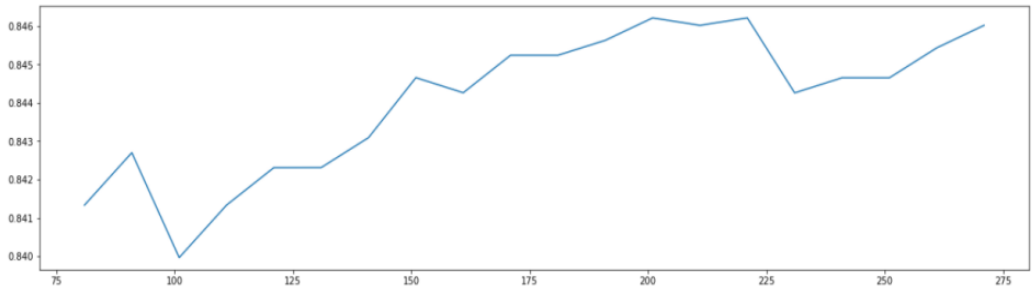


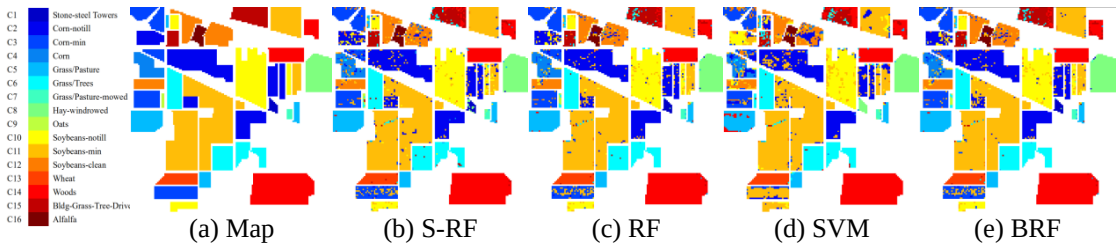
Figure 1: Learning curve of Indian dataset

4. Experiment

In order to verify the effectiveness of the improved random forest model in the previous section, three hyperspectral datasets are selected for experiments in this section, namely, the Indian dataset, the KSC dataset and the Salinas dataset. The classification evaluation metrics used are macro average accuracy (*MA*), weighted average accuracy (*WA*) and Kappa coefficient (*k*).

4.1. Analysis of feature recognition

4.1.1. Indian dataset



(a) Map (b) S-RF (c) RF (d) SVM (e) BRF

Figure 2: Distribution of typical features in Indian

The Indian data are hyperspectral remote sensing images of the Indian Pines test site in Indiana, USA. After the water vapor noise bands were removed, the effective bands were 200. The original images are

ISSN 2616-5775 Vol. 4, Issue 6: 67-71, DOI: 10.25236/AJCIS.2021.040611

labeled with different colors according to feature categories, as shown in Figure 2(a). The image data contains 16 categories of features.

Figure 2(b,c,d,e) shows the hyperspectral image classification results, S-RF represents the spectral domain, and the rest of the methods are based on spectral-space domain data.We can see that the classification results of hyperspectral images based on the S-RF model are poor.

In Table 1, it is easy to see that the classification results using spectral-space features are better than those using single features.Moreover, the classification accuracy of BRF model proposed is higher than the results of SVM, with a 15% improvement in *MA*, 7.6% improvement in *WA* and 9.7% improvement in *k* compared with the SVM method.

Table 1: Comparison of classification results of hyperspectral images

Indian					KSC					Salinas				
Category	S-RF	RF	SVM	BRF	Category	S-RF	RF	SVM	BRF	Category	S-RF	RF	SVM	BRF
C1	1.0000	0.9980	1.0000	1.0000	C1	0.9367	0.9538	0.9503	0.9385	C1	1.0000	0.9980	1.0000	1.0000
C2	0.9989	0.9995	0.9942	1.0000	C2	0.8850	0.8761	0.9744	0.9145	C2	0.9989	0.9995	0.9942	1.0000
C3	0.9872	0.9839	0.9521	0.9838	C3	0.9113	0.9392	0.8627	0.9180	C3	0.9872	0.9839	0.9521	0.9838
C4	0.9923	0.9874	0.9928	0.9940	C4	0.6879	0.8162	0.7358	0.8521	C4	0.9923	0.9874	0.9928	0.9940
C5	0.9949	0.9902	0.9870	0.9935	C5	0.7778	0.8841	0.9811	0.9118	C5	0.9949	0.9902	0.9870	0.9935
C6	0.9995	0.9995	1.0000	0.9995	C6	0.7579	0.7368	0.8810	0.8273	C6	0.9995	0.9995	1.0000	0.9995
C7	1.0000	1.0000	0.9983	0.9989	C7	0.7843	0.8571	0.8983	0.7869	C7	1.0000	1.0000	0.9983	0.9989
C8	0.8461	0.8558	0.7556	0.8614	C8	0.8375	0.9104	0.8978	0.9248	C8	0.8461	0.8558	0.7556	0.8614
C9	0.9968	0.9971	0.9959	0.9914	C9	0.9341	0.8630	0.9176	0.9192	C9	0.9968	0.9971	0.9959	0.9914
C10	0.9678	0.9723	0.9668	0.9739	C10	0.9808	0.9950	0.9843	0.9851	C10	0.9678	0.9723	0.9668	0.9739
C11	0.9815	0.9749	0.9756	0.9642	C11	0.9948	0.9951	0.9712	0.9951	C11	0.9815	0.9749	0.9756	0.9642
C12	0.9842	0.9842	0.9687	0.9816	C12	0.9708	0.9538	0.9240	0.9837	C12	0.9842	0.9842	0.9687	0.9816
C13	0.9717	0.9777	0.9643	0.9773	C13	1.0000	1.0000	1.0000	1.0000	C13	0.9717	0.9777	0.9643	0.9773
C14	0.9868	0.9773	0.9843	0.9961	*MA*	0.8815	0.9062	0.9214	**0.9198**	C14	0.9868	0.9773	0.9843	0.9961
C15	0.8396	0.8499	0.8374	0.8538	*WA*	0.9204	0.9319	0.9358	**0.9433**	C15	0.8396	0.8499	0.8374	0.8538
C16	0.9978	0.9909	0.9921	0.9967	*K*	0.9104	0.9231	0.9249	**0.9364**	C16	0.9978	0.9909	0.9921	0.9967
MA	0.9716	0.9712	0.9603	**0.9729**						*MA*	0.9716	0.9712	0.9603	**0.9729**
WA	0.9411	0.9442	0.9190	**0.9457**						*WA*	0.9411	0.9442	0.9190	**0.9457**
K	0.9345	0.9379	0.9049	**0.9396**						*K*	0.9345	0.9379	0.9049	**0.9396**

4.1.2. KSC dataset

The KSC data were taken by the AVIRIS sensor at the Kennedy Space Center in Florida .After water vapor noise band removal, the effective band is 176.There are 13 categories representing the various land cover types occurring in this environment, as shown in Figure 3(a).

(a) Map (b) S-RF (c) RF (d) SVM (e) BRF

Figure 3: Distribution of typical features in KSC

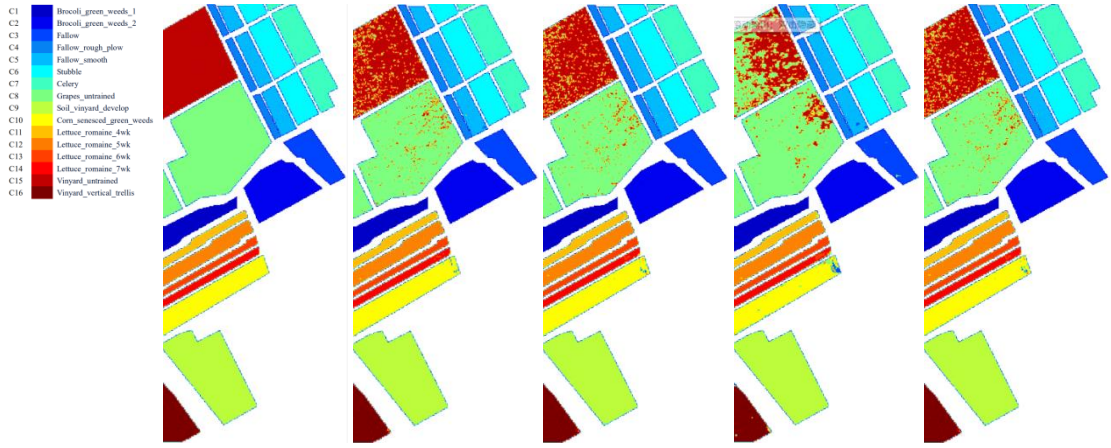
The classification results of different methods are given in Figure 3(b,c,d,e). The introduction of space information improves the classification accuracy and achieves a complement to the spectral information. The classification results of class 4 and class 5 using combined features are significantly better than those obtained by using single spectral features.

As seen in Table 1, the model in this paper outperforms the classification based on the spectral feature approach in terms of classification accuracy and kappa coefficient, with a 2.8% improvement in macro-average accuracy, 1.2% improvement in weighted average accuracy and 1.4% improvement in kappa coefficient. This indicates that the inclusion of texture features leads to a certain degree of improvement in classification accuracy.

4.1.3. Salinas dataset

Salinas was captured by the AVIRIS sensor in Salinas Valley, CA. After removing the bands with heavy water vapor absorption, 204 bands remain. The original image is labeled with different colors according to feature categories, as shown in Figure 4(a), which contains 16 crop categories.

ISSN 2616-5775 Vol. 4, Issue 6: 67-71, DOI: 10.25236/AJCIS.2021.040611



(a) Map (b) S -RF (c) RF (d) SVM (e) BRF

Figure 4: Distribution of typical features in Salinas

The classification results of different methods are given in Figure 4(b,c,d,e). It can be seen that class 8 and class 15 are relatively similar in spectral features, so they will easily misclassify each other when classified, so among the 16 classes of crop classification, only class 8 and class 15 are significantly lower in classification accuracy than other crops. After the introduction of texture features, the classification accuracy was improved.

As seen in Table 1, the introduction of texture features helps to distinguish more accurately between class 8 and class 15 crops, but it also leads to a decrease in accuracy due to misclassification of a very small portion of the data, such as class 3 and class 5. However, *MA*, *WA* and *k* all improved by 0.1%, 0.5% and 0.5%, respectively, which is a significant improvement in the overall classification accuracy on the surface.

4.2. Analysis of texture feature recognition

A texture primitive is considered to be a collection of several pixel points that, when combined together, can form a local image structure that recurs throughout the entire image, with the smallest texture primitive being a single pixel point.

Figure 5 shows the texture features of the Indian image, where Original is the original image, Mean represents the mean value; Std represents the standard deviation; Contrast represents the contrast, reflecting the sharpness of the image and the depth of the texture grooves; Dissimilarity represents the phase dissimilarity; Homogeneity represents the inverse disparity, measuring the local uniformity of the image Asm stands for angular second-order moment, which is used to describe the uniformity of the image grayscale distribution and the coarseness of the texture; Energy stands for energy, which is the squared sum of the values of the elements of the grayscale co-occurrence matrix, reflecting the uniformity of the image grayscale distribution and the coarseness of the texture; Entropy stands for entropy, which measures the randomness of the image texture; Max stands for maximum probability. Gray is the grayscale image, and Lbp is the LBP feature image obtained from the grayscale image operation.

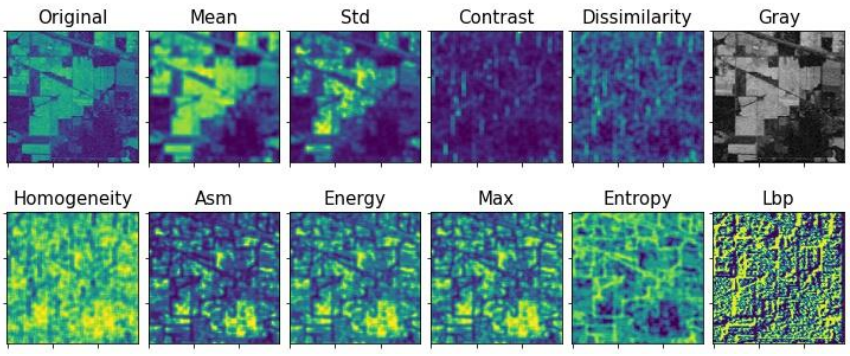


Figure 5: Texture features of Indian images

ISSN 2616-5775 Vol. 4, Issue 6: 67-71, DOI: 10.25236/AJCIS.2021.040611

5. Conclusion

The classification accuracy of hyperspectral images is affected by two factors: the processing of spectral data and the selection of classification models. For the processing of spectral data, this paper proposes that the classification based on the superposition of spectral and space information, which can improve the classification accuracy; and for the selection of classification model, this paper determines the optimal combination of parameters by choosing the optimal number of decision trees, maximum depth, minimum number of leaves and other parameters on the original better-performing RF model to obtain the BRF model, which further improves the classification effect. Finally, this paper uses the Indian, KSC and Salinas datasets to evaluate the proposed method with three metrics: *MA*, *WA* and *k*, and the results prove that the classification effect of BRF is optimal. However, the research work in this paper has many areas for improvement, such as how to improve the space domain data for better classification of spectral-space domain data, and how to choose other classification models, such as KNN, CNN, etc., to obtain better classification results.

References

[1] Tong Q X, Zhang B and Zheng L F. Hyperspectral Remote Sensing: the Principle, Technology and Application [M]. Beijing: Higher Education Press, 2006a.
[2] Kruse F A, Boardman J W and Huntington J F. Comparison of airborne hyperspectral data and EO-1 Hyperion for mineral mapping [J]. IEEE Transactions on Geoscience and Remote Sensing, 2003, 41(6): 1388-1400
[3] Yao Y J, Qin Q M, Zhang Z L, et al. Research progress of hyperspectral technology applied in agricultural remote sensing [J]. Transactions of the CSAE, 2008, 24(7): 301-306.
[4] Jiang D M. Approach to High Spectral Resolution Infrared Remote Sensing of Atomospheric Temperature and Humidity Profiles [D]. Nanjing University of Information Science and Technology, 2007.
[5] Xu J J, Zhao H. Research on Feature Fxtraction and Flassification of Fyperspectral Femote Fensing Image-based on Discrete Cosine Transform (DCT) and Support Vector Machine Technology [J]. Journal of Jiamusi University (Natural Science Edition), 2006(04): 468-470+475.
[6] Guo C Y. Hyperspectral Imagery classification based on support vector machine [D]. Harbin Engineering University, 2007.
[7] Wang X L. Ensemble methods for Spectral-space classification of urban hyperspectral data [D]. Jilin University, 2010.
[8] Song X F, Jiao L C. Classification of Hyperspectral Remote Sensing Image Based on Sparse Representation and Spectral Information [J]. Journal of Electronics and Information, 2012, 34(02): 268-272.
[9] Tan Y M, Xia W. Optimum Band Combination Based Hyperspectral Remote Sensing Image Classification [J]. Mapping and space Geographic Information, 2014, 37(04): 19-22.
[10] Liu J M, Luo F L, Huang H, Liu Y Z. Classification of Hyperspectral remote sensing images using correlation neighbor LLE [J]. Optical Precision Engineering, 2014, 22(06): 1668-1676.
[11] Li L, Ren Y M. Classification of hyperspectral data based on random forest [J]. Computer Engineering and Applications, 2016, 52(24): 189-193