Hyperspectral Image Classification Based on Convolutional Neural Network and Dimension Reduction

Xuefeng LIU, Qiaoqiao SUN, Bin LIU, Biao HUANG
College of Automation & Electronic Engineering
Qingdao University of Science and Technology
Qingdao, China

Min FU
College of Information Science & Engineering
Ocean University of China
Qingdao, China

***Abstract*—One of the most important ways to explore the information in hyperspectral images (HSIs) is accurate classification of targets. Deep learning algorithm has made a great breakthrough in many areas due to its strong ability of data mining. Typical deep learning models such as convolutional neural network (CNN), deep belief network (DBN) and so on, not only combines the advantages of unsupervised and supervised learning but also have a good performance in large data classification. In this paper, a hybrid classification method combined CNN with dimension reduction (DR) operated by principal component analysis (PCA) is proposed, which fully takes the spatial information and the spectral characteristics of HSI into account. Furthermore, in order to solve the problem of sample imbalance, virtual samples are introduced to the experiments. Numerical results show that the proposed DR-CNN method, especially with virtual samples (named as DR-CNN-vs method in this paper) has promising prospect in the field of HSI classification.**

Keywords-classification; convolutional neural network (CNN); dimension reduction (DR);principal component analysis (PCA); virtual samples, remote sensing image

I. INTRODUCTION

With the development of hyperspectral sensors and imaging spectrometer, hyperspectral imagery has become one of the most important remote sensing technologies, which has been widely used in military, agriculture, geology, resources [1-4] and so on. To accurately explore the rich information in HSIs, some processing methods have been developed and applied, for instance, classification, target detection and unmixing [5], etc. Due to the high complexity of HSI in both spatial and spectral domain, such as intra-class differences, strong relevance among bands and insufficient training samples [6-8], etc. the classification results obtained from traditional methods need to be further improved.

CNN has gained widespread attention because it shown strong analysis capabilities of big data and can help to extract inherent laws and characteristics of data [9]. It has made a great breakthrough in pattern recognition [10-11], image classification [12] and language processing [13], etc.

Recently, CNN was introduced into HSI processing and achieved some preliminary results [14-16]. To make full use of the spatial and spectral information of HSIs, a hybrid classification method based on dimension reduction and convolutional neural network is proposed in this paper. Firstly, F-norm is used to reduce the dimension of HSI roughly and spectral information could be extracted according to the spectral bands of HSI. Next, PCA is used to further reduce the dimension and spatial characteristics can be obtained. Then, reshaping the 1-D vector (combined the spatial characteristics with the spectrum information of HSI) to a 2-D array as the input of CNN. Finally, the targets can be classified through the trained CNN network. The proposed classification method, named as DR-CNN method, has been applied to the real-world HSI and compared to other considered methods in the experiments.

The remainder of the paper is organized as follows: CNN network is overviewed in Section II; the proposed DR-CNN classification method for HSI is introduced in Section III; some experimental results are contained in Section IV, and Section V concludes the paper.

II. OVERVIEW OF CNN

Deep learning algorithm can improve the accuracy of classification and prediction by constructing models of multiple hidden layers and massive training data [17]. However, with the increase in the number of layers, the network parameters will be in a great augment which will cause a large amount of computation and overfitting.

In order to reduce the network parameters, CNN has a special architecture with local connection and shared weights [18] as shown in Fig. 1.

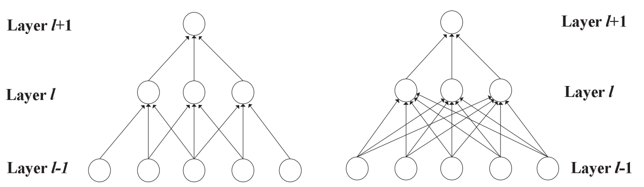


Figure 1. Local connection (left) and full connection

The work was supported by the National Natural Science Foundation of China (Grant No. 61401244, 61773227); the Shandong Provincial Science and Technology Development Project, China (Grant No. 2013YD01033) and the Science and Technology Project of Shandong Provincial University (Grant No. J15LN39).

978-1-5386-3524-7/17/$31.00 ©2017 IEEE

The number of network parameters can be reduced effectively through the local connection, but this is not enough. if the same weights and biases are used for all neurons in each feature map, named as shared weights, the network parameters will be further reduced. As shown in Fig. 2, the weights are encoded by color.

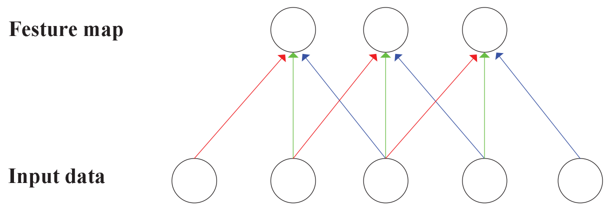


Figure 2. Shared weights

Conventional CNN framework is mainly composed of input layer, convolution layers, pooling layers and full connection layers as shown in Fig. 3. Different feature maps can be obtained by convolution input data with different convolution kernels, and each feature map represents a feature.

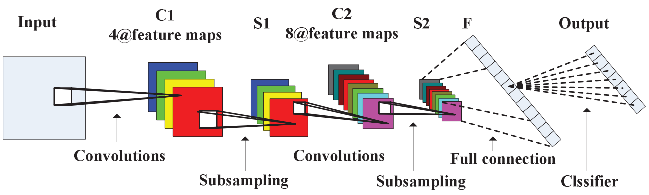


Figure 3. Conventional CNN framework

In convolution layer, the value of a neuron at position (x, y) of the mth feature map in the lth layer is given by [19]

$$v_{xy}^{lm} = f\left(\sum_{p}\sum_{n1=0}^{N_1-1}\sum_{n2=0}^{N_2-1} w_{lmp}^{n_1 n_2} v_{(l-1)p}^{(x+n_1)(y+n_2)} + b_{lm}\right) \tag{1}$$

where p indexes the feature map in the (l − 1)th layer connected to the current (mth) feature map, $w_{lmp}^{n_1 n_2}$ is the weights of position (n_1, n_2) connected to the pth feature map, N_1 and N_2 are the height and the width of the convolution kernel, and b_{lm} is the bias of the mth feature map in the lth layer. f is the activation function, the commonly used activation function such as relu, sigmoid and tanh, etc.

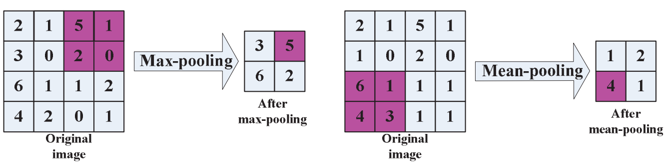


Figure 4. Max-pooling and mean-pooling (with 2×2 kernels and stride 2)

Subsampling can offer invariance by reducing the resolution of the feature maps [20], and pooling is a common method of subsampling, for example max-pooling, mean-pooling (see Fig. 4) and so on. After several layers of convolution and pooling, the feature maps can be flattened into a 1-D vector, which is fully connected to the front layer. In the last layer a softmax or other classifiers is selected as the classifier.

The whole CNN network is trained by back propagation (BP) algorithm.

III. CLASSIFICATION

A. Spectral Information

Hyperspectral remote sensing technology mainly uses a lot of very narrow band to obtain the relevant three-dimensional image data from the region of interest. Each pixel can be represented by a spectral curve, so, the target can be classified through their spectral information.

Different dimensions of HSI correspond to different energies represented by the square of F-norm:

$$\left\|R_1(:,:,i_3)\right\|_F^2 = \sum_{i_3=1}^{I_3}\left|R(:,:,i_3)\right|^2 \qquad (i_3=1,2,\ldots I_3) \tag{2}$$

where the raw HSI with I_1 rows, I_2 columns and I_3 spectral bands. The smaller the corresponding F-norm value of the image, the lower the corresponding energy and the less useful information contained. The dimensional of image with the square of F-norm less than a certain value σ with more external disturbance such as water vapor will be removed. Therefore, the dimension of HSI could be reduced by F-norm to improve the classification efficiency. The new HSI can be represented by $R_2(:,:,I_4)$ and I_4 is less than I_3. Each pixel in R_2 can be represented by a 1-D (1×I_4 and I_4 being the spectral dimension) vector.

B. Spatial Characteristics

A HSI not only reflects the spectral information with hundreds of bands but also contains the spatial characteristics. In order to obtain spatial characteristics, PCA is used to further reduce the dimension based on principal components (PCs) of a HSI. If the number of PCs of a HSI is I_5 and S_1×S_2 neighbor pixels are selected, a pixel block can be represented as S_1×S_2×I_5 members [21]. The spatial characteristics of each pixel can be obtained by unfolding the S_1×S_2×I_5 tensor into a 1-D vector with $S_1S_2I_5$ elements.

C. Clasiification with DR-CNN

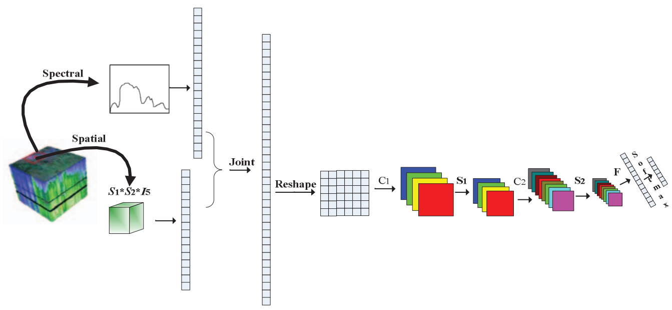


Figure 5. Classification based on DR-CNN

To make full use of the information in both spatial and spectral domains of HSI, a DR-CNN classification method based on dimension reduction and CNN is proposed, as shown in Fig. 5.

Firstly, the dimension of HSI could be roughly reduced by F-norm and the spectrum information could be extracted. Secondly, PCA is used to further reduce the dimension and the spatial characteristics could be obtained. Thirdly, the spatial characteristics are combined with the spectrum information of HSI to form a new 1-D vector and reshaping the 1-D vector to 2-D array as the input of CNN. Finally, a softmax is selected as the classifier in the last layer and a BP network is set up to fine-tune the entire CNN network. Once the network is trained, the target can be classified by the network.

IV. Experiment

A real-world hyperspectral data, HYDICE (Hyperspectral Digital Imagery Collection Experiment) HSI, is considered in this paper to evaluate the classification capability of the proposed method. As shown in Fig. 6, it has 141 rows, 126 columns and 210 spectral bands.

Figure 6. HYDICE HSI

A. *Spectral-spatial Information Extraction*

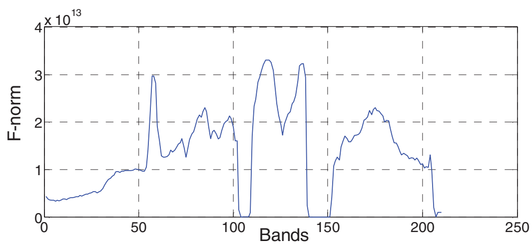


Figure 7. F-norn of HYDICE HSI

According to the F-norm of the raw HYDICE HSI in Fig. 7, the spectral bands with the square of F-norm in (2) less than 0.5×10^{13} are removed and about 142 bands have been remained. the spectral information of each pixel can be extracted as a 1-D vector with 1×142 elements from HYDICE HSI based on the spectral bands as introduced in subsection III. *A*. Then, PCA could be used to further reduce the dimension and 6 PCs are selected for the HYDICE image. For each class in HYDICE HSI, a box with 3×3×6 is used and unfolded to a 1-D vector. So, the spatial characteristics of each pixel can be represented by a 1-D vector with 1×54 elements as mentioned in subsection III. *B*. Finally, the spatial and spectral information are combined to a new 1-D vector with 1×196 elements and reshaping the 1-D vector to a 2-D (14×14) array as the input of CNN.

B. *Network Construction*

The classification results will be affected by the depth of the network, the number of convolution kernels and the size of kernel, etc. Taking the classification results and running time into consideration, a CNN network as shown in Table I is established based on experiments. Sigmoid is selected as the activation function in the convolution layer and max-pooling is chosen in the pooling layer.

TABLE I. Kernal Size And Feature Maps Of Proposed Method

Layer	**Input**	**C_1**	**S_1**	**C_2**	**S_2**	**F**	**Output**
Kernel size	14×14	5×5	2×2	4×4	2×2	1×32	1×12
Feature maps	1	16	16	32	32	1	1

C. *HYDICE HSI Classification*

The DR-CNN model is pre-trained by the training samples and then, BP algorithm is used to fine-tune the network. 13 land-cover classes of the HYDICE HSI are listed in Fig. 8, the epoch is 100 and the learning-rate is 1 in this experiment.

Figure 8. Classes in the HYDICE HSI

To evaluate the classification performance of the proposed DR-CNN method with spectral and spatial information, CNN with only spectral information, deep belief network (DBN) and support vector machine (SVM) are compared in this paper. DBN is also a typical deep learning model, by stacking the single restricted boltzmann machine (RBM) to extract features more efficiently [22]. A DBN network is constructed with 196-1000-2000-4000-12 units, there are 50 epochs of pre-training and 100 epochs of fine-tuning with BP algorithm in this experiment. SVM is one of the machine learning methods, it has unique advantages in the small sample, nonlinear and high dimensional pattern recognition. Radial basis function (RBF) is selected as kernel function of SVM, and the penalty parameter is 0.05 and gamma in kernel function is 100.

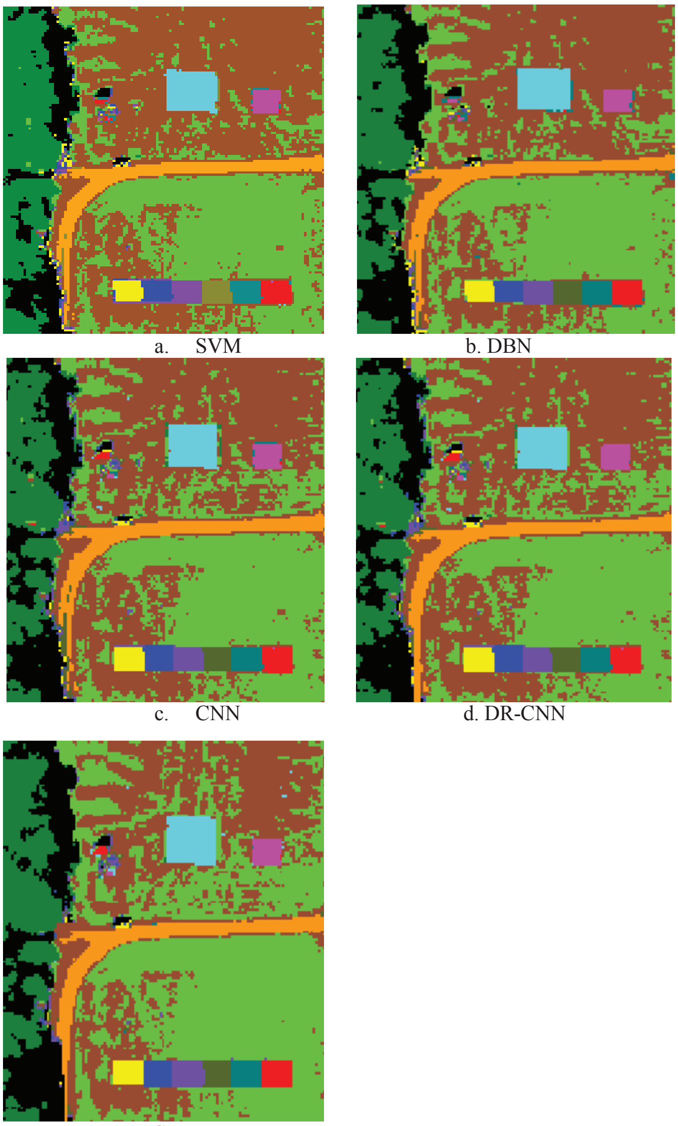
50 training samples in class 1 to 9 and 500 training samples in class 10 to 13 are selected respectively. The whole image is

used as a test set. It can be seen that the number of samples in the first nine classes is much less than that of the latter three classes, i.e. there is a sample imbalance problem [23]. The network fits better on classes with more samples while fits poor on classes with less samples, which will have an impact on the classification results. In order to solve the problem, virtual samples are introduced to DR-CNN, which is named as DR-CNN-vs:

$$v = \alpha + n \tag{3}$$

where α is a sample in training set, n is the Gaussian noise. Generally, the parameters of the noise are set according to the characteristics of the actual sample. The setting of the parameters makes the virtual sample similar to the sample in same class and different from the other classes. The noise with a mean of 0 and variance of 0.001 is selected in the experiment.

The classification results are illustrated in Fig. 9 and the overall accuracy (OA) and Kappa coefficient are listed in Table II.



a. SVM

b. DBN

c. CNN

d. DR-CNN

e. DR-CNN-vs

Figure 9. Classification results of HYDICE HSI

From Fig. 9, it can be seen that the first three methods have more misclassified points, and the classification results of the edge of Class 1 to 8 are particularly not ideal (see Fig. 9a, b and c). The road in the lower left has a wrong line caused by misclassifying, but after the combination of spectral and spatial characteristics, the wrong line in the road disappears and the wrong points at the edge are also significantly reduced (see Fig. 9d). The DR-CNN-vs method gives much clearer meadows texture above class 8 (see Fig. 9e) compared other considered methods.

TABLE II. OA AND KAPPA VALUES

Accuracy / Classifier	OA (%)	Kappa
SVM	91.28	0.8828
DBN	92.63	0.9070
CNN	92.56	0.9065
DR-CNN	94.29	0.9321
DR-CNN-vs	96.68	0.9552

It can be seen from Tab II that DR-CNN performs better than the first three methods. Therefore, taking into account the spectral information and spatial characteristics of HSIs can effectively improve the classification accuracy. Under the condition of unbalanced samples, the introduction of virtual samples will make the classification results more ideal.

V. CONCLUSION

To make full use of spectral and spatial characteristics and improve the classification of HSI and improve the classification, a method based on dimension reduction and CNN is proposed in this paper. F-norm and PCA are introduced to reduce the dimension of HSI. Then, CNN is used to the HSI classification. Furthermore, virtual samples are used to solve the problem of unbalanced samples. The experimental results show that the proposed DR-CNN and DR-CNN-vs methods, have promising prospect in HSI classification compared to other considered methods.

The results encourage us to extend our experiments on target detection and unmixing of HSI.

REFERENCES

[1] Gao Zhenheng, "Research on Classification Techine for Hyperspectral Remote Sensing Imagery," National University of Defense Technology, 2011.

[2] K. Makantasis, K. Karantzalos, "Deep supervised learning for hyperspectral data classification through convolutional neural networks," IEEE International Geoscience and Remote Sensing Symposium., pp. 4959-4962, 2015.

[3] Y. Zhang, JP. Zhang, M. Jin, M.D. Desai, "Adaptive subspace decomposition and classification for hyperspectral image," Chinese Journal of Electronics, vol. 9. No. 1, pp. 82-88, 2000.

[4] S. Bernabe, S. Lopez, A. Plaza, R. Sarmiento, "GPU implementation of an automatic target detection and classification algorithm for hyperspectral image analysis," IEEE Geoscience & Remote Sensing Letters, vol. 10, pp. 221-225 , 2013.

[5] Z. Chen, B. Wang, Y Niu, W Xia, T QiuZhang, B Hu, "Change detection for hyperspectral images based on tensor analysis," IEEE Geoscience and Remote Sensing Symposium. , 2015:1662-1665.

[6] A. Romero, C. Gatta, G. Camps-Valls, “Unsupervised Deep Feature Extraction for Remote Sensing Image Classification,” IEEE Transactions on Geoscience & Remote Sensing, vol. 54, no. 3, pp.1-14, 2015.

[7] B. Liao, J. Xu, J. Lv, S. Zhou, “An Image Retrieval Method for Binary Image Based on DBN and Softmax Classifier,” Iete Technical Review, vol. 32, no. 4, pp. 294-303, 2015.

[8] Chen Jin, “On Classification Method of Hyperspectral Image,” National University of Defense Technology, 2010.

[9] A. Krizhevsky, I. Sutskever, G. Hinton, “ImageNet classification with deep convolutional neural networks,” International Conference on Neural Information Processing Systems, pp. 1097-1105, 2012.

[10] Eric Tatulli, Thomas Hueber, “Feature extraction using multimodal convolutional neural networks for visual speech recognition,” IEEE Acoustics, Speech and Signal Processing, pp. 2971-2975, 2017.

[11] Y. Zhong, J. Sullivan, H. Li, “Face attribute prediction using off-the-shelf CNN features,” International Conference on Pattern Recognition, pp. 1-7, 2016.

[12] W. Zhao, L. Jiao, W. Ma,J. Zhao, J. Zhao, H. Liu, et al, “Superpixel-Based Multiple Local CNN for Panchromatic and Multispectral Image Classification,” IEEE Transactions on Geoscience & Remote Sensing, pp.1-16, 2017.

[13] R.Socher, C.Y. Lin, A.Y. Ng, CD Manning, “Parsing Natural Scenes and Natural Language with Recursive Neural Networks,” International Conference on Machine Learning, ICML 2011, Bellevue, Washington, Usa, pp. 129-136, 2011

[14] Michael Kampffmeyer, Arnt-Borre Salberg, Robert Jenssen, “Samantic Segmentation of Small Objects and Modeling of Uncertainty in Urban Remote Image Using Deep Convolutional Neural Networks,” IEEE Conference on Computer Vision and Pattem Recognition Worksshops, 2016.

[15] R. Kirthika, D.T.K. Shanthi, G.A. Manikandan, “Novel Method for Hyperspectral Image Classification Using CNN Based Laplacian Eigenmap Pixels Distribution Flow,” International Journal for Scientific Research & Development, vol. 2, pp. 833-836, 2014.

[16] J. Yang, Y. Zhao, C.W. Chan, C. Yi, “Hyperspectral image classification using two-channel deep convolutional neural network,” IEEE Geoscience and Remote Sensing Symposium., pp. 5079-5082, 2016.

[17] Y. Lecun, Y. Bengio, G. Hinton, “Deep learning,” Nature, vol. 521,pp.436-444, 2015.

[18] Y. Lécun, L. Bottou, Y. Bengio, P. Haffner, “Gradient-based learning applied to document recognition.” Proceedings of the IEEE, vol.86, no.11, pp. 2278-2324, 1998.

[19] Chen Yushi, Jiang Hanlu, Li Chunyang, Jia Xiuping, Pedram Ghamisi, “Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks,” IEEE Transactions on Geoscience & Remote Sensing, vol.54, no.10, pp. 6232-6251, 2016

[20] Z. Zuo, B. Shuai, G. Wang, X. Liu, X. Wang, B. Wang, Y. Chen, “Learning Contextual Dependence With Convolutional Hierarchical Recurrent Neural Network,” IEEE Transactions on Image Processing, vol.25 no. 27, pp. 2983-2996, 2016.

[21] Y. Chen, X. Zhao, X. Jia, “Spectral–Spatial Classification of Hyperspectral Data Based on Deep Belief Network,” IEEE Journal of Selected Topics in Applied Earth Observations & Remote Sensing, vol. 8,no. 62381-2392, 2015.

[22] G. Hinton , S. Osindero, Y.W. Teh, “A Fast Learning Algorithm for Deep Belief Nets,” Neural Computation, vol.18, no.7, pp. 1527-1554, 2014.

[23] H. He, E.A. Garcia, “Learning from Imbalanced Data,” IEEE Transactions on Knowledge & Data Engineering, vol.21,no. 9 pp. 1263-1284, 2009.