Self-Supervised Learning With A Dual-Branch ResNet for Hyperspectral Image Classification

Tianrui Li, Xiaohua Zhang, *Member, IEEE*, Shuhan Zhang, and Li Wang

***Abstract*—Deep learning methods have made considerable progress in many fields, but most of them rely on a large amount of sample. In the hyperspectral image classification task, many unlabeled data and few labeled data exist, so it is necessary to use a small number of training samples to achieve good results. In this paper, in order to fuse spectral and spatial information, a dual-branch residual neural network (ResNet) is proposed, with one branch for extracting spectral features and one branch for extracting patch features. Further, according to the properties of the hyperspectral image, self-supervised learning training methods are designed for these two branches. When spectral information is used for training, the image is artificially divided into several parts, with each part being a category for the classification task. When patch features are used for training, the task is to recover the spectral information of the intermediate pixels. After the pretext task training is completed, a pre-training weight will be provided for classification task training. Experiments with a small number of samples of two public data sets show that this method has better classification performance than existing methods.**

***Index Terms*—Deep learning, hyperspectral image classification, self-supervised learning, small-sample learning**

I. Introduction

HYPERSPECTRAL image (HSI) classification is an important technique in remote sensing. Because each pixel of an image has hundreds or even thousands of bands, the same area can be characterized by spectral and spatial information, so that richer and more refined information can be obtained. HSI classification has important uses in geological prospecting [1], urban planning [2], ecological assessment [3], and other fields.

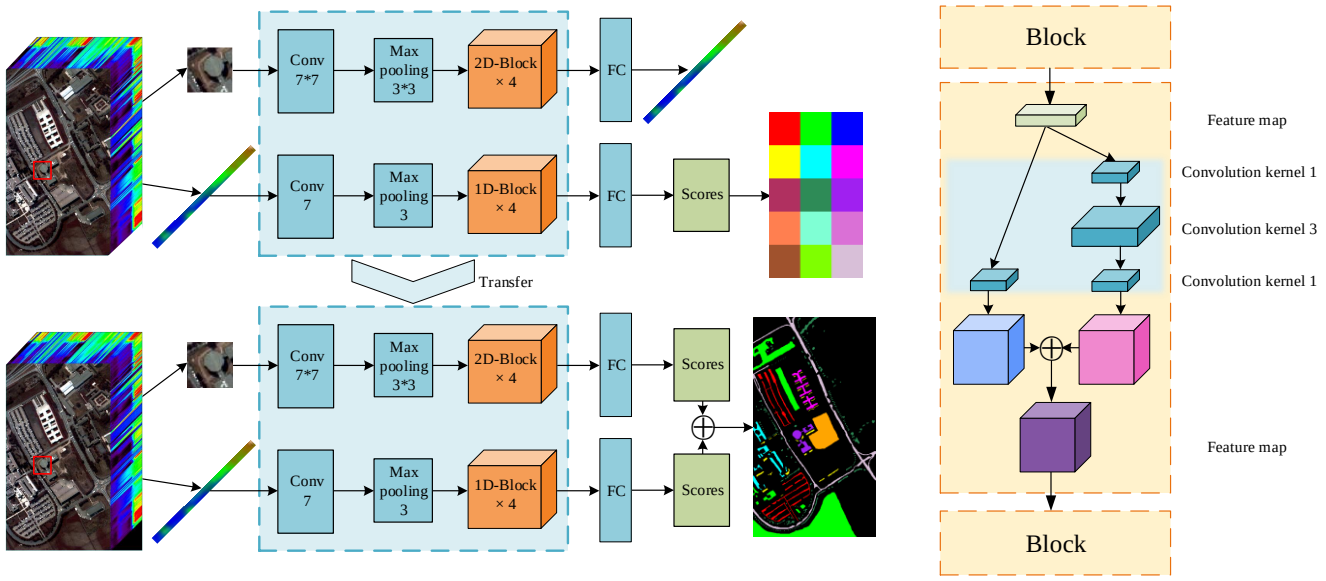
Traditional machine learning algorithms, such as the k-nearest neighbors (KNN) [4], support-vector machine (SVM) [5], and graph learning [6], have been used for classification. The traditional method uses a fixed feature extraction method, and the features extracted by similar spectrum vectors are relatively similar, and the phenomenon of the same spectrum in different categories is prone to appear in the HSI. The classification performance is relatively poor. These methods are not as effective as learning-based feature extraction technique.

With the development of deep learning, convolutional neural networks (CNNs) have made great progress in image classification [7], object detection [8], and action recognition [9]. CNNs are increasingly used for HSI classification [10]-[12]. These methods either use only spectral information or only spatial information. Although some are used at the same time, there is no synergy between them. Moreover, the number of channels of HSI data is large, and such complex data types usually cause more network parameters. Training a large network generally requires a large number of labeled samples, but there are usually fewer labeled samples and a large number of unlabeled samples in HSI data. Therefore, the use of deep learning for HSI classification needs to solve two problems: 1. How to use spatial information and spectral information jointly. 2. How to train a classification network with promotion ability based on a small number of labeled samples and a large number of unlabeled samples.

In order to combine spatial and spectral information, Xu et al. [13] fused the spectral features extracted by long short-term memory (LSTM) with the spatial features extracted by a CNN for classification. Roy et al. [14] first used 3D-CNN to extract spatial-spectral features, and then used 2D-CNN to further extract spatial features. These methods use feature extraction methods of different dimensions to extract spatial and spectral features. However, these models are limited by a small number of training labeled samples, so the network design is usually simple, which makes the network generalization not strong. In order to use unlabeled samples, self-supervised learning is proposed. Self-supervised learning [15] mainly uses the characteristics of the data to design pretext tasks, followed by transfer learning, in which the pre-training weights are used in downstream tasks to improve the learning process. Nowadays, many pretext tasks have been designed [16]-[18]. These schemes provide ideas for the pretext task design in HSI classification. Yue et al. [19] have proposed a three-dimensional transformation method. Transformation of data can increase the diversity of data and enhance the generalization ability of the network. Some methods [20], [21] used different levels of neighborhood information (resembling sub-pixel, pixel, and super-pixel features) to provide supervision information. However, these methods of using super-pixels first need to generate super-pixels based on data features, and the features of these data are what we need to learn and use for classification, so the spatial information obtained is based on spectral information as prior knowledge.

This work was supported in part by the National Natural Science Foundation of China under Grant 61877066. *(Corresponding author: Xiaohua Zhang.)*

Tianrui Li, Xiaohua Zhang, Shuhan Zhang, and Li Wang are with the School of Artificial Intelligence, Xidian University, Xi'an 710071, China, and also with the Key Laboratory of Intelligent Perception and Image Understanding of Ministry of Education, Xi'an, 710071, China. (e-mail: trli@stu.xidian.edu.cn; Xh_zhang@mail.xidian.edu.cn; zhangshuhan136@-163.com; 2512522206@qq.com)



(a) (b)

Fig. 1. (a) The overall framework of the HSI classification network proposed in this paper. The upper part is the self-supervised learning part, and the lower part is the classification part. (b) Details of the block in (a).

The classification results are largely affected by the spectral characteristics. Li et al. [22] combined the two samples (local cube centered at the training sample), and whether the two samples are of the same type is used as supervision information. However, this method is to design pseudo-labels in the part of the HSI data that has the category label. This leads to the fact that the rich features in the unlabeled data are not fully utilized. The key to self-supervised learning is how to design pretext tasks that are more effective for downstream tasks.

Therefore, in order to be able to learn features from all data and effectively combine the spectral and spatial features, this paper combines the characteristics of HSI classification to design pretext tasks for the spectral and spatial domains. Spatial information is introduced when spectral training is used, and spectral information is introduced when image patch training is employed. At the same time, in order to improve the ability of the network to extract features, we use the residual neural network (ResNet) [23] as the backbone network and propose a dual-branch ResNet (DBRNet) to fuse the two kinds of information.

II. Proposed Method

A. Overall Framework

The proposed HSI classification framework is divided into two branches. The first branch is the network for extracting spectral features. Using the network structure of Resnet50, we replaced the two-dimensional convolution, pooling, and batch normalization (BN) layers with one-dimensional layers so as to adapt to the input of spectral data. The second branch is the network that extracts the features of the image patches. The network structure of Resnet50 is used.

In the training phase, the pretext task is trained first, and the spectral branch is trained for classification by the pseudo-label of each spectrum. The loss function is the cross-entropy loss function:

$$\text{loss} = -\frac{1}{n}\sum_{i=1}^{n} y_i \ln x_i \tag{1}$$

Where y is the artificially designed category label, and x is the label predicted by the network.

The spatial branch is used to train the task of reconstructing the spectrum from the image patch, and the loss function is the mean squared error (MSE) loss function:

$$\text{loss} = \frac{1}{n}\sum_{i=1}^{n} \| R_i - S_i \|_2 \tag{2}$$

Where R is the spectral vector of the central pixel reconstructed by the network, and S is the true spectral vector of the central pixel vector.

After the two branches are trained, the trained model is transferred to the downstream task (HSI classification). The last fully connected (FC) layers of the two branches are replaced to perform HSI classification training, and each branch uses the cross-entropy loss function. In addition, considering that the spectral information and spatial information of the same pixel should be similar after the feature is extracted by the network, the feature before the last FC layer of each branch is taken out to calculate the MSE loss:

$$\text{loss} = -\frac{1}{n}\sum_{i=1}^{n} y_i(\ln x_i + \ln x_p) + \frac{\alpha}{n}\sum_{i=1}^{n} \| f_i - f_p \|_2 \tag{3}$$

where y_i is the label; x_i and x_p are the predicted category results obtained by the spectral branch and the spatial branch, respectively; f_i and f_p are the features of the spectral branch and the spatial branch, respectively, before they are passed through to the FC layer; and α is a hyperparameter.

Finally, the results obtained from the two branches of FC layers are added to obtain the classification result. The overall

TABLE I
HSI CLASSIFICATION RESULTS (%) BY SELECTING 20 LABELED SAMPLES FROM EACH CLASS OF THE UNIVERSITY OF PAVIA DATA

Class	SVM	3D-CNN	HybridSN	SSAD	DBRNet	Pretext tasks + DBRNet
1	65.11	73.21	59.70	82.71	81.86	**94.12**
2	78.64	82.03	96.11	93.91	90.20	**97.68**
3	72.05	71.47	94.56	**96.72**	84.07	87.32
4	91.65	91.03	84.82	**97.70**	96.84	96.27
5	98.94	99.62	99.54	**100**	99.69	99.94
6	69.93	68.75	82.15	87.12	93.47	**97.92**
7	92.82	86.56	**100**	98.24	92.44	92.31
8	80.93	83.61	84.27	**91.58**	86.20	90.20
9	99.89	99.46	78.53	99.56	**100**	**100**
OA	78.52	79.15	86.76	92.02	89.70	**95.86**
AA	83.07	83.25	86.63	94.17	91.64	**95.08**
K	0.72	0.73	0.82	0.89	0.86	**0.94**

TABLE II
HSI CLASSIFICATION RESULTS (%) BY SELECTING 50 LABELED SAMPLES FROM EACH CLASS OF THE HOUSTON DATA

Class	SVM	3D-CNN	HybridSN	SSAD	DBRNet	Pretext tasks + DBRNet
1	96.83	97.66	93.92	94.75	**99.33**	95.33
2	90.28	99.50	**99.91**	99.66	**99.91**	98.75
3	98.60	99.07	99.69	**100**	**100**	**100**
4	92.37	96.06	92.46	99.24	94.13	**100**
5	98.07	99.74	**100**	**100**	**100**	**100**
6	96.36	98.18	**100**	98.54	**100**	**100**
7	80.29	81.11	91.62	96.30	95.81	**99.09**
8	52.84	82.49	86.34	25.62	**95.64**	87.43
9	75.37	76.20	88.15	89.51	92.49	**96.83**
10	82.15	93.62	94.39	8.66	97.19	**99.32**
11	69.87	87.93	97.63	91.98	95.69	**97.80**
12	60.60	89.94	96.61	94.25	90.87	**98.73**
13	38.42	73.74	95.94	**98.32**	94.51	94.98
14	98.67	96.82	**100**	**100**	99.73	**100**
15	98.52	99.34	**100**	**100**	**100**	**100**
OA	81.14	91.01	94.92	83.30	96.58	**97.61**
AA	81.95	91.43	95.78	86.45	97.02	**97.88**
K	0.79	0.90	0.94	0.81	0.96	**0.97**

process is shown in Fig. 1, (a) is the overall network structure, (b) is the residual structure of ResNet.

B. Spectral Domain Pretext Task

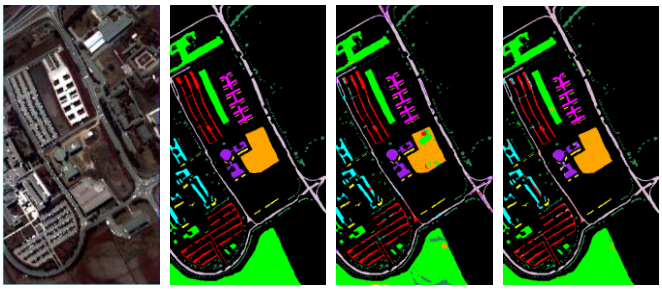
Classification using only spectral information is easily affected by different spectra of the same object and different objects of the similar spectra. When a small number of samples is used for training, it is difficult for the samples to cover all situations of the same category, so it is hard to obtain a good classification result. Therefore, spatial information should be introduced when classification is based on spectral information. Considering that most of the pixels of the same category in HSI classification are concentrated in the adjacent area, the HSI can simply be divided into several small blocks, with each small block being a category. In this way, although the pixels of the same category in some different areas are marked as different categories, and the pixels of different categories are marked as the same category, in general, the network has seen all the data, and the probability is high that the category is divided correctly when the size is divided appropriately. The task we designed is as follows.

First, the spectral data of each channel is normalized:

$$P_{norm}^{i} = \frac{P^{i} - P_{\min}^{i}}{P_{\max}^{i} - P_{\min}^{i}}, i = 1,2,3...n \tag{4}$$

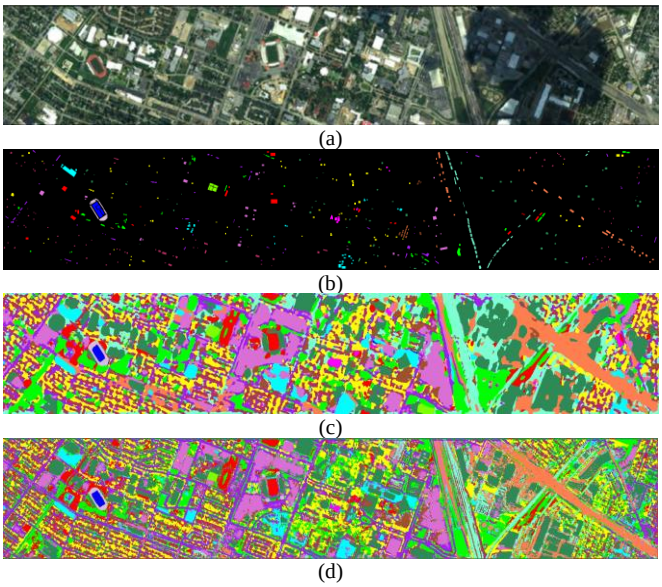
where P is the value of each element of the spectral data, n is the total number of channels in the spectrum, and i is the index of the channel.

Then, the HSI is divided into $r \times c$ rectangular blocks, with r rows and c columns, and the pixels in each rectangular block are regarded as belonging to the same category. In this way, a classification task is trained first. When a downstream task is trained, the parameter of the last FC layer of the network is replaced with the number of HSI data categories.



(a) (b) (c) (d)

Fig. 2. Classification maps of PaviaU. (a) False color image. (b) Ground truth. (c) SSAD. (d) Pretext tasks + DBRNet.



(a) (b) (c) (d)

Fig. 3. Classification maps of Houston. (a) False color image. (b) Ground truth. (c) HybridSN. (d) Pretext tasks + DBRNet.

C. Spatial Domain Pretext Task

When $n \times n$ image patches are used for classification, the neighborhood information of the pixels can be considered, but other categories of pixels in the image patch will affect the category of the current patch. Moreover, when n is large, there will be many pixels of different categories in an image patch expanded by pixels at the boundary, and the image patch obtained by the expansion of two adjacent pixels of different categories would have many of the same pixels, thereby affecting the classification accuracy. Therefore, in the pretext task, spectral information needs to be introduced. Considering that the feature extracted by the network from the image patch should be closer to the feature of the central pixel, other pixels, as a kind of auxiliary information, cannot affect the features extracted by the network too much. Therefore, the network should have the ability to see the image patch and know which pixel the image patch corresponds to. The task we designed is as follows.

First, channel normalization of the original data and

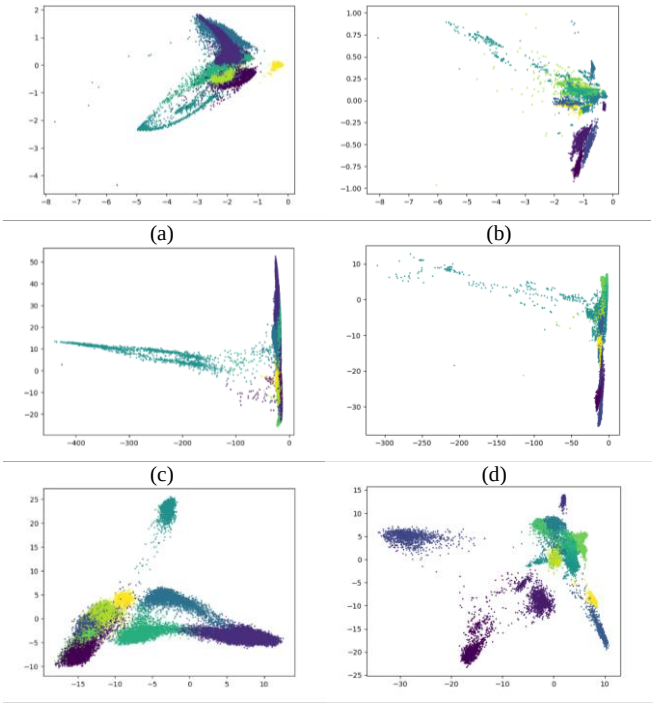


Fig. 4. The feature vector extracted by the OTVCA and network. (a), (c) and (e) are the features vector extracted on the PaviaU data set by OTVCA, DBRNet, and Pretext tasks + DBRNet respectively. (b), (d) and (f) are the features vector extracted on the Houston data set by OTVCA, DBRNet, and Pretext tasks + DBRNet respectively.

dimensionality reduction by principal component analysis (PCA) are performed, and then each pixel is expanded from the center to an $n \times n$ patch. The category of this patch is the category of the center point pixel. For the pixels at the edge, the center-symmetric pixels are used for expansion. The parameter of the last FC layer of the branch for extracting spatial features is set to the number of dimensions of the spectral vector, and the network performs training on the task of reconstructing the normalized spectral vector from the image patch. This training can use spectral information to supervise the spatial information. After the training is completed, the parameters of the final FC layer are set to the number of HSI data categories for the training of the classification task.

III. Experiments

A. Data Set and Training Details

In our research, two widely used public data sets are used to verify the effectiveness of the proposed method. They are the University of Pavia data set (PaviaU) and the Houston data set.

The PaviaU data is an area of Pavia city, Italy with a size of 610 × 340 pixels, which was taken by the German airborne Reflective Optical Spectral Imaging System (ROSIS-03). There are 9 categories and 103 bands in each pixel after removing the bands affected by noise.

The Houston data is a part of the University of Houston campus and the neighboring urban area with a size of 349 × 1905 pixels taken by the ITRES-CASI 1500 sensor. The data set has 15 categories. Each pixel has 144 bands.

TABLE III
The Classification Results (%) of using SVM to classify the extracted features

Method	PaviaU (20 samples per class)			Houston (50 samples per class)		
	OA	AA	K	OA	AA	K
OTVCA	63.45± 2.75	74.23± 1.92	0.54± 0.04	71.36± 1.08	72.42± 0.72	0.68± 0.02
DBRNet	52.80± 3.30	60.58± 0.84	0.42± 0.03	67.34± 1.60	69.17± 1.53	0.64± 0.02
Pretext tasks + DBRNet	**80.51± 0.90**	**77.43± 2.13**	**0.74± 0.01**	**81.57± 0.76**	**84.14± 0.43**	**0.79± 0.01**

All experiments were performed on an HP desktop computer equipped with an Intel Xeon(R) Silver4114 CPU, two GTX 2080 Ti graphics cards, and 128 GB memory. When image patches are taken from the PaviaU data, the data is first reduced by PCA to obtain 18 bands. The patch size is 17 × 17 pixels. The spectral self-supervised method divides the image into 40 × 20 categories. When image patches are taken from the Houston data, the data is first reduced by PCA to obtain 9 bands. The patch size is 17 × 17 pixels. The spectral self-supervised method divides the image into 17 × 85 categories. During the training of the pretext task, both the training set and the validation set use all the sample. With the Adam optimizer, the learning rate is 0.005, the batch size is 256, and the number of training rounds is 200. The Adam optimizer is also used for the training of the classification task; the learning rate is 0.01, the batch size is 64, the hyperparameter α is 0.2, the number of training rounds is 200, and the best performing model is selected as the final model. The transfer process consists of directly loading the network trained by the pretext tasks into the network of the classification task (except the last FC layer) and setting the parameters of the last FC layer of the classification task network to the number of HSI data categories.

B. Classification Results

In order to evaluate the method presented in this paper, we compare the algorithm with SVM [5], 3D-CNN [12], hybrid spectral CNN (HybridSN) [14], self-supervised learning method with adaptive distillation (SSAD) [19]. Among them, SVM is the representative of traditional classification algorithms, 3D-CNN is the basic algorithms for classification using deep learning, HybridSN is a space-spectrum joint classification algorithm using deep learning, and SSAD is advanced self-supervised HSI classification methods. The evaluation indicators s elected are the overall accuracy (OA), the average accuracy (AA), and Cohen's kappa coefficient (K). The results of SVM, 3D-CNN, HybridSN, SSAD and the method proposed in this paper are the average of 5 runs of randomly selected samples.

For the PaviaU data, 20 samples are randomly selected from each sample category for training. For the Houston data, 50 samples are randomly selected from each sample category for training. The results are shown in Tables I and II. The best value is shown in bold. The classification maps are shown in Fig. 2. and Fig. 3. We show the best maps in the comparison method and our maps. From the data in the table, we can see that the

deep learning method using a combination of spatial and spectral information is better than the traditional method and that the method using self-supervised learning generally has better classification results because it can extract additional information from the data. However, the extra information extracted may interfere with the classification task. As shown in the 8th and 10th categories of SSAD in Table II. Compared with other highest methods, our method is more effective. For PaviaU data, the OA increased from 92.02% to 95.86%, the AA increased from 94.17% to 95.08%, and K increased from 0.89 to 0.94. For the Houston data, the OA increased from 94.92 % to 97.61%, the AA increased from 95.78% to 97.88%, and K increased from 0.94 to 0.97. The above results show that our method is superior to other methods and that it can use a smaller number of samples to achieve a more accurate classification.

C. Validation of the Self-Supervised Learning Method

It can be seen from the data in the Tables I and II that training the pretext task first can effectively improve DBRNet classification performance. In addition, the feature vectors extracted by the network are reduced by PCA to draw a scatter plot (Fig. 3). It can be clearly seen from the figure that the features extracted by the network using the pretext task are more distinguishable in different categories.

In order to quantitatively represent the gap between with or without pretext task, we use the network to extract the features of the labeled samples, and use PCA to reduce the features to 2 dimensions. In order to prove the effectiveness, we also compared with an advanced feature extraction method called Total Variation Component Analysis (OTVCA) [24]. The dimension of each feature is 2. Then use a simple classifier (SVM) to classify using only these features. The result is shown in Table III.

The classification results show that the feature vectors extracted by the network are easier to be classified by an SVM after using the pretext task, which indicates that the difference between different samples becomes larger in the feature space, which is useful for classification tasks.

IV. Conclusion

In this paper, we propose a dual-branch network structure, which can effectively extract spatial and spectral information. In addition, for such a structure, we design a pretext task for each branch, which makes the network more suitable for training with a small number of samples. Our method does not perform any data enhancement or data expansion and uses only a small number of samples for training. Excellent classification results were obtained for both data sets, which confirmed the superiority of our method.

References

[1] F. van der Meer, "Analysis of spectral absorption features in hyperspec-tral imagery," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 5, no. 1, pp. 55-68, Feb. 2004.

[2] G. Abbate, L. Fiumi, C. De Lorenzo, and R. Vintila, "Evaluation of remote sensing data for urban planning. Applicative examples by means of multispectral and hyperspectral data," in *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, 2003, pp. 201-205.

[3] U. Heiden, K. Segl, S. Roessner, and H. Kaufmann, "Ecological evaluation of urban biotope types using airborne hyperspectral HyMap data," in *2nd GRSS/ISPRS Joint Workshop on Remote Sensing and Data Fusion over Urban Areas*, 2003, pp. 18-22.

[4] J. M. Yang, P. T. Yu, and B. C. Kuo, "A nonparametric feature extraction and its application to nearest neighbor classification for hyperspectral image data," *IEEE Trans. Geosci. Remote Sens.*, vol. 48, no. 3, pp. 1279-1293, Mar. 2010.

[5] F. Melgani and L. Bruzzone, "Classification of hyperspectral remote sensing images with support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 8, pp. 1778-1790, Aug. 2004.

[6] F. Luo, L. Zhang, B. Du and L. Zhang, "Dimensionality Reduction With Enhanced Hybrid-Graph Discriminant Learning for Hyperspectral Image Classification," *IEEE Trans. Geosci. Remote Sens*., vol. 58, no. 8, pp. 5336-5353, Aug. 2020.

[7] C. Szegedy et al., "Going deeper with convolutions," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2015, pp. 1-9.

[8] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards real-time object detection with region proposal networks," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 39, no. 6, pp. 1137-1149, Jun. 2017.

[9] D. Tran, H. Wang, L. Torresani, J. Ray, Y. LeCun, and M. Paluri, "A closer look at spatiotemporal convolutions for action recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2018, pp. 6450-6459.

[10] L. Fang, W. Zhao, N. He and J. Zhu, "Multiscale CNNs Ensemble Based Self-Learning for Hyperspectral Image Classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 9, pp. 1593-1597, Sept. 2020.

[11] Y. Jiang, Y. Li and H. Zhang, "Hyperspectral Image Classification Based on 3-D Separable ResNet and Transfer Learning," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 12, pp. 1949-1953, Dec. 2019.

[12] Y. Chen, H. Jiang, C. Li, X. Jia, and P. Ghamisi, "Deep feature extraction and classification of hyperspectral images based on convolutional neural networks," *IEEE Trans. Geosci. Remote Sens*., vol. 54, no. 10, pp. 6232-6251, Oct. 2016.

[13] Y. Xu, L. Zhang, B. Du, and F. Zhang, "Spectral-spatial unified networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens*., vol. 56, no. 10, pp. 5893-5909, Oct. 2018.

[14] S. K. Roy, G. Krishna, S. R. Dubey and B. B. Chaudhuri, "HybridSN: Exploring 3-D–2-D CNN Feature Hierarchy for Hyperspectral Image Classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 2, pp. 277-281, Feb. 2020.

[15] C. Doersch, A. Gupta, and A. A. Efros, "Unsupervised visual rep-resentation learning by context prediction," in *Proc. IEEE Int. Conf. Comput. Vis. (ICCV)*, Dec. 2015, pp. 1422-1430.

[16] R. Zhang, P. Isola, and A. A. Efros, "Colorful image colorization," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 649-666.

[17] M. Noroozi and P. Favaro, "Unsupervised learning of visual representions by solving jigsaw puzzles," in *Proc. Eur. Conf. Comput. Vis. (ECCV)*, 2016, pp. 69-84.

[18] D. Pathak, P. Krahenbuhl, J. Donahue, T. Darrell, and A. A. Efros, "Context encoders: Feature learning by inpainting," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 2536-2544.

[19] J. Yue, L. Fang, H. Rahmani, and P. Ghamisi, "Self-supervised learning with adaptive distillation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*.

[20] Y. Wang et al., "Self-supervised low-rank representation (SSLRR) for hyperspectral image Classification," *IEEE Trans. Geosci. Remote Sens*., vol. 56, no. 10, pp. 5658-5672, Oct. 2018.

[21] Y. Wang et al., "Self-supervised feature learning with CRF embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens*., vol. 57, no. 5, pp. 2628-2642, May 2019.

[22] Y. Li, L. Zhang, W. Wei and Y. Zhang, "Deep Self-Supervised Learning for Few-Shot Hyperspectral Image Classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2020, pp. 501-504.

[23] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770-778.

[24] B. Rasti, M. O. Ulfarsson and J. R. Sveinsson, "Hyperspectral Feature Extraction Using Total Variation Component Analysis," *IEEE Trans. Geosci. Remote Sens*., vol. 54, no. 12, pp. 6976-6985, Dec. 2016