Hyperspectral image classification based on spectral-spatial kernel principal component analysis network

Yanguo Fan¹, Shizhe Hou¹, and Dingfeng Yu²

¹College of oceanography and space informatics, China University of Petroleum (East China), Qingdao, China

²Institute of Oceanographic Instrumentation, Shandong Academy of Sciences, 37 Miaoling Road, Qingdao, China

Abstract. Hyperspectral imagery contains both spectral information and spatial relationships among pixels. How to combine spatial information with spectral information effectively has always been a research hotspot of hyperspectral image classification. In this paper, a Spatial-Spectral Kernel Principal Component Analysis Network (SS-KPCANet) was proposed. The network is developed from the original structure of Principal Component Analysis Network. In which PCA is replaced by KPCA to extract more nonlinear features. In addition, the combination of spatial and spectral features also improves the performance of the network. At the end of the network, neighbourhood correction is added to further improve the classification accuracy. Experiments on three datasets show the effectiveness of the proposed method. Comparison with state-of-the-art deep learning-based methods indicate that the proposed method needs less training samples and has better performance.

1 Introduction

Hyperspectral remote sensing started from multispectral remote sensing in the 1980s, with the rapid development of information processing technology, hyperspectral has gradually become a hot research field of remote sensing. The hyperspectral sensor can obtain the continuously spectral of the target object with high spectral resolution. In addition, the hyperspectral sensor can also obtain the spatial information of the position relationship, shape and structure between objects. Therefore, hyperspectral image contains both two-dimensional spatial information and one-dimensional spectral information, which has more obvious advantages in distinguishing different features and is widely used in recognition and classification tasks [1].

In the early stage of hyperspectral image classification, the statistical pattern recognition methods that performed well in multispectral image classification are used, including K-nearest-neighbours, maximum likelihood method and minimum distance method [2,3]. However, the above methods performed poorly in hyperspectral image classification. Later scholars began to design new hyperspectral image classification algorithm and attempted to use spatial-spectral information [4-7].

In recent years, the image classification algorithm based on deep learning has achieved good performance. Typical deep learning models include convolutional neural network (CNN) [8], deep belief network (DBN) [9], and stacked auto encoder (SAE) [10], etc. Among them, CNN is widely used in hyperspectral image classification due to its advantages in high-dimensional data expression [11-12]. However, CNN still has problems such as time-

consuming parameter training process and difficult parameter adjustment. Based on this, Chan [13] proposed a relatively simple principal component analysis network (PCANet) which can adapt to different tasks and different data. The essence of PCANet is a simplified CNN model, in which two-level cascaded principal component analysis is used to replace the convolutional filter in CNN. And such a simple PCANet is on par with some state of art methods on the tasks of image classification [13-14]. At present, PCANet is mainly used in handwritten numeral recognition, face recognition, object recognition and other recognition for single picture, seldom used in hyperspectral classification.

In this paper, we proposed a Spatial-Spectral Kernel Principal Component Analysis Network (SS-KPCANet) for hyperspectral image classification inspired by PCANet. Hyperspectral image contains certain nonlinear information while PCANet is mainly linear operation, considering the extraction of nonlinear feature may enhance the performance of the network, we improve the original structure of PCANet by the introduction of kernel function. In addition, the depth features of hyperspectral images are extracted by using both spatial and spectral information in the proposed method. Experiments on three datasets show the effectiveness of the proposed method. Comparison with state-of-the-art deep learning-based methods indicate that the proposed method needs less training samples and has better performance.

The rest of this paper is organized as follows. In Section 2 we will describe the details of the proposed SS-KPCANet. The experimental results and comparison are provided in Section 3. Conclusions are drawn in Section 4.

* Corresponding author: author@e-mail.org

2 Spatial-Spectral Kernel Principal Component Analysis Network

PCANet is applied to the tasks of single image classification, while the pixels in hyperspectral image are represented as one-dimensional spectral vectors. In this

paper, spectral imaging is employed to reshape the spectral vectors to two-dimensional images. Our method extracts spectral and spatial features respectively through the above pattern. Then the extracted features are fused for SVM classification. Finally, a neighbourhood correction is adopted and the final results is obtained. The flow chart of the SS-KPCANet is shown in Figure 1.

Fig. 1. Flow chart of the SS-KPCANet

2.1. The structure of the Network

KPCANet can structure cascaded KPCA filters, assuming that there are $\mathbf{N}$ training picture samples, each sample size is $m \times n$, and the filter size of each layer is $k_1 \times k_2$, our model only need to learn KPCA filter core from the training sample set.

2.1.1 Cascaded KPCA filter extraction

For each sample, all the results of sampling with size $k_1 \times k_2$ are transformed into column vectors, then the mean-moved vectors make up matrix $\mathbf{X}_i \in \mathbf{R}^{(k_1 k_2) \times (\tilde{m} \tilde{n})}$, where $\tilde{m} = m - (k_1 - 1)$, $\tilde{n} = n - (k_2 - 1)$. All samples can construct a matrix :

$$\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N] \in \mathbf{R}^{(k_1 k_2) \times (N \tilde{m} \tilde{n})} \quad (1)$$

Let the number of filters in the first layer be L_1 , the purpose of this algorithm is to minimize the reconstruction error by a family of orthonormal filters:

$$\min_{\mathbf{V} \in \mathbf{R}^{k_1 k_2 \times L_1}} \|\mathbf{X} - \mathbf{V}\mathbf{V}^T \mathbf{X}\|_F^2, \text{ s.t. } \mathbf{V}^T \mathbf{V} = \mathbf{E}_{L_1} \quad (2)$$

where $\mathbf{E}_{L_1}$ is identity matrix of size $L_1 \times L_1$, the solution $\mathbf{V}$ is known as the L_1 principal eigenvectors of $\mathbf{X}\mathbf{X}^T$. Hence the KPCA filters can be expressed as:

$$\mathbf{W}_l^1 \doteq \text{mat}_{k_1, k_2} \{q_l[K(\mathbf{X})]\} \in \mathbf{R}^{k_1 k_2}, \quad l = 1, 2, \dots, L_1 \quad (3)$$

where $\text{mat}_{k_1, k_2}(\cdot)$ is a function that can reshape a vector to a matrix, $q_l[K(\mathbf{X})]$ means the l th principal eigenvector of $K(\mathbf{X})$, $K(\mathbf{X})$ denotes matrix $\mathbf{X}$ with kernel transformation.

Then, the output of the first layer can be obtained by

$$\mathbf{I}_i^1 = \mathbf{I}_i * \mathbf{W}_i^1, \quad i = 1, 2, \dots, N \quad (4)$$

where $\mathbf{I}_i$ is the input image of the first layer.

The second layer almost repeating the same process as the first layer. Suppose the number of filters in the second

layer is L_2 , KPCANet will have $L_1 \times L_2$ output images through two layers. One also can stack multiple layers so that deeper features can be extracted. In [13], the authors suggested that two layers are enough for most tasks.

2.1.2 Output layer: hashing and histogram

For each of the L_1 input images, we can obtain L_2 output images $\mathbf{O}_i^l = (\mathbf{I}_i^l * \mathbf{W}_i^l)_{l=1}^{L_2}$ from the second layer. Then a binary quantization is followed and $H(\mathbf{O}_i^l)$ acquired, where $H(\cdot)$ is a Heaviside step (like) function that can divide positive numbers to one and others to zero.

Then turn the L_2 binary bits to decimal number by convert the outputs into a single integer-valued "image":

$$\mathbf{T}_i^l = \sum_{l=1}^{L_2} 2^{l-1} H(\mathbf{O}_i^l) \quad (5)$$

so that every sample is an integer in the range $[0, 2^{L_2} - 1]$.

Subsequently, each of the L_1 images is separated into B blocks. The histogram of the values in each block is computed, and all the B histograms were collected into one vector $Bhist(\mathbf{T}_i^l)$. The L_1 vectors is the features expression for the input $\mathbf{I}_i$:

$$\mathbf{f}_i = [Bhist(\mathbf{T}_i^1), \dots, Bhist(\mathbf{T}_i^{L_1})]^T \in \mathbf{R}^{(2^{L_2}) L_1 B} \quad (6)$$

2.2 Spatial-spectral information

In SS-KPCANet, two kinds of features are extracted, features of original pixel spectral and features of pixel combined with its neighbourhood. Then the extracted features are fused for SVM classification. We choose a size of 3×3 to take advantage of neighbourhood information, which is widely adopt to describe the spatial information.

Furthermore, at the end of the network, we use neighbourhood correction to further improve the accuracy

of classification. Select the most categories by the statistic number of categories in the region with size of $k \times k$.

2.3 Comparison with nonlinear spectral-spatial Network

Nonlinear spectral-spatial Network (NSSNet) [14] is an early attempt to apply PCANet to hyperspectral image classification, authors enhance the original PCANet by adding nonlinear mapping and spatial-spectral method. The experimental results indicated that the NSSNet achieved better performance than state-of-the-art methods. But the NSSNet ignored the influence of different parameters on the results and used up to 50% samples for training, which leading to the training process is still time-consuming.

The SS-KPCANet has a similar structure to the NSSNet but performs better than it. We use the same imaging strategy to extract the two kinds of features, while the NSSNet used different strategies, which may affect the correct expression of the features. In the parameter setting stage, our method determines the optimal parameters through a series of comparative experiments: 2 layers, 9×9 patch size, 8 filters for each layer, 7×7 size of histogram block and 0.8 overlap ratio, rather than the default parameters of PCANet.

3 Experiment and discussion

In the experiment, three hyperspectral datasets were selected, namely Indian Pines, Salinas and Pavia University scenes.

Indian Pines dataset was obtained by Airborne Visible Infrared Imaging Spectrometer (AVIRIS) in northwestern Indiana, with 25 meters pixel size, 145×145 image size and 200 available bands. It includes 16 categories that can be used for training and classification, such as farmland, forest, highway and housing.

Salinas dataset was collected by AVIRIS in Salinas Valley, California, with 3.7 meters spatial resolution, 512×217 image size and 204 available bands. There are 16 categories for training and classification, including vegetables, bare soil and vineyards.

Pavia University dataset was gained by Reflective Optics System Imaging Spectrometer (ROSIS) at the University of Pavia, Italy, with 1.3 meters pixel size, 610×340 image size and 103 available bands. There are 9 categories including asphalt pavement, shadow, grassland and building.

The parameters of the network are set as mentioned above: 2 layers, 9×9 patch size, 8 filters for each layer, 7×7 size of histogram block and 0.8 overlap ratio. All samples were randomly divided into training set and test set according to the ratio of 1:9.

Firstly, we have carried out a comparative experiment of the SS-KPCANet, the original PCANet and Support vector machine (SVM) with Radial basis function (RBF) in hyperspectral image classification. The original PCANet shared the same parameters with our SS-KPCANet, and the parameters of SVM(RBF) are obtained by Crossvalidation. All the methods are added with

neighbourhood correction to ensure the fairness of the experiment. The classification results are displayed in Figure 2, Figure 3 and Figure 4, and the classification accuracies are revealed in Table 1.

Fig.2. Classification results of Indian Pines dataset. (a) The ground truth. (b) Results of SS-KPCANet. (c) Results of original PCANet. (d) Results of SVM(RBF).

Fig.3. Classification results of Salinas dataset. (a) The ground truth. (b) Results of SS-KPCANet. (c) Results of original PCANet. (d) Results of SVM(RBF).

Fig.4. Classification results of Pavia University dataset. (a) The ground truth. (b) Results of SS-KPCANet. (c) Results of original PCANet. (d) Results of SVM(RBF).

Tab.1 Classification accuracy with different methods of three datasets

	Indian Pines			Salinas			Pavia University		
	OA	AA	K	OA	AA	K	OA	AA	K
SS-KPCANet	99.24	99.35	0.994	99.18	99.16	0.991	99.05	98.45	0.987
PCANet	97.23	98.01	0.971	97.25	98.24	0.976	96.59	97.48	0.955
SVM(RBF)	86.25	89.87	0.84	84.06	81.58	0.82	91.43	80.41	0.885

As can be seen from figures 3 to 5, the performance of our method is better than the original PCANet. Which shows that the combination of spatial-spectral information and the utilization of kernel function improve the ability to extract depth feature of the PCANet in hyperspectral images. The accuracy evaluation indexes of the experiment include overall accuracies (OA), average accuracies (AA) and kappa coefficients (K) which are displayed in table 1. Compared with the original PCANet, the overall accuracy of SS-KPCANet is improved by 2.01% in Indian Pines dataset, 1.93% in Salinas dataset and 2.46% in Pavia University dataset. Note that the high accuracy of the original PCANet is corrected by neighbourhood, without this correction, the accuracy is only 88.31%, 86.76% and 91.11% respectively, which also indicates the effectiveness of our improvements. Similar results have been obtained from AA and K. Comparison with traditional machine learning algorithms SVM(RBF) shows that the SS-KPCANet reduces the occurrence of misclassification and makes the same region more uniform, OA increased by 12.99%, 15.12% and 7.62% on three datasets. Training time is also analysed in our paper, the SS-KPCANet spends more time than the PCANet because of the use of spatial information and the introduction of kernel function, but the high precision shows that the increase of training time is acceptable.

Table 2. Comparison between SS-KPCANet and some state-of-art algorithms

	Indian Pines			Pavia University			training sample
	OA (%)	AA (%)	K	OA (%)	AA (%)	K	
SAE-LR	93.98	93.81	0.931	98.46	97.67	0.980	50%
DBN-LR	95.91	94.20	0.953	98.99	98.38	0.987	50%
CNN	91.36	78.28	0.901	96.21	84.48	0.950	10%
DCNN-LR	95.97	93.23	0.954	99.03	98.19	0.987	50%
3D-CNN	99.07	98.66	0.989	99.39	98.85	0.992	50%
SS-KPCANet	99.24	99.35	0.994	99.05	98.45	0.987	10%

The comparison with some state-of-the-art deep learning-based methods advanced algorithms on Indian Pines and Pavia University can be seen in Table 2, namely: SAE-LR[15], DBN-LR[16], CNN[17], DCNN-LR[18] and 3D-CNN[19]. In terms of classification accuracy, the SS-KPCANet shows the best performance on Indian Pines dataset, and it is only 0.34% lower than 3D-CNN on Pavia University dataset. When considering the training

sample ratio, only CNN and our SS-KPCANet are trained with 10% samples, and SS-KPCANet is better than CNN. In addition, the training process based on deep learning method is quite time-consuming, usually takes tens of minutes or even hours, and also requires a strict operating environment (such as GPUs). SS-KPCANet performs similarly, and even better, with less training time. For example, the training time of SS-KPCANet is only 210.44 seconds on Indian Pines and 446.47 seconds on Pavia University. In summary, SS-KPCANet has better performance in hyperspectral image classification, with a simple structure and fewer training samples, and it is an ideal model for hyperspectral image classification tasks.

4 Conclusion

In this paper, a spatial-spectral combined kernel principal component analysis network SS-KPCANet is proposed based on the original principal component analysis network. SS-KPCANet is a simplified deep learning model based on convolutional neural network, in which the original convolution filter is replaced by a simpler KPCA filter. Our method shows excellent performance in the classification of hyperspectral images by combining spatial-spectral information and introducing kernel function. The accuracy of the classification was further improved by the neighbourhood correction at the end of the classification stage. Experiments show that the SS-KPCANet also shows obvious classification advantages in comparison with some state-of-the-art algorithms. Further work should focus on optimizing the algorithms and discussing the impact of spectral imaging on the original feature expression.

Acknowledgement

The study was financially supported by the Key R & D project of Shandong Province (grants No.2019GHY112017)

References

1. P. J. Du, J. S. Xia, Z. H. Xue, K. Tan, H. J. Su, R. Bao, Journal of Remote Sensing. **20(2)**, 236–256 (2016)
2. G. Camps-Valls, D. Tuia, L. Bruzzone, J. A. Benediktsson, IEEE Signal Process. Mag. **31(1)**, 45–54 (2014).
3. A. Agathos, J. Li, D. Petcu, A. Plaza, IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens. **7(6)**, 2281–2296 (2014)
4. M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, J. C. Tilton, Proc. IEEE. **101(3)**, 652–675 (2013)
5. L. Zhang, L. Zhang, D. Tao, X. Huang, IEEE Trans. Geosci. Remote Sensing. **50(3)**, pp. 879–893 (2012)
6. X. Kang, S. Li, J. A. Benediktsson, IEEE Trans. Geosci. Remote Sensing. **52(5)**, pp. 2666–2677 (2014)

7. J. Li, X. Huang, P. Gamba, J. M. Bioucas-Dias, L. Zhang, J. Atli Benediktsson, and A. Plaza, *IEEE Trans. Geosci. Remote Sensing*. **53(3)**, pp. 1592–1606 (2015)
8. A. Krizhevsky, I. Sutskever, G. E. Hinton, *Commun. ACM*. **60(6)**, 84-90 (2017)
9. G. E. Hinton, S. Osindero, Y. W. Teh, *Neural Comput.* **18(7)**, 1527-1554 (2014)
10. G. E. Hinton, R. R. Salakhutdinov, *Science*. **313(5786)**, 504-507 (2006)
11. S. Q. Yu, S. Jia, C. Y. Xu, *Neurocomputing*. **219**, 88-98 (2017)
12. H. Wei, Y. Huang, W. Li, F. Zhang, H. C. Li, *J. Sens.* 258619 (2015)
13. T. H. Chan, K. Jia, S. H. Gao, J. W. Lu, Z. N. Zhang, Y. Ma, *IEEE Trans. Image Process.* **24(12)**, 5017-5032 (2014)
14. B. Pan, Z. W. Shi, N. Zhang, S. B. Xie, *IEEE Geosci. Remote Sens. Lett.* **13(12)**, 1782-1786 (2016)
15. Y. S. Chen, Z. H. Lin, X. Zhao, G. Wang, Y. F. Gu, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **7(6)**, 2094-2107 (2014)
16. Y. S. Chen, X. Zhao, X. P. Jia, *IEEE J. Sel. Top. Appl. Earth Observ. Remote Sens.* **8(6)**, 2381-2392 (2015)
17. Y. Lu, L. D. Wu, J. Zhu, *Computer Engineering and Design*. **39(09)**, 144-149 (2018)
18. J. Yue, W. Z. Zhao, S. J. Mao, H. Liu, *Remote Sens. Lett.* **6(6)**, 468-477 (2015)
19. Y. Li, H. K. Zhang, Q. Shen, *Remote Sens.* **9(1)**, 67 (2017)