

PAPER • OPEN ACCESS

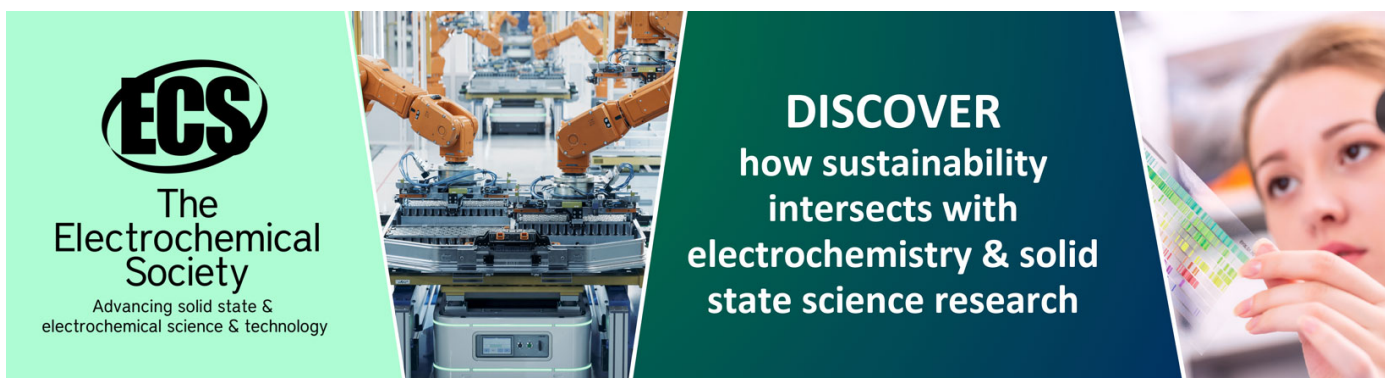
Hyperspectral Image Classification of Convolutional Neural Network Combined with Valuable Samples

To cite this article: Lixin Hu *et al* 2020 *J. Phys.: Conf. Ser.* **1549** 052011

View the [article online](#) for updates and enhancements.

You may also like

- [Predictive analysis on hyperspectral images using deep learning](#)
E Sarah Rajkumar and V Pattabiraman
- [Adaptive and automatic red blood cell counting method based on microscopic hyperspectral imaging technology](#)
Xi Liu, Mei Zhou, Song Qiu et al.
- [A Hyperspectral Image Classification Method with CNN Based on attention-enhanced Spectral and Spatial Features](#)
Yangming Zhang, Kun Yang and Lei Yuan



ECS
The
Electrochemical
Society
Advancing solid state &
electrochemical science & technology

DISCOVER
how sustainability
intersects with
electrochemistry & solid
state science research

Hyperspectral Image Classification of Convolutional Neural Network Combined with Valuable Samples

Lixin Hu^{1,a}, Xiaobo Luo^{2,3} Yufan Wei¹

¹Institute of Computer Science and Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

²Chongqing Engineering Research Center for Spatial Big Data Intelligent Technology, Chongqing University of Posts and Telecommunications, Chongqing, China

³Chongqing Institute of Meteorological Sciences, Chongqing, China

^ahulixxx2@163.com

Abstract. Aiming at the problem that the manual labeling of samples in the hyperspectral image classification is expensive and laborious, a large number of unlabeled samples are not effectively utilized and the classification results are not ideal. A method which can provide valuable samples and employ convolutional neural network to extract spectral spatial features for classification is proposed. Active learning method is used to construct a valuable training sample set by iteratively selecting the most uncertain samples through support vector machine which performs well in small sample classification, and labeling them. Then the 3D convolutional neural network is used to extract the spectral spatial features of hyperspectral image. The experimental results of the hyperspectral classification on Indian Pines and PaviaU datasets show that the proposed method (3D VS-CNN) is better than traditional classification methods.

1. Introduction

One of the most important tasks in the field of remote sensing in recent decades is image classification, in which the purpose is to identify objects in an image. The biggest characteristic of hyperspectral image (HSI) is that there are many spectral, which can provide the detailed spectral features of different objects, and is very useful for image classification [1].

Recently, convolutional neural network (CNN) has shown exciting results in the field of image classification [2,3]. Liu [4] employed CNN to encode the spectral and spatial information of pixels and classify them with multilayer perceptrons. Chen [5] proposed a convolutional framework for extracting spectral and spatial features with different convolution kernel sizes for HSI classification. The feature maps generated in each layer of CNN are connected directly to subsequent layers through connection modes for classification [6].

With sufficient labeled data, CNN can learn more effective features from raw pixels directly. In addition, due to the characteristics of weight sharing and sparse connection, CNN can extract the spatial information of an image, and requires fewer parameters. However, CNN adopts supervised learning classification methods, which often requires a large number of labeled data to achieve better classification results. But the process of training data collection requires a lot of time and labor, which makes the HSI classification full of challenges.



In this study, in order to provide valuable training samples for CNN, active learning (AL) is introduced in the paper. Through support vector machine (SVM), which performs well in small sample classification, the most uncertain samples are selected iteratively for labeling, and valuable training sample set is constructed. The spectral spatial information of HSI is extracted by 3D CNN. And the method (3D VS-CNN) proposed in the paper achieved good classification results on Indian Pines and PaviaU datasets.

2. Methodology

2.1. Active Learning (AL)

AL [7-9] only needs a small number of initial label samples to join the training set to extend the dataset by interacting with users. The key idea of AL is to employ selection criterion to find valuable samples in order to reduce the size of training set and avoid the disadvantages of blind target annotation [10]. The core of AL algorithm is to find valuable samples to construct training set by selection criterion, which is more purposeful than selecting samples randomly as training set. This paper mainly employs BvSB [11] method:

The probability of the first and second most probable class for a sample in the pool of unlabeled samples P is calculated and arranged in an ascending order. The first N samples with low probability difference are selected by each iteration. The lower BvSB value, the more uncertain is the sample. BvSB is defined as

$$\text{BvSB} = \arg \min_{x_i \in P} (P(y_{\text{First}} | x_i) - P(y_{\text{Second}} | x_i)) \quad (1)$$

The algorithm flow of AL is shown in Algorithm 1.

Algorithm 1 AL Algorithm Flow

- 1: Inputs: Initially training set L , pool of unlabeled samples P , and the number of samples N selected at each iteration.
- 2: Output: Valuable samples.
- 3: Repeat: Training SVM model with the training set L .
- 4: Employing selection criterion Equation (1) to evaluate the samples P and sort them, respectively.
- 5: Selecting N samples from P and labeling them as P_n .
- 6: Adding the labeled samples P_n to the training set L , while removing them from P .
- 7: Until the predefined condition is satisfied.

2.2. Proposed Method

In this paper, we mainly employ spectral spatial information of HSI to design the model, and the structure of our proposed 3D VS-CNN is shown in Figure 1. AL and SVM classifier are employed in the first part to make full use of spectral information of HSI to select valuable samples. Then, convolutional kernels are exploited to extract the spectral spatial features of HSI.

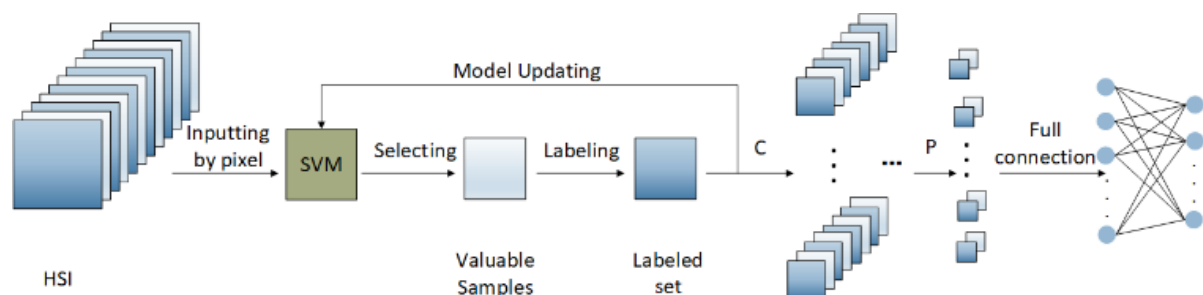


Figure 1. Illustration of our proposed 3D VS-CNN framework

The convolutional layer (C) is the core of 3D VS-CNN. Convolution operation is performed on multiple convolutional kernels and local feature maps, then the bias is added, and the obtained result is nonlinearly processed by the activation function ReLU as the input of the next layer [12]. The convolution operation can be expressed as:

$$O_3^{(l)} = A^{(l)} \left(\sum_k \sum_{i=0}^j \sum_{m=0}^h \sum_{n=0}^h W_{imn,k}^{(l)} * X^{(l-1)} + B^{(l)} \right) \quad (2)$$

where $A^{(l)}$ represents the activation function in the l th layer, $W_{imn,k}^{(l)}$ denotes the weights of convolutional filters at layer l , $X^{(l-1)}$ is the output in the $(l-1)$ th layer, $B^{(l)}$ denotes the bias, $*$ represents the convolution action, and imn represents the position connected to the feature map. The convolutional layer is usually followed by a pooling layer (P), which can reduce the size of the parameter matrix, enhance the feature information and reduce the number of parameters [13]. The pooling method has Max Pooling, Mean Pooling, etc. Here, the Max Pooling are employed, which is as follows:

$$\text{MaxP} = \max(a \ c_{(i,j)}) \quad (3)$$

where a represents the whole feature matrix, and $c_{(i,j)}$ denotes a window function to the $i \times j$ patch of feature matrix.

Then, softmax with strong non-linear classification ability as classifier is adopted in the full connected layer (FC), which makes the probabilities of all output neurons sums to 1. The probability of the input belongs to class i is as follows:

$$S(WX + B) = \frac{e^{W_i X + B_i}}{\sum_j e^{W_j X + B_j}} \quad (4)$$

where X denotes the input vector, W and B are the weight and bias.

3. Experiments and results

In order to evaluate the performance of our developed model, all experiments were implemented on Ubuntu 16.04 using Python language and Keras library, and based on the hardware configuration of Intel Core i7 with 3.70GHz and Nvidia Geforce GTX 1070.

3.1. The datasets

The first dataset is the Indian Pines dataset which was gathered by AVIRIS sensor in 1992 in Northwest Indiana. Indian Pines contains 145×145 pixels with a spatial resolution of 20m and 220 spectral bands. After removing the noise bands, 200 useful bands are retained. As shown in Figure 2a, this dataset contains a total of 21025 pixels, with the exception of background pixels, the remaining 10249 labeled pixels of 16 classes.

The second dataset is PaviaU dataset which was gathered by ROSIS sensor in 2003 in the city of Pavia, Italy. PaviaU contains 610×340 pixels with a spatial resolution of 1.3m and 115 spectral bands. After removing the noise bands, 103 useful bands are retained. As shown in Figure 2b, this dataset contains a total of 207400 pixels, with the exception of background pixels, the remaining 42776 labeled pixels of 9 classes.

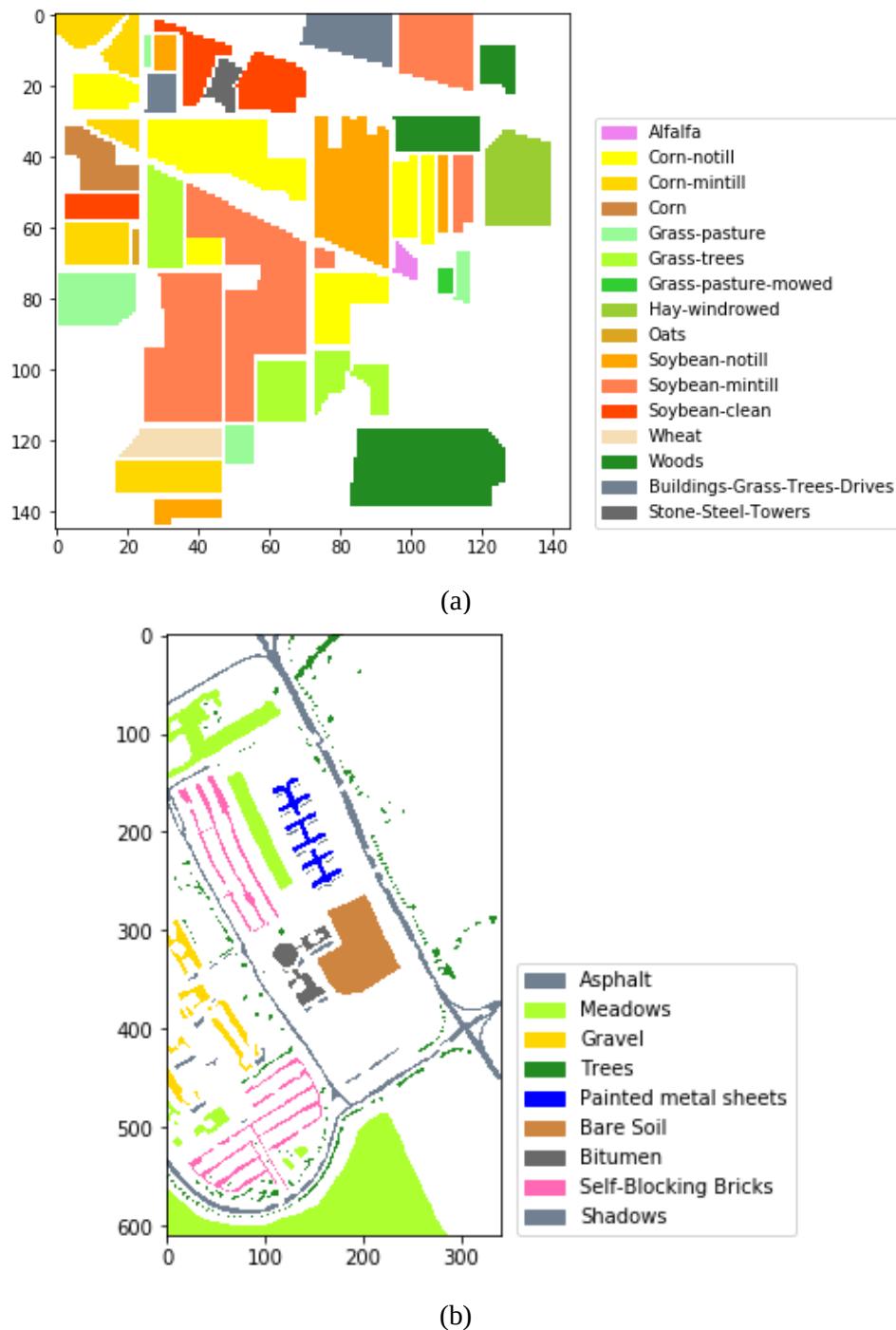


Figure 2. Ground truth. (a) Indian Pine dataset; (b) PaviaU dataset

3.2. Experiment Setting

For the two dataset, all the available samples are divided into two parts according to the ration of 5:5 firstly. Then, the training sample starts with $5 \times C$ labeled samples, while the remaining samples are utilized to initialize the pool set. After the initial samples and pool set are create, 95% of the rest are employed for testing and 5% for validation. TABLE I shows the architecture of 3D VS-CNN for two datasets. The () content indicates the difference between Indian Pine dataset and PaviaU dataset, and the rest is the same.

Table 1. The architecture of 3D VS-CNN on two datasets

3D VS- CNN	Kernel	Activation	Pooling	Dropout
	$20 \times 3 \times 3 \times 9$ ($20 \times 3 \times 3 \times 3$)	ReLU	$2 \times 2 \times 1$	-
	$40 \times 3 \times 3 \times 9$ ($40 \times 3 \times 3 \times 3$)	ReLU	$2 \times 2 \times 1$	5%
	Neurons	Activation	Dropout	
	80	ReLU	5%	
	16(9)	Softmax	-	

Moreover, PCA method is applied in the experiment. For Indian Pines dataset, 30 principal component bands are extracted from 200 bands by principal component analysis (PCA). These 30 bands contain more than 99% of the original spectral information and retain clear spatial information. And the PaviaU dataset leaves 10 bands. In addition, the spectral-spatial patches have been cut to 13 with the aim of extracting spectral and spatial information from neighboring pixels in two datasets.

3.3. Experiment Results

In order to evaluate the effectiveness of the model we proposed, passive learning algorithm (random method) and AL algorithm are experimented on two hyperspectral datasets, respectively. The random sample selection method and BvSB method have been applied in our experiments.

For the Indian Pines dataset, starting from 80 samples, SVC algorithm in SVM utilizes Gaussian kernel function. AL is applied to select 20 valuable samples for each iteration, and there are 100 iterations. For the PaviaU dataset, starting from 45 samples, and 100 valuable samples are selected for each iteration, and there are 60 iterations. Then, Figure 3 shows the classification performance on two datasets. (a) (c) represent the random method, and (b) (d) represent the method proposed in this paper. We can see clearly that the 3D VS-CNN can get good classification results.

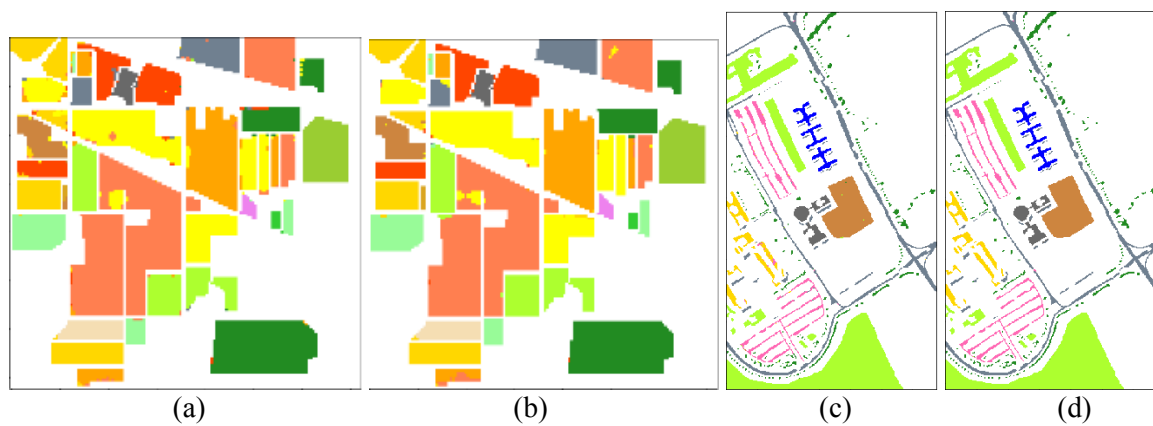


Figure 3. Classification maps for the two dataset. (a) Indian Pines (3DCNN); (b) Indian Pines (3D VS-CNN); (c) PaviaU (3DCNN); (d) PaviaU (3D VS-CNN)

In addition, we also compared our method with traditional machine learning methods. TABLE II shows the overall classification accuracy of different methods under the same training test ratio. In comparison, 3D VS-CNN achieves the highest overall classification accuracy, 98.27% in Indian pines and 99.48% in PaviaU.

Table 2. Overall classification accuracy of different methods using the same training and testing ratio

	KNN	SVM	1D CNN	2D CNN	3D CNN	3D VS-CNN
Indian Pines	69.59	69.89	85.54	92.13	96.43	98.27
PaviaU	86.50	81.56	92.36	98.46	98.86	99.48

4. Conclusions

Aiming at the problem of HSI classification, this paper proposes a HSI classification method based on CNN with valuable samples. 3D VS-CNN employs AL to construct valuable training samples, and employs CNN to extract spectral-spatial information of HSI to improve classification accuracy. Experiments on Indian Pines and PaviaU datasets show that this method is superior to other methods. In the future work, we plan to employ our proposed model for other remote sensing images, not just for HSI. In addition, we can combine multiple selection criteria to improve the classification performance, rather than just a single selection criterion. Finally, the model proposed in this paper can be combined with post-classification processing to improve the classification results, which deserves further study.

Acknowledgment

This work was supported by the National Natural Science Foundation of China (41871226), the Special Foundation of Postdoctoral Scientific Research Project of Chongqing (Xm2016081), the Open Fund Project of Chongqing Meteorological Bureau (KFJJ201602), and the Key Projects of Chongqing Application Development Plan (cstc2014yykfB30003).

References

- [1] W. Pin, Xu S and Li Y, "Hyperspectral image classification based on joint sparsity model with low-dimensional spectral-spatial features." *J Appl Remote Sens.* Vol. 11, 2017: 015010.
- [2] C. Yushi, X. Zhao, and X. Jia, "Spectral-Spatial Classification of Hyperspectral Data Based on Deep Belief Network." *IEEE J. Sel. Top. Appl. Earth Observ.*, vol. 8, 2015, pp.1-12.
- [3] Paoletti M. E, Haut, J. M. Fernandez-Beltran, R. Plaza, J. Plaza, A. Li, J, "Capsule networks for hyperspectral image classification." *IEEE Trans. Geosci. Remote.*, vol. 57, 2018, pp.1-16.
- [4] Feng J, Chen J, Liu L, Cao X, Zhang X, and Jiao, L, "CNN-based multilayer spatial-spectral feature fusion and sample augmentation with local and nonlocal constraints for hyperspectral image classification." *IEEE J. Sel. Top. Appl. Earth Observ.*, vol. 12, 2019, pp.1299-1313.
- [5] Liu P, Choo K. K. R, Wang L, and Huang F, "Svm or deep learning? A comparative study on remote sensing image classification." *Soft Comput.*, 2016.
- [6] Chen Y, Jiang H, Li C, Jia X and Ghamisi P, "Deep Feature Extraction and Classification of Hyperspectral Images Based on Convolutional Neural Networks." *IEEE Trans. Geosci. Remote.*, vol. 54, 2016, pp.6232-6251.
- [7] Paoletti M. E, Haut J. M, Fernandez-Beltran R, Plaza J, Plaza A, and Li J, "Capsule networks for hyperspectral image classification." *IEEE Trans. Geosci. Remote.* vol. 57, 2018, pp.1-16.
- [8] H. Lianzhi, and T. Ping, "A graph based active learning method for classification of remote sensing images." *International Journal of Wavelets, Multiresolution and Information Processing*, 2017.
- [9] Y. Jihai, L. Shijun, and X. Wenning, "Active learning for visual image classification method based on transfer learning." *IEEE Access*, vol. 6, 2018, pp.187-198.
- [10] J. Yuan, "Multi-criteria active deep learning for image classification." *Knowledge-Based Systems*, vol. 172, 2019, pp.86-94.
- [11] Zhang H.K, Li Y, Jiang Y.A, Wang P, Shen Q and Shen C.H, "Hyperspectral Classification Based on Lightweight 3-D-CNN With Transfer Learning." *IEEE Trans. Geosci. Remote.*, vol. 57, 2019, pp.5813-5828.
- [12] Z. Ce, X. Pan, Li H, Gardiner A and Atkinson P. M, "A hybrid mlp-cnn classifier for very fine resolution remotely sensed image classification." *ISPRS-J. Photogramm. Remote.*, vol. 140,

- 2018, pp.133-144.
- [13] C. Gong, Y. Ceyuan, Y. Xiwen, Guo L and Han J, “When deep learning meets metric learning: remote sensing image scene classification via learning discriminative CNNs.” IEEE Trans. Geosci. Remote, vol. 56, 2018, pp.1-11.