Locally Homogeneous Covariance Matrix Representation for Hyperspectral Image Classification

Xinyu Zhang, Yantao Wei, Huang Yao, Zhijing Ye, Yicong Zhou, *Senior Member, IEEE*, and Yue Zhao

***Abstract*—Combining spectral and spatial information has been proven to be an effective way for hyperspectral image (HSI) classification. However, making full use of spectral–spatial information of HSI still remains an open problem, especially when only a small number of labeled samples are available. In this article, a new spectral–spatial feature extraction method called locally homogeneous covariance matrix representation (CMR) is proposed for the fusion of spectral and spatial information. Specially, to make use of neighborhood homogeneity of land covers, original HSI is first segmented into many superpixels using modified entropy rate superpixel segmentation. Then, to acquire the most similar pixels, we propose to construct neighborhoods of each pixel from the overlapping areas between the corresponding superpixels and the sliding window centered on it. Subsequently, CMRs of different pixels can be obtained. In the classification stage, we fed the obtained CMRs into SVM with Log-Euclidean-based kernel for classification. Compared to the traditional approach that utilizes neighboring information only within a fixed window, the proposed local homogeneity strategy can absorb more discriminative spectral–spatial features. Experimental results from a series of available HSI datasets show that our proposed method is superior to several state-of-the-art methods, especially when the training set is very limited.**

***Index Terms*—Covariance matrix (CM), entropy rate superpixel (ERS) segmentation, feature extraction, hyperspectral image (HSI) classification.**

Manuscript received April 30, 2021; revised July 4, 2021 and August 4, 2021; accepted August 25, 2021. Date of publication September 8, 2021; date of current version September 30, 2021. This work was supported in part by the National Natural Science Foundation of China under Grant 61502195, in part by the Natural Science Foundation of Hubei Province under Grant 2018CFB691, in part by the Self-Determined Research Funds of CCNU From the Colleges' Basic Research and Operation of MOE under Grant CCNU20TD005, in part by the Science and Technology Development Fund, Macau SAR (File 189/2017/A3), and in part by the Research Committee at University of Macau under Grant MYRG2016-00123-FST and Grant MYRG2018-00136-FST. *(Corresponding author: Yantao Wei.)*

Xinyu Zhang, Yantao Wei, and Huang Yao are with the Hubei Research Center for Educational Informationization, Faulty of Artificial Intelligence in Education, Central China Normal University, Wuhan 430079, China (e-mail: zxy.one@mails.ccnu.edu.cn; yantaowei@mail.ccnu.edu.cn; yaohuang@mail.ccnu.edu.cn).

Zhijing Ye is with the School of Science, Wuhan University of Technology, Wuhan 430070, China (e-mail: xkinghust@163.com).

Yicong Zhou is with the Faculty of Science and Technology, University of Macau, Macau 999078, China (e-mail: yicongzhou@um.edu.mo).

Yue Zhao is with the School of Computer Science and Information Engineering, Hubei University, Wuhan 430062, China (e-mail: zhaoyhu@hubu.edu.cn).

Digital Object Identifier 10.1109/JSTARS.2021.3110779

I. Introduction

DIFFERENT from ordinary RGB images, hyperspectral images (HSIs) usually contain hundreds of spectral channels from ultraviolet to infrared, which provides valuable information for detailed material analysis [1], [2]. Therefore, HSIs have been widely applied in many fields, such as environmental monitoring, agriculture [3], medical diagnosis, and target detection [1], [4], [5]. In the last few decades, HSI classification technology, which assigns a unique class label to each pixel, has attracted great attention in the field of remote sensing [6]. However, it is very challenging to obtain satisfactory results due to the limited labeled samples and existing noise [2], [7]–[9].

In the past few years, a large number of HSI classification methods have been designed [10], [11]. Due to the presence of a large number of bands in the HSI data, many spectral dimensionality reduction-based HSI classification methods have been proposed, such as the methods based on principal component analysis (PCA) [12] and maximum noise fraction (MNF) [13]. In [14], an unsupervised classification framework based on robust manifold matrix factorization, which can address the high dimensionality of HSI, has been proposed. However, since the high intraclass variability and interclass similarity in HSI data [7], resulting from the influence of variation of light, climate, and other uncontrolled factors, classification performances produced by these spectral-based methods are usually unsatisfactory [15].

To tackle this issue, a variety of spatial–spectral frameworks, which consider both spectral and spatial information, have been widely investigated [16]–[20]. For example, extended morphological profiles are put forward to adaptively extract spatial characteristics in HSI [17]. Kang *et al.* [21] proposed an edge-preserving filtering-based framework to improve the classification performance obtained by the pixelwise classifier. In [22], superpixel-based classification is adopted to utilize spatial–spectral features via multiple kernels (SCMK). Meanwhile, Li *et al.* [23] proposed a method based on local binary patterns (LBPs) to exploit the contextual information of HSI. To overcome the oversmoothing phenomenon caused by the Gabor filtering, a method called spectral–spatial range Gabor filtering has also been developed in [24]. Huang *et al.* [25] proposed a method called local linear spatial–spectral probabilistic distribution by

constructing a multiclass probability map to make full use of spatial correlation in HSI.

In addition, owing to the successful application in computer vision over the past few years, deep-learning-based methods have also been employed in HSI classification [26]–[30]. For example, a regularized deep feature extraction and classification method is proposed in [29], which develops a 3-D convolutional neural network (CNN) to explore spatial–spectral features. However, the 3-D convolution model has a high time cost and computational complexity [31]. To address this problem, Roy *et al.* [32] have achieved competitive performance by the cooperation between 2-D CNN and 3-D CNN. Although these methods based on deep learning can achieve good classification performance, they usually require a large number of samples for training to avoid overfitting problem. But acquiring labeled samples is very time-consuming in the field of remote sensing [33], [34]. Although some improvement measures have been put forward, i.e., random patches networks [35] and deep metric learning based framework [36], it is still an open matter.

Recently, the covariance matrix representation (CMR) has been proven to be an effective feature representation method that can make use of the correlation between different features [37]–[39]. Each nondiagonal element in the covariance matrix (CM) represents the covariance between the two different features. Taking the advantages of CMR, Fang *et al.* [37] proposed a spatial–spectral feature extraction method called local CMR (LCMR). But there is much noise in the classification map. Although Zhang *et al.* [40] proposed a method that combines multiscale adaptive weighted filtering with LCMR to obtain more smooth classification maps, the local neighbor selection is still an important problem to be solved. The main idea of LCMR is to obtain the discriminative features of different pixels by computing CMR from the local neighborhood determined by a sliding window. However, the pixels in the same sliding window may not belong to the same class, which leads to noisy classification maps [41]. Therefore, this article proposes a new representation method called locally homogeneous CMR (LHCMR) method to make full use of local-homogeneity-based spatial–spectral information to enhance the classification performance. Specially, we select the most similar pixels in the overlapping area between the superpixel and the sliding window, where it is called locally homogeneous area. This article extends our previous conference paper [40], which just contains very preliminary results. Compared to it, this article first adopts our modified entropy rate superpixel (ERS) segmentation to segment HSIs into many homogeneous regions based on spatial–spectral information and then obtains CMRs based on the local homogeneity strategy, which provides a new solution for similar neighboring pixels selecting.

The remainder of this article is organized as follows. In Section II, some related works are briefly reviewed. Section III presents the proposed LHCMR-based method for HSI classification. Section IV presents experimental results using four well-known HSI datasets and shows the results of the comparison with other state-of-the-art methods. Finally, Section V concludes this article.

II. Related Work

A. ERS Segmentation

ERS segmentation is one of the commonly used preprocessing methods for HSIs [22], [42]. It is a graph-based segmentation algorithm with high efficiency [43]. Specifically, the ERS first constructs a graph $G = (V, E)$ on the image to be segmented. $\mathbf{V}$ is the vertex set corresponding to pixels in the image and $\mathbf{E}$ is the edge set consists of the pairwise similarities between adjacent pixels. The weight $W_{i,j}$ between vertexes v_i and v_j is defined using Gaussian kernel as

$$W_{i,j} = \exp\left(\frac{d(v_i, v_j)^2}{2\sigma^2}\right) \tag{1}$$

where $d(v_i, v_j)$ is the spatial distance. Then, the graph is segmented into some smaller connected subgraphs by selecting a subset of edges $\mathbf{A} \subseteq \mathbf{E}$. To balance compactness and similarities of cluster sizes, an entropy rate term $\mathbf{H}(\mathbf{A})$ and a balancing term $\mathbf{B}(\mathbf{A})$ are incorporated into the objective function

$$\max_{\mathbf{A}}\{\mathbf{H}(\mathbf{A}) + \alpha\mathbf{B}(\mathbf{A})\}, \qquad \mathbf{s.t.} \quad \mathbf{A} \subseteq \mathbf{E}. \tag{2}$$

where α is used to control the contributions of the entropy rate term $\mathbf{H}(\mathbf{A})$ and the balancing term $\mathbf{B}(\mathbf{A})$. Then, a 2-D superpixel map can be generated. The optimization problem in (2) can be effectively solved by a greedy algorithm [42].

B. CM Representation

Suppose an image is defined as $\mathbf{I}$, and a given region $\mathbf{R} \subseteq \mathbf{I}$ has K pixels. Let $\{\mathbf{x_i}\}_{i=1}^{K}$ denote the feature of the ith pixel within $\mathbf{R}$ and $\overline{\mathbf{x}}$ denote the mean vector of them. The CMR of region $\mathbf{R}$ can be then obtained as follows:

$$\mathbf{C}_R = \frac{1}{K-1}\sum_{i=1}^{K}(\mathbf{x}_i - \overline{\mathbf{x}})(\mathbf{x}_i - \overline{\mathbf{x}})^{\mathbf{T}}. \tag{3}$$

In the CM, each nondiagonal entry represents the covariance between two different features. Since the obtained CM is a symmetric positive definite matrix, which is located on a nonlinear Riemannian manifold space. So, a manifold-based distance metric, the Logarithmic Euclidean distance metric, can be used to measure the CM. Specifically, for two given CMs $\mathbf{C_1}$ and $\mathbf{C_2}$, the Log-Euclidean distance between them is defined by the following formula:

$$\mathbf{d}(\mathbf{C_1}, \mathbf{C_2}) = \| \operatorname{logm}(\mathbf{C_1}) - \operatorname{logm}(\mathbf{C_2}) \|_F \tag{4}$$

where $\| \cdot \|_F$ is the Frobenius norm and $\operatorname{logm}(\cdot)$ denotes the ordinary matrix logarithm operator [38]. For a given symmetric positive definite matrix $\mathbf{C}$, let $\mathbf{C} = \mathbf{U}\mathbf{\Sigma}\mathbf{U}^{\mathbf{T}}$ represent its eigendecomposition, then its logm can be computed by

$$\mathbf{logm}(\mathbf{C}) = \mathbf{U}\log(\mathbf{\Sigma})\mathbf{U}^{\mathbf{T}}. \tag{5}$$

III. Proposed Method

To overcome the problem with LCMR based on a fixed sliding window, the LHCMR-based spectral–spatial feature extraction method is described in this section, where its framework is shown

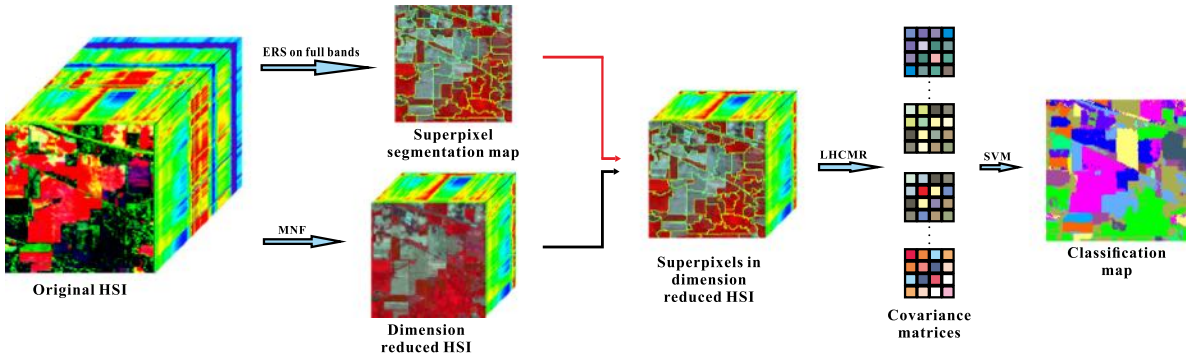

Fig. 1. Flowchart of proposed LHCMR method.

in Fig. 1. Specially, we first introduce the superpixel segmentation to reveal the homogeneity of land covers with different sizes. Then, the CM is computed based on the most similar pixels selected from the overlapping area of the superpixel and sliding window. After that, we feed the obtained CM sets into SVM with Log-Euclidean-based kernel for final classification. The proposed method takes the advantage of superpixel segmentation and obtains CM within the locally homogeneous area, which is conducive to enhance classification performance.

A. Superpixel Segmentation

Usually, in order to apply superpixel segmentation on HSI, a common way is to simply apply PCA to extract the first or three principle components of HSI. However, HSIs have the advantage of high spectral resolution, the PCA step before segmentation may make it lost. Thus, we carry out the ERS segmentation on full bands to more accurately capture the structure of HSI. In (1), both the vertexes $\mathbf{v}_i$ and $\mathbf{v}_j$ are n-dimensional spectral vectors in our modified version. Then, the weight $W_{i,j}$ is calculated as

$$\mathbf{W_{i,j}} = \exp\left(\frac{\mathbf{d_n}(\mathbf{v_i}, \mathbf{v_j})^2}{\mathbf{2}\sigma^2}\right) \tag{6}$$

where $\mathbf{d}_n(v_i, v_j)$ can be defined as

$$\mathbf{d_n}(\mathbf{v_i}, \mathbf{v_j}) = \sqrt{(\mathbf{v_i} - \mathbf{v_j})^{\mathrm{T}}(\mathbf{v_i} - \mathbf{v_j})}. \tag{7}$$

Then, we perform ERS on the original HSI I_{hsi} to produce superpixel map. Let N denote the number of superpixels, the superpixel segmentation results can be defined as

$$I_{\text{hsi}} = \bigcup_{s}^{N} \eta_s \qquad \text{s.t.} \quad \eta_s \cap \eta_j = \oslash, (s \neq j) \tag{8}$$

where $\boldsymbol{\eta}_s$ is the $\mathbf{s}$th superpixel in the superpixel map.

Through the segmentation algorithm, HSI can be divided into several meaningful subregions. In this way, we can effectively extract spatial information in HSI, and then combine the spectral features for further representation.

B. Construction of LHCMR

After that, LHCMR will be derived based on the results of superpixel segmentation. We can calculate the similarity between each pixel and its neighbors, and then select the most similar $K-1$ pixels to construct the CM.

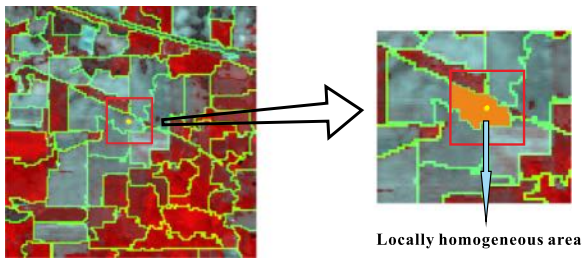

Fig. 2. Locally homogeneous area: the overlapping area of superpixel and the window.

However, constructing CMs directly from superpixels may mix pixels belonging to different classes. One of the reasons is that the pixels in superpixel may not be uniform as expected [44]. For example, the pixels in the same superpixel may be far away from each other in some cases, this is inconsistent with the local assumption. Therefore, we need to make a constraint to make full use of local spatial information and ensure local homogeneity.

To this end, we introduce a sliding window based on superpixel. Although directly selecting neighbors in a fixed window will cause some noise, constructing neighbors in the overlapping area between the superpixel and the sliding window can provide a solution to the aforementioned problem. Fig. 2 shows the overlapping area, where we will select the most similar pixels. Assume that there are m pixels in the superpixel η_s. For each pixel $p_i \in \eta_s, (i = 1, 2, \ldots, m)$, the overlapping area of the $T * T$ window W centered on the given pixel and the superpixel η_s is determined by

$$D_{ol} = \{p_i | p_i \in \eta_s \cap p_i \in W, i \leq \min\{m, T^2\}\}. \tag{9}$$

Subsequently, for each pixel, we employ the cosine-distance-based KNN to select the $K-1$ most similar pixels. Suppose there are M pixels in the overlapping area $\mathbf{D_{ol}}$, the cosine distance between the pixel ρ_1 and its neighboring pixels can be calculated by

$$\cos(\rho_1, \rho_i) = \frac{\langle \rho_1, \rho_i \rangle}{\| \rho_1 \|_2 \cdot \| \rho_i \|_2}, \quad \mathbf{i = 1, 2, \ldots, M} \tag{10}$$

TABLE I
DETAILS OF INDIAN PINES, PAVIA UNIVERSITY, SALINAS, AND PAVIA CENTER DATASET

	Indian Pines			Pavia University			Salinas			Pavia Center	
Class	Name	Samples	Class	Name	Samples	Class	Name	Samples	Class	Name	Samples
1	Alfalfa	46	1	Asphalt	6631	1	Brocoli_green_weeds_1	2009	1	Water	65278
2	Corn-notill	1428	2	Meadows	18649	2	Brocoli_green_weeds_2	3726	2	Trees	6508
3	Corn-mintill	830	3	Gravel	2099	3	Fallow	1976	3	Meadows	2905
4	Corn	237	4	Trees	3064	4	Fallow_rough_plow	1394	4	Self-Blocking Bricks	2140
5	Grass-pasture	483	5	Painted metal sheets	1345	5	Fallow_smooth	2678	5	Bare Soil	6549
6	Grass-trees	730	6	Bare Soil	5029	6	Stubble	3959	6	Asphalt	7585
7	Grass-pasture-mowed	28	7	Bitumen	1330	7	Celery	3579	7	Bitumen	7287
8	Hay-windrowed	478	8	Self-Blocking Bricks	3682	8	Grapes_untrained	11271	8	Tiles	3122
9	Oats	20	9	Shadows	947	9	Soil_vinyard_develop	6203	9	Shadows	2165
10	Soybean-notill	972				10	Corn_senesced_green_weeds	3278			
11	Soybean-mintill	2455				11	Lettuce_romaine_4wk	1068			
12	Soybean-clean	593				12	Lettuce_romaine_5wk	1927			
13	Wheat	205				13	Lettuce_romaine_6wk	916			
14	Woods	1265				14	Lettuce_romaine_7wk	1070			
15	Buildings-Grass-Trees-Drives	386				15	Vinyard_untrained	7268			
16	Stone-Steel-Towers	93				16	Vinyard_vertical_trellis	1807			
Total		10249	Total		42776	Total		54129	Total		103539

where $\langle\cdot\rangle$ and $\|\cdot\|$ represent the inner product and Frobenius norm, respectively. Based on the position index of these pixels, we can associate them with HSI.

To reduce noise and computational complexity, we first apply the MNF method for dimension reduction. After obtaining the neighbor pixels in the reduced-dimensional HSI, we can derive the LHCMR of K pixels according to (3). In this way, the local homogeneity-based spatial–spectral information of HSI will be naturally integrated into the LHCMR. Each off-diagonal element in the matrix represents the correlation between different spectral bands of this type of material, which helps improve the classification performance.

C. Classification

Finally, a series of CMRs obtained in the locally homogeneous area are fed into the SVM for training and classification. In this article, a commonly used kernel function called Logarithm-Euclidean kernel function [45] is adopted to map the CM from Riemannian manifold space to Euclidean space, it can be defined by

$$k_{\text{logm}}(C_1, C_2) = \text{trace}[\text{logm}(C_1) \cdot \text{logm}(C_2)] \qquad (11)$$

where C_1 and C_2 are two given CMs.

IV. EXPERIMENTAL RESULTS AND DISCUSSION

To demonstrate the effectiveness of the proposed LHCMR-based method, extensive experiments have been performed on four well-known HSI datasets. The classification performance is objectively evaluated by three widely used quality indicators: overall accuracy (OA), average accuracy (AA), and Kappa coefficient k.

A. Datasets

The first experimental dataset is Indian Pines. The number of spectral bands has been reduced to 200 after 24 water absorption bands are removed and each band has the size of 145×145 pixels. The ground truth available is designated into 16 classes. The second HSI data are the Pavia University image and were acquired by ROSIS sensor over the campus at the University of Pavia, Italy. This scene is of size $610 \times 340 \times 103$ after the noise-corrupted bands were removed. The ground-truth is composed of 9 land cover classes. The third HSI used in the experiment is Salinas, which was also gathered by AVIRIS sensor over the Salinas Valley, California. This image has a size of $512 \times 217 \times 204$ after 20 water absorption bands were discarded. And this dataset contains 16 classes. The fourth HSI data are the Pavia Center image, which with the size of 1096×492 and 102 spectral bands. The ground-truth map covers 9 classes. The details of all the datasets are tabulated in Table I.

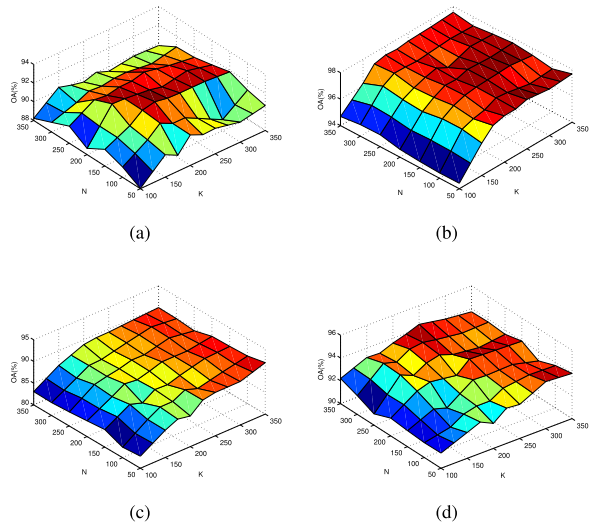

Fig. 3. Analysis of parameters *N* and *K* in our proposed LHCMR method on the (a) Indian Pines image, (b) Pavia University image, (c) Salinas image, and (d) Pavia Center image.

B. Parameter Analysis

The effect of the number of superpixels *N*, the size of sliding window *T*, and the number of local neighboring pixels *K* has been seriously investigated in this section. First, we will discuss the effect of *N* (from 50 to 350 with step 50) and *K* (from 100 to 350) on the performance of the proposed method (*T* is fixed). Fig. 3 shows the OAs achieved by different *N* and

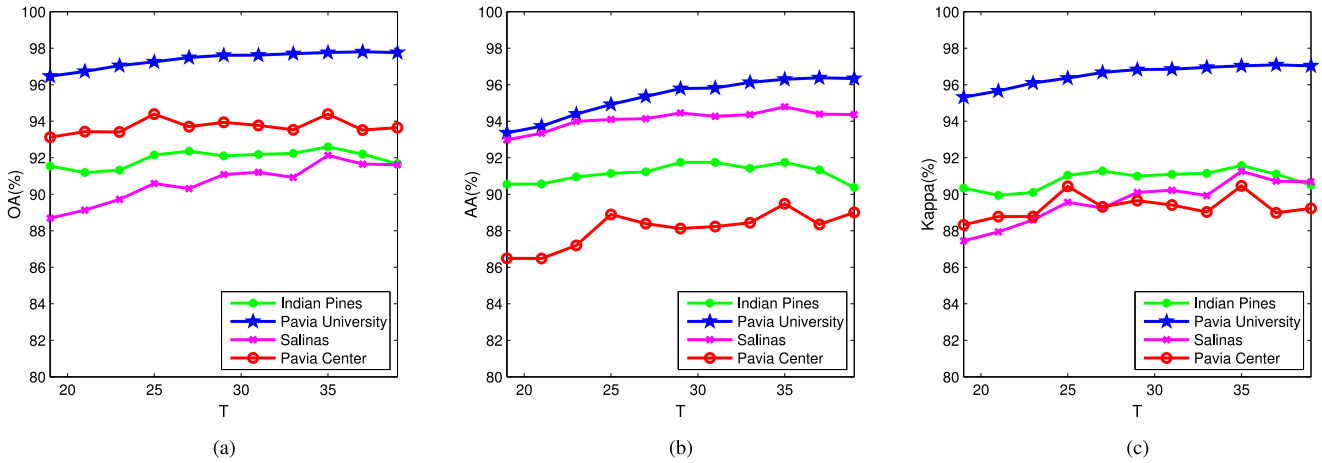

Fig. 4. Analysis of the parameters T in our proposed LHCMR method. (a) OAs. (b) AAs. (c) Kappa.

TABLE II
PARAMETERS OF THE PROPOSED LHCMR METHOD ON FOUR DATASETS

Dataset	N	K	T
Indian Pines	150	200	
Pavia University	50	250	35
Salinas	50	350	
Pavia Center	350	250	

K on four datasets. For the number of superpixels N, we can observe that the classification performance is sensitive to the number of superpixels on Indian Pines image. It is because that a small number of superpixels is not sufficient to capture the image structure, whereas a large number of superpixels leads to oversegmentation for Indian Pines with small size. Furthermore, the number of superpixels partly reflects the details of the image. Thus, it can be seen from Fig. 3 that competitive results can be obtained when dividing a small number of superpixels on the Salinas dataset, whereas a large number of superpixels need to be divided for Pavia Center dataset. For the number of local neighboring pixels K, it is obvious that the OAs on all datasets are particularly low when the number of local neighboring pixels is small. The reason for this is that a small number of local neighboring pixels may not be enough to extract sufficient spatial information. It can be also observed that, with the increase of K, the classification accuracies of the proposed method have a trend of rising on Pavia University, Salinas, and Pavia Center image, whereas the OAs first improve and then decrease on the Indian Pines image. This is mainly because that the size of the Indian Pines image is smaller, a large number of local neighboring pixels may mix in more dissimilar pixels, which will hamper the classification performance. As a result, the parameters K and N on the four datasets are detailed in Table II, respectively.

Then, parameter T, which is used to ensure local homogeneity, is also tested with parameters N and K fixed. In our experiment, T ranges from 19 to 39 with step 2. The effects of different T on accuracies obtained on four datasets are shown in Fig. 4. As can be seen, OAs, AAs, and Kappa show an upward trend when T grows from 19 to 35 on all the images. In addition, the best performance is achieved when T is 35 and the accuracies will decrease when T is larger than 35 on Indian Pines, Salinas, and Pavia Center images. On Pavia University image, the best results are obtained when T is 37. In the proposed method, the sliding window plays a role in ensuring spatial similarity of superpixels. This means that small size will lead to few neighboring pixels in the overlapping area, which causes the extracted features to be not representative. However, when the window is too large (larger than the size of the superpixel block), it will become meaningless. As a result, the value of T is set to 35 for all datasets.

TABLE III
AVERAGE RESULTS (%) OF TEN EXPERIMENTS AT THE INDIAN PINES DATASET WITH TEN TRAINING SAMPLES PER CLASS

Class	LBP	SCMK	deepNRD	RPNet	LCMR	S-DMM	HybridSN	LHCMR
1	**100.00**	97.50	87.22	97.22	99.44	94.72	83.33	95.28
2	65.85	75.50	35.79	52.81	76.78	38.86	32.72	**77.11**
3	76.40	72.65	47.61	70.68	69.41	47.48	56.22	**78.16**
4	**97.00**	94.36	92.25	90.00	94.41	88.50	94.27	87.71
5	82.98	80.42	45.67	86.26	**90.66**	77.36	85.84	90.53
6	89.11	95.89	72.31	91.75	92.38	94.60	98.75	**99.94**
7	**100.00**	97.78	**100.00**	93.89	**100.00**	99.44	**100.00**	**100.00**
8	99.79	98.59	92.78	91.56	99.76	94.38	**100.00**	98.65
9	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**
10	83.43	62.12	35.41	71.10	81.48	69.53	60.19	**86.19**
11	70.03	71.80	54.21	66.48	65.89	47.80	50.10	**72.54**
12	80.74	73.26	46.93	70.94	**85.78**	41.46	64.15	84.00
13	98.26	99.95	94.15	96.36	99.44	99.85	**100.00**	99.08
14	94.06	83.15	80.21	85.35	**97.63**	92.06	78.88	97.57
15	**97.55**	92.26	73.46	88.94	**94.76**	67.50	81.91	93.22
16	98.43	98.31	92.05	89.52	99.40	**100.00**	**100.00**	98.55
OA	80.80	78.53	57.69	74.17	81.77	64.01	64.78	**84.72**
AA	89.60	87.10	71.88	83.93	90.45	78.35	80.40	**91.16**
k	78.42	75.71	52.04	70.92	79.45	59.77	61.03	**82.70**

C. Comparison With Different Methods

In this article, the performance of our proposed LHCMR algorithm is compared with seven state-of-the-art methods. Traditional spatial–spectral methods include the LBP-based method [23], superpixel-based classification method via multiple kernels (SCMK) [22], and LCMR [37]. In addition, some advanced deep-learning-based methods, random patches network (RPNet) [35], an improved 2-D CNN-based approach called deepNRD [30], deep metric learning based model (S-DMM) [36], and hybrid spectral CNN (HybridSN) [32] are also included. For the deepNRD, its parameters are set to the best, and the parameters of other methods are the default parameters.

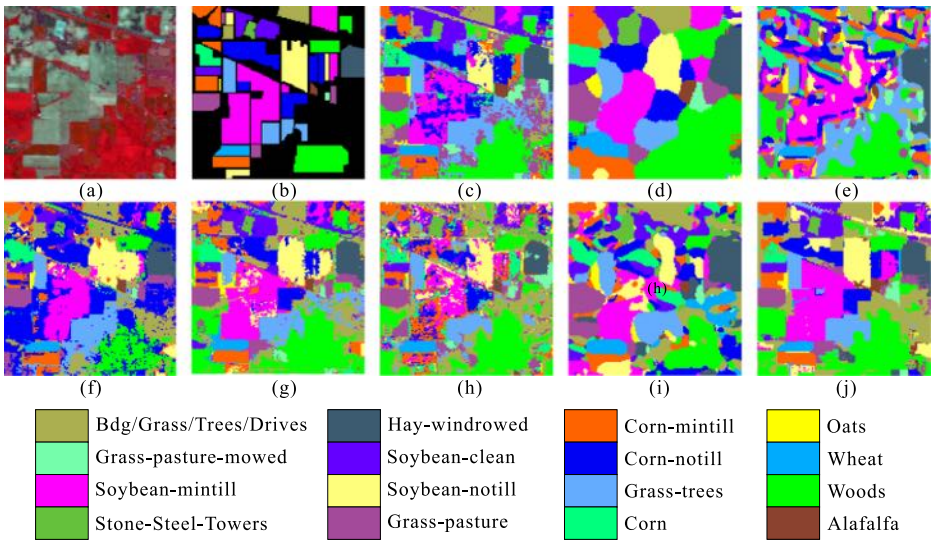

Fig. 5. Classification maps on the Indian Pines dataset. (a) Three-band color composite of the HSI. (b) Ground-truth. (c) SCMK. (d) LBP. (e) deepNRD. (f) RPNet. (g) LCMR. (h) S-DMM. (i) HybridSN. (j) Proposed LHCMR.

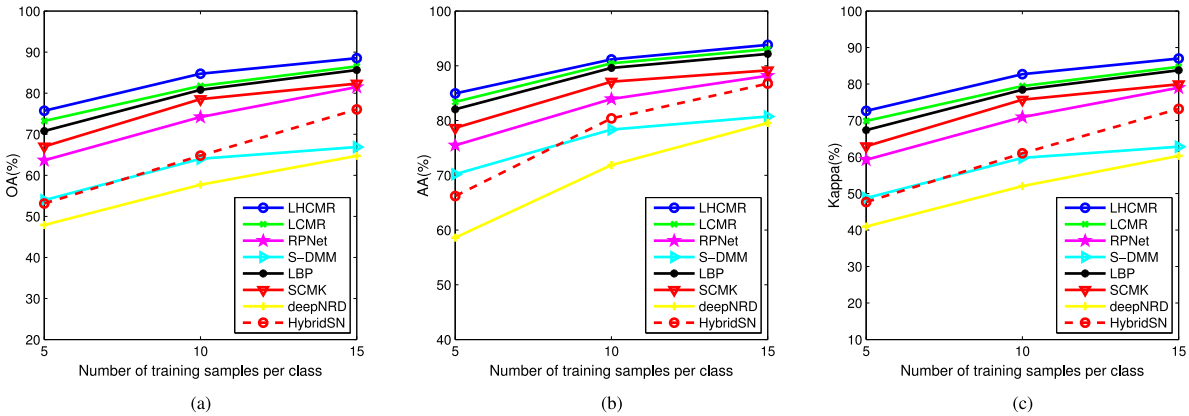

Fig. 6. Classification results of different numbers of training samples on the Indian Pines dataset. (a) OAs. (b) AAs. (c) Kappa.

The first experiment was conducted on the Indian Pines dataset. In this experiment, ten training samples per class are randomly selected, and the remaining are used as test samples. The average classification results for various methods over ten trials with different randomly selected training samples are reported in Table III. The best results are highlighted in bold font. As can be observed, deep-learning-based methods achieved lower accuracies, mainly because these methods suffer from overfitting problem when very few samples are chosen for training. Among them, RPNet achieves better results because its convolution kernels are randomly selected from feature images without any training. However, the proposed LHCMR performs the best among other comparison algorithms in terms of OA, AA, and Kappa coefficient. In addition, the proposed method significantly enhances the classification performance of LCMR. This demonstrates that the local homogeneity strategy can obtain more similar neighboring pixels and extract more effective spatial–spectral information than just using a fixed window.

Fig. 5 illustrates visual classification results obtained by different methods on Indian Pines image. It can be observed that when using limited training samples, the classification maps generated by the SCMK, RPNet, LCMR, and S-DMM have much noise, whereas the LBP, deepNRD, and HybridSN lead to oversmoothed maps. However, the proposed LHCMR can not only preserve the structure of HSI, but also can classify pixels more accurately. The reason for this is that the superpixel segmentation strategy can naturally capture the structure of HSI. Furthermore, we select the representative neighboring pixels from locally homogeneous areas.

In addition, we have also investigated the effect of different numbers of training samples on the performance of the proposed LHCMR and other methods. Fig. 6 shows the OAs, AAs, and Kappa obtained by different methods. We randomly select the numbers of training samples from 5 to 15 in a step of 5 from each class (since there are only 20 samples of a class in Indian Pines image). As can be seen, the proposed

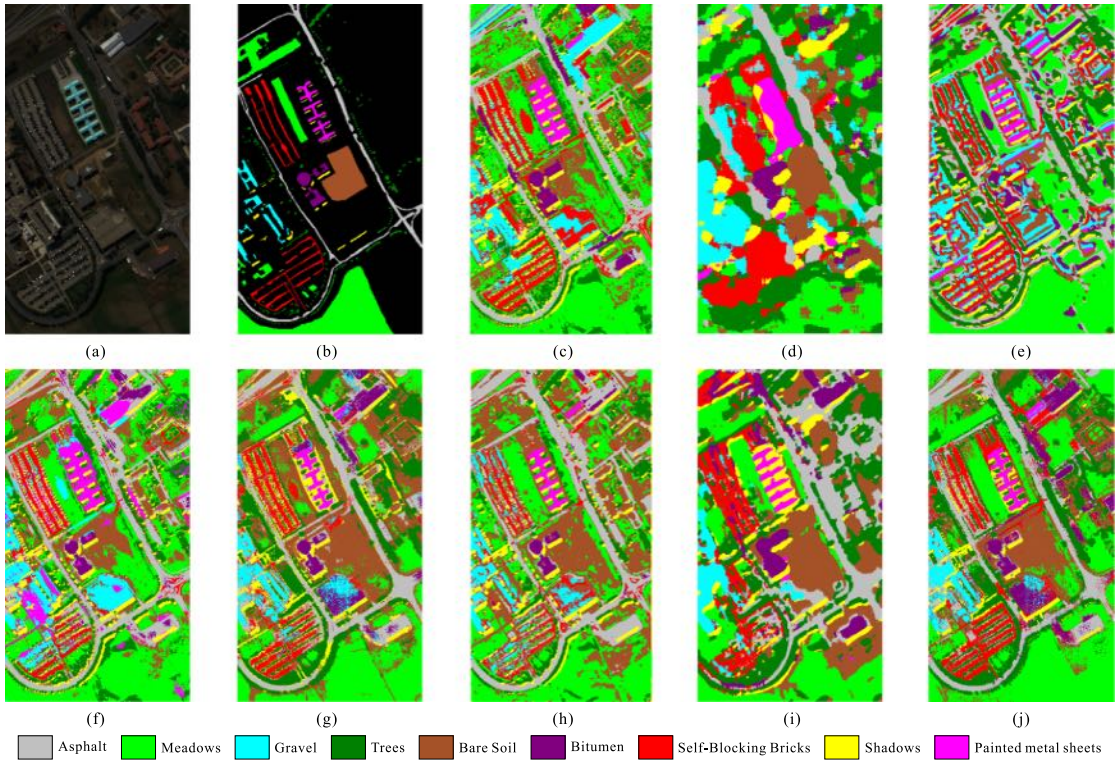

Fig. 7. Classification maps of the Pavia University dataset. (a) Three-band color composite of the HSI. (b) Ground-truth. (c) SCMK. (d) LBP. (e) deepNRD. (f) RPNet. (g) LCMR. (h) S-DMM. (i) HybridSN. (j) Proposed LHCMR.

TABLE IV
AVERAGE RESULTS (%) OF TEN EXPERIMENTS AT THE PAVIA UNIVERSITY DATASET WITH TEN TRAINING SAMPLES PER CLASS

Class	LBP	SCMK	deepNRD	RPNet	LCMR	S-DMM	HybridSN	LHCMR
1	57.70	84.67	62.74	82.73	82.32	**92.13**	66.76	86.40
2	70.43	74.73	89.16	75.71	84.85	69.06	80.89	**90.56**
3	87.67	85.23	45.85	81.99	87.66	88.99	72.62	**93.03**
4	65.17	93.09	88.36	90.75	**96.90**	93.09	89.72	94.72
5	93.85	99.60	82.58	99.81	95.74	**100.00**	87.72	97.24
6	90.93	88.93	36.56	81.37	91.84	84.82	**97.95**	91.61
7	84.68	96.40	55.49	92.80	95.28	98.14	**100.00**	95.20
8	88.01	75.78	83.47	81.21	79.93	**90.77**	77.40	85.19
9	63.72	**99.99**	97.18	98.01	93.19	99.96	87.51	94.58
OA	73.87	81.86	75.14	81.09	86.7	81.60	81.58	**90.44**
AA	78.02	88.71	71.27	87.15	89.75	90.77	84.5	**92.06**
k	67.41	77.04	66.96	76.04	82.89	76.88	76.54	**87.54**

LHCMR always performs the best among other comparison methods.

The second experiment was performed on the Pavia University image, and ten labeled samples per class are randomly chosen for training. Table IV tabulates the classification results (averaged over ten experiments) of diverse methods and Fig. 7 shows the classification maps. It can be seen that the proposed LHCMR achieves the best classification accuracy among other comparison methods. As shown in Fig. 7, the proposed LHCMR method yield a more smooth and accurate map. Furthermore, we also randomly choose 10, 15, 20, 25, and 30 samples per class for training, and the average classification results are shown in Fig. 8. It can be seen that with the increase of training samples, the performance of all competitive methods is on the rise, but our method is still the best. These results objectively verify the effectiveness of local homogeneity strategy for HSI classification.

TABLE V
AVERAGE RESULTS (%) OF TEN EXPERIMENTS AT THE SALINAS DATASET WITH FIVE TRAINING SAMPLES PER CLASS

Class	LBP	SCMK	deepNRD	RPNet	LCMR	S-DMM	HybridSN	LHCMR
1	93.73	98.59	59.78	98.39	99.94	99.31	**100.00**	99.92
2	87.71	98.17	81.03	99.34	81.80	98.09	**100.00**	87.62
3	84.67	90.62	86.61	93.32	97.95	92.97	**99.85**	99.53
4	97.10	94.33	83.87	99.57	**100.00**	89.50	82.22	99.98
5	91.03	96.17	81.18	93.90	96.79	**98.55**	79.12	98.38
6	82.31	**99.74**	91.48	97.67	98.79	98.86	99.72	98.69
7	84.45	**99.86**	92.81	98.37	86.27	99.80	99.47	93.67
8	75.42	59.41	63.25	52.02	76.19	69.61	79.30	**87.76**
9	86.21	**99.99**	87.29	99.75	87.05	99.19	99.77	87.72
10	76.73	77.63	66.33	82.11	88.60	78.21	83.35	**92.20**
11	91.51	93.91	97.55	92.38	**99.91**	96.06	90.12	99.28
12	83.32	95.37	75.08	97.90	96.71	**99.83**	67.95	**99.83**
13	86.31	98.56	93.96	98.64	95.97	**99.76**	97.91	97.26
14	80.31	88.73	80.00	91.27	93.00	95.41	90.33	**94.23**
15	82.42	67.58	59.03	75.16	85.19	72.60	65.51	**88.73**
16	93.91	94.73	45.28	94.59	87.53	89.42	**100.00**	91.90
OA	83.46	84.24	74.35	84.01	87.64	87.19	86.91	**92.13**
AA	86.07	90.84	77.78	91.52	91.98	92.32	89.66	**94.79**
k	81.69	82.52	71.57	82.30	86.30	85.77	85.45	**91.26**

The third experiment is performed on the Salinas dataset. In this dataset, only five labeled samples per class are randomly selected for training. The average classification accuracies obtained by different methods have been tabulated in Table V and the full classification maps are shown in Fig. 9. As can also be observed, the proposed LHCMR method delivers better performance than other compared methods, in terms of quantitative metrics and visual results.

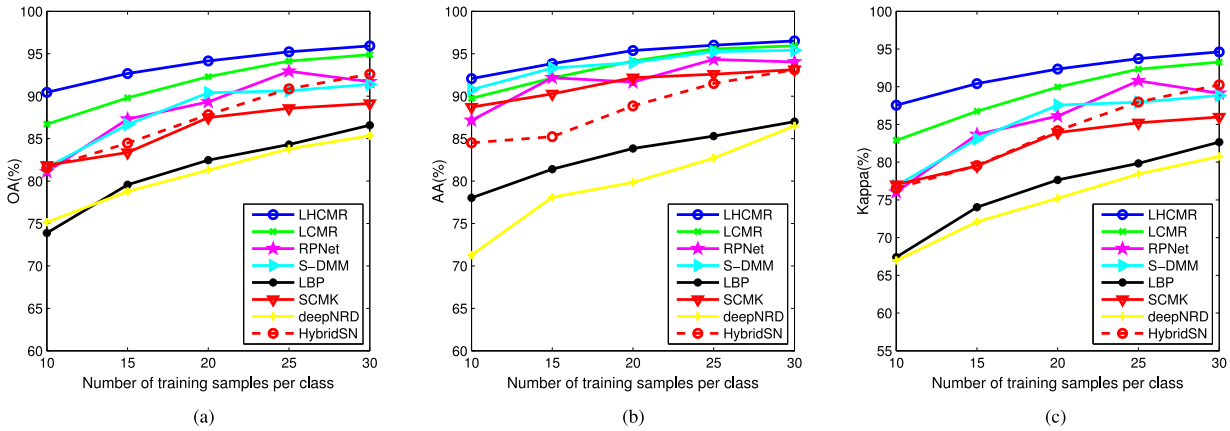

Fig. 8. Classification results of different numbers of training samples on the Pavia University Image. (a) OAs. (b) AAs. (c) Kappa.

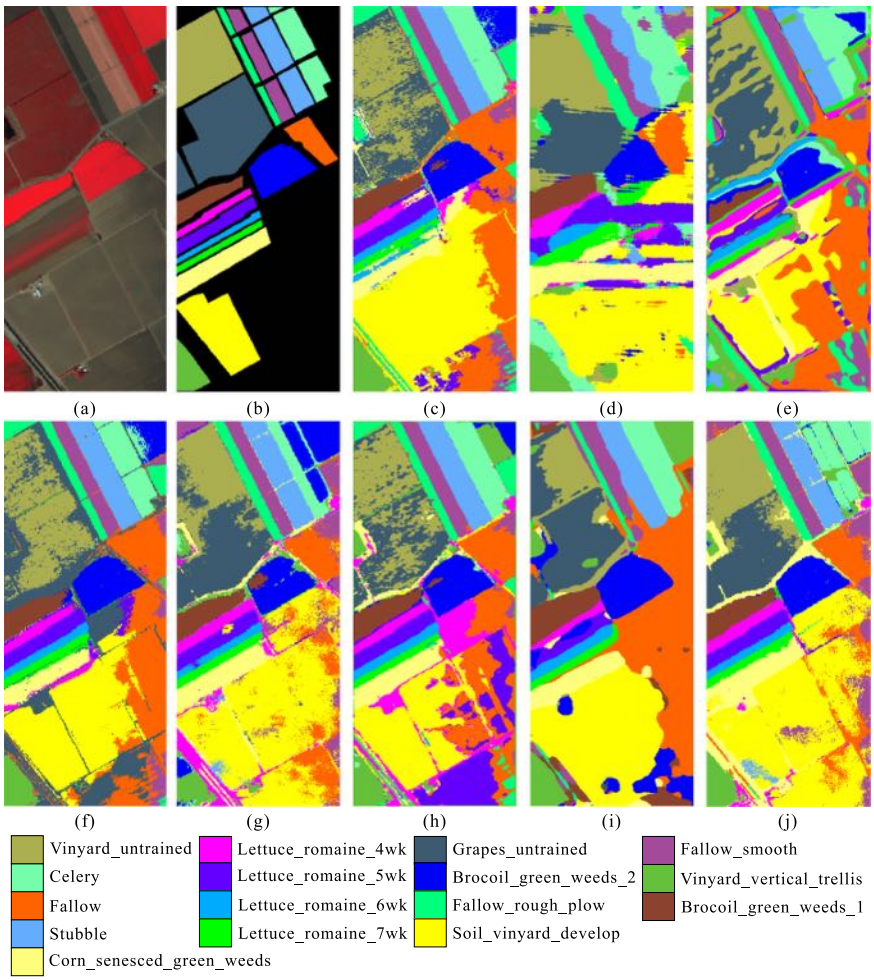

Fig. 9. Classification maps on the Salinas dataset. (a) Three-band color composite of the HSI. (b) ground-truth. (c) SCMK. (d) LBP. (e) deepNRD. (f) RPNet. (g) LCMR. (h) S-DMM. (i) HybridSN. (j) Proposed LHCMR.

Fig. 10 also illustrates OAs, AAs, and Kappa coefficients achieved by diverse methods using different training samples. It can be seen that the proposed method always achieves the best results, especially with a small number of training samples. The results demonstrate that the obtained locally homogeneous CM sets have provided more discriminative features.

The fourth experiment is performed on Pavia Center dataset. In this case, only five labeled samples per class are randomly chosen for training. Table VI reports the quantitative classification accuracies of different methods and Fig. 11 illustrates the full classification maps.

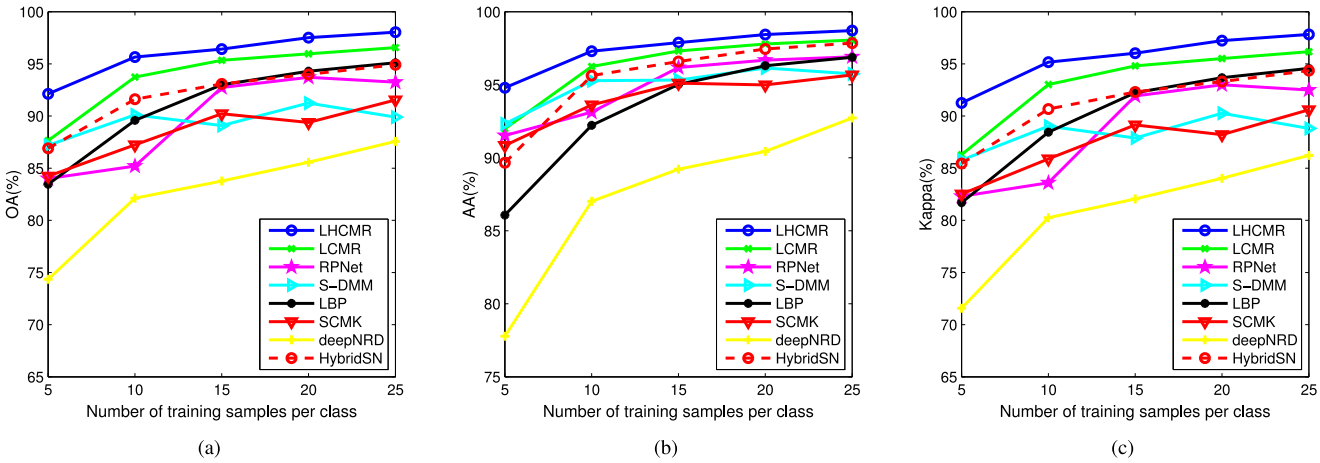

Fig. 10. Classification results of different methods with different numbers of training samples on the Salinas Image. (a) OAs. (b) AAs. (c) Kappa.

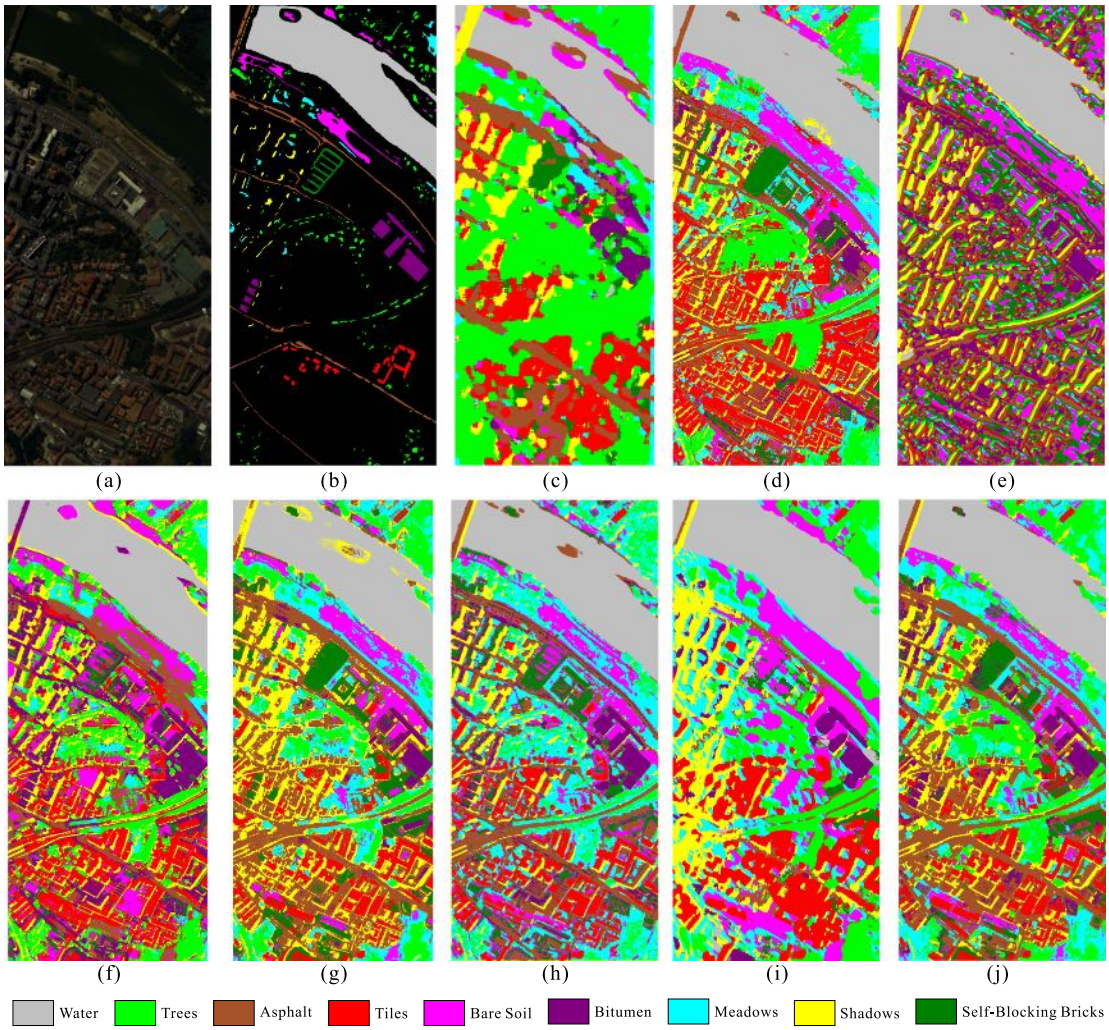

Fig. 11. Classification maps on the Pavia Center dataset. (a) Three-band color composite of the HSI. (b) Ground-truth. (c) SCMK. (d) LBP. (e) deepNRD. (f) RPNet. (g) LCMR. (h) S-DMM. (i) HybridSN. (j) Proposed LHCMR.

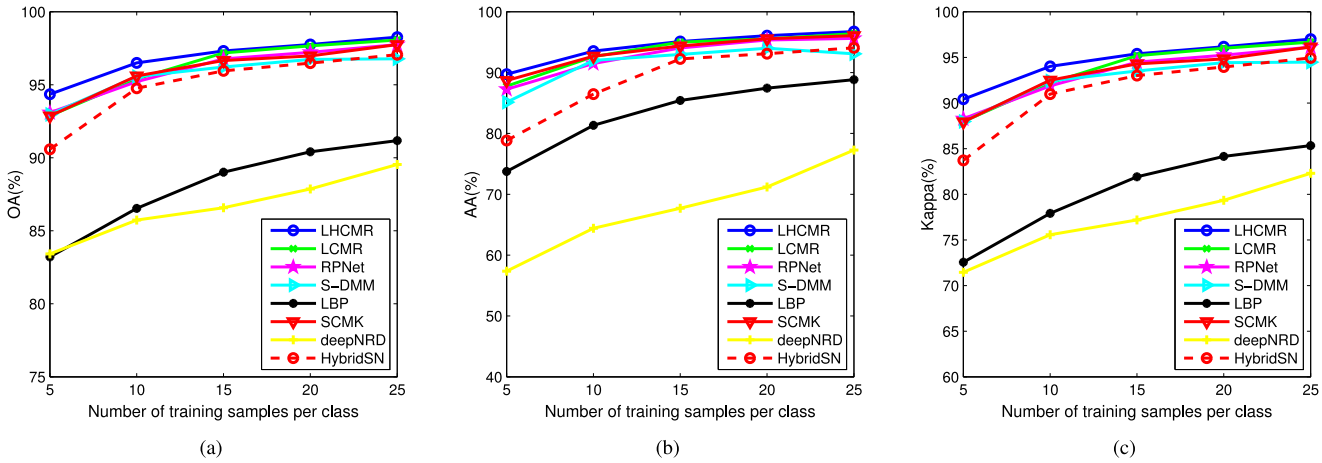

Fig. 12. Classification results of different numbers of training samples on the Pavia Center dataset. (a) OAs. (b) AAs. (c) Kappa.

TABLE VI
AVERAGE RESULTS (%) OF TEN EXPERIMENTS AT THE PAVIA CENTER DATASET WITH FIVE TRAINING SAMPLES PER CLASS

Class	LBP	SCMK	deepNRD	RPNet	LCMR	S-DMM	HybridSN	LHCMR
1	91.25	97.36	99.80	98.39	97.57	99.01	**99.99**	98.26
2	84.26	72.68	27.74	83.24	83.50	**85.92**	83.01	80.40
3	51.98	80.26	13.24	80.21	83.60	80.48	**89.10**	79.81
4	88.71	**99.16**	37.52	91.04	93.17	75.14	36.81	96.22
5	56.18	**87.85**	82.79	74.23	82.35	69.74	79.31	84.13
6	55.26	91.04	59.05	89.16	89.87	**96.30**	39.14	94.05
7	75.97	77.85	72.29	80.88	76.27	73.47	**87.98**	86.55
8	85.76	96.62	38.92	90.68	92.06	86.29	**99.10**	94.83
9	74.39	95.26	85.00	97.46	91.42	**99.84**	95.09	93.42
OA	83.21	92.86	83.42	93.09	92.88	92.97	90.58	**94.35**
AA	73.75	88.68	57.37	87.26	87.76	85.13	78.83	**89.74**
k	72.55	87.94	71.45	88.27	87.96	87.98	83.71	**90.41**

Fig. 12 also shows the effect of the number of training samples (ranging from 5 to 25, with a step of 5 per class) on the performance of different methods. The same conclusion can be drawn from this image. The proposed local-homogeneity-based approach can deliver competitive results with a few training samples.

V. CONCLUSION

In this article, LHCMR-based HSI classification method has been proposed. Not only the proposed method overcomes the drawbacks of nearest neighbors selection in a fixed window, but also achieves a competitive classification performance. The quantitative and visual results on homogeneous and nonhomogeneous HSI datasets illustrate that the proposed method is superior to several existing state-of-the-art classification methods. The reason for this may be twofold: first, by using a superpixel segmentation strategy on full bands, the HSI can be accurately segmented into many homogeneous regions based on spatial–spectral information. Second, locally homogeneous CMRs calculated based on the locally homogeneous area (overlapping area of the superpixels and the sliding window) have fully utilized the correlation among different spectral bands. For future works, we will introduce various promising distance metrics to exploit spectral–spatial information residing in HSI.

REFERENCES

[1] S. Jia, X. Deng, J. Zhu, M. Xu, J. Zhou, and X. Jia, "Collaborative representation-based multiscale superpixel fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7770–7784, Oct. 2019.

[2] S. Li, W. Song, L. Fang, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Deep learning for hyperspectral image classification: An overview," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 9, pp. 6690–6709, Sep. 2019.

[3] K. L. M. Ang and J. K. P. Seng, "Big data and machine learning with hyperspectral information in agriculture," *IEEE Access*, vol. 9, pp. 36699–36718, 2021.

[4] M. J. Khan, H. S. Khan, A. Yousaf, K. Khurshid, and A. Abbas, "Modern trends in hyperspectral image analysis: A review," *IEEE Access*, vol. 6, pp. 14118–14129, 2018.

[5] J. Aravinth, A. Bharadwaj, K. Harikrishna, and N. Vignajeeth, "Classification of urban objects from HSR-HTIR data using CNN and random forest classifier," in *Proc. 3rd Int. Conf. Commun. Electron. Syst.*, 2018, pp. 388–391.

[6] S. Jia, S. Jiang, Z. Lin, N. Li, M. Xu, and S. Yu, "A survey: Deep learning for hyperspectral image classification with few labeled samples," *Neurocomputing*, vol. 448, pp. 179–204, 2021.

[7] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Deep learning classifiers for hyperspectral imaging: A review," *ISPRS J. Photogrammetry Remote Sens.*, vol. 158, pp. 279–317, 2019.

[8] J. Jiang, J. Ma, Z. Wang, C. Chen, and X. Liu, "Hyperspectral image classification in the presence of noisy labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 851–865, Feb. 2019.

[9] S. Xu and J. Li, "Combining contextual information for subspace based hyperspectral image classification," in *Proc. Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens.*, Sep. 2018, pp. 1–5.

[10] J. Peng, X. Jiang, N. Chen, and H. Fu, "Local adaptive joint sparse representation for hyperspectral image classification," *Neurocomputing*, vol. 12, pp. 239–248, 2019.

[11] W. Yang, J. Peng, W. Sun, and Q. Du, "Log-Euclidean kernel-based joint sparse representation for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 12, pp. 5023–5034, Dec. 2019.

[12] S. Prasad and L. M. Bruce, "Limitations of principal components analysis for hyperspectral target recognition," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 625–629, Oct. 2008.

[13] A. A. Green, M. Berman, P. Switzer, and M. D. Craig, "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Trans. Geosci. Remote Sens.*, vol. GRS-26, no. 1, pp. 65–74, Jan. 1988.

[14] L. Zhang, L. Zhang, B. Du, J. You, and D. Tao, "Hyperspectral image unsupervised classification by robust manifold matrix factorization," *Inf. Sci.*, vol. 485, pp. 154–169, 2019.

[15] L. He, J. Li, C. Liu, and S. Li, "Recent advances on spectral-spatial hyperspectral image classification: An overview and new guidelines," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 3, pp. 1579–1597, Mar. 2018.

[16] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral-spatial hyperspectral image segmentation using subspace multinomial logistic regression and Markov random fields," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 3, pp. 809–823, Mar. 2012.
[17] J. A. Benediktsson, J. A. Palmason, and J. R. Sveinsson, "Classification of hyperspectral data from urban areas based on extended morphological profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 3, pp. 480–491, Mar. 2005.
[18] Y. Zhou and Y. Wei, "Learning hierarchical spectral-spatial features for hyperspectral image classification," *IEEE Trans. Cybern.*, vol. 46, no. 7, pp. 1667–1678, Jul. 2016.
[19] Y. Wei, Y. Zhou, and H. Hong, "Spectral-spatial response for hyperspectral image classification," *Remote Sens.*, vol. 9, no. 3, 2017, Art. no. 203.
[20] J. Li *et al.*, "Multiple feature learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1592–1606, Mar. 2015.
[21] X. Kang, S. Li, and J. A. Benediktsson, "Spectral-spatial hyperspectral image classification with edge-preserving filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 5, pp. 2666–2677, May 2014.
[22] L. Fang, S. Li, W. Duan, J. Ren, and J. A. Benediktsson, "Classification of hyperspectral images by exploiting spectral-spatial information of superpixel via multiple kernels," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 12, pp. 6663–6674, Dec. 2015.
[23] W. Li, C. Chen, H. Su, and Q. Du, "Local binary patterns and extreme learning machine for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 7, pp. 3681–3693, Jul. 2015.
[24] L. He, C. Liu, J. Li, Y. Li, S. Li, and Z. Yu, "Hyperspectral image spectral-spatial-range Gabor filtering," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4818–4836, Jul. 2020.
[25] H. Huang, Y. Duan, H. He, and G. Shi, "Local linear spatial-spectral probabilistic distribution for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1259–1272, Feb. 2020.
[26] M. Zhu, L. Jiao, F. Liu, S. Yang, and J. Wang, "Residual spectral-spatial attention network for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 449–462, Jan. 2021.
[27] C. Shi and C. Pun, "Multiscale superpixel-based hyperspectral image classification using recurrent neural networks with stacked autoencoders," *IEEE Trans. Multimedia*, vol. 22, no. 2, pp. 487–501, Feb. 2020.
[28] X. Han, B. Shi, and Y. Zheng, "SSF-CNN: Spatial and spectral fusion with CNN for hyperspectral image super-resolution," in *Proc. IEEE Int. Conf. Image Process.*, Oct. 2018, pp. 2506–2510.
[29] 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.
[30] M. E. Paoletti, J. M. Haut, J. Plaza, and A. Plaza, "Neighboring region dropout for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 6, pp. 1032–1036, Jun. 2020.
[31] C. Yu, R. Han, M. Song, C. Liu, and C. I. Chang, "A simplified 2D-3D CNN architecture for hyperspectral image classification based on spatial-spectral fusion," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 2485–2501, Apr. 2020.
[32] 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.
[33] J. Peng, W. Sun, L. Ma, and Q. Du, "Discriminative transfer joint matching for domain adaptation in hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 6, pp. 972–976, Jun. 2019.
[34] S. Zhong and Y. Zhang, "An iterative training sample updating approach for domain adaptation in hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, Jul. 2020, doi: 10.1109/LGRS.2020.3007021.
[35] Y. Xu, B. Du, F. Zhang, and L. Zhang, "Hyperspectral image classification via a random patches network," *ISPRS J. Photogrammetry Remote Sens.*, vol. 142, pp. 344–357, 2018.
[36] B. Deng, S. Jia, and D. Shi, "Deep metric learning-based feature embedding for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 1422–1435, Feb. 2020.
[37] L. Fang, N. He, S. Li, A. J. Plaza, and J. Plaza, "A new spatial-spectral feature extraction method for hyperspectral images using local covariance matrix representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 6, pp. 3534–3546, Jun. 2018.
[38] G. Zhao, N. Li, B. Tu, G. Zhang, and W. He, "Density peak covariance matrix for feature extraction of hyperspectral image," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 3, pp. 534–538, Mar. 2020.
[39] N. He *et al.*, "Feature extraction with multiscale covariance maps for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 2, pp. 755–769, Feb. 2019.
[40] X. Zhang, Y. Wei, H. Yao, and Y. Zhou, "Improved local covariance matrix representation for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, 2020, pp. 68–71.
[41] T. Zhan, Z. Lu, M. Wan, and G. Yang, "Multiscale superpixel kernel-based low-rank representation for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 17, no. 9, pp. 1642–1646, Sep. 2020.
[42] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai, and L. Wang, "SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 8, pp. 4581–4593, Aug. 2018.
[43] M. Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy rate superpixel segmentation," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2011, pp. 2097–2104.
[44] A. M. Saranathan and M. Parente, "Uniformity-based superpixel segmentation of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 3, pp. 1419–1430, Mar. 2016.
[45] R. Wang, H. Guo, L. S. Davis, and Q. Dai, "Covariance discriminative learning: A natural and efficient approach to image set classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, 2012, pp. 2496–2503.

Xinyu Zhang is currently pursuing the M.S. degree in computer science and technology with the School of Artificial Intelligence, Xidian University, Xi'an, China.

He was with the Hubei Research Center for Educational Informationization, Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, China. His research interests include computer vision and machine learning.

Yantao Wei received the B.Sc. degree in information and computing science from the Qingdao University of Science and Technology, China, in 2006, the M.S. degree in computational mathematics from the Huazhong University of Science and Technology, China, in 2008 and the Ph.D. degree in control science and engineering from Huazhong University of Science and Technology, China, in 2012.

He is currently an Associate Professor with the Hubei Research Center for Educational Informationization, Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, China. His research interests include educational artificial intelligence, computer vision, and machine learning.

Huang Yao received the B.Sc. degree in electronic publishing from the Wuhan University, China, in 2005, the M.S. degree in graphical communication engineering from the Wuhan University, China, in 2007 and the Ph.D. degree in photogrammetry and remote sensing from Wuhan University, China, in 2011.

He is currently a Lecturer with the Hubei Research Center for Educational Informationization, Faculty of Artificial Intelligence in Education, Central China Normal University, Wuhan, China. His research interests include educational artificial intelligence, computer vision, and machine learning.

Zhijing Ye received the B.Sc. and Ph.D. degrees in mathematics from the Huazhong University of Science and Technology, Wuhan, China, in 2011 and 2016, respectively.

He is currently with the School of Science, Wuhan University of Technology, Wuhan. His research interests include statistical learning, pattern recognition, and hyperspectral image processing.

Yicong Zhou (Senior Member, IEEE) received the B.S. degree from Hunan University, Changsha, China, in 1992, and the M.S. and Ph.D. degrees from Tufts University, Medford, MA, USA, in 2008 and 2010, respectively, all in electrical engineering.

In 2011, he joined as an Assistant Professor the Department of Computer and Information Science, University of Macau, Taipa, Macau, where he is currently a Full Professor and the Director of the Vision and Image Processing Laboratory. His research interests include image processing, computer vision, machine learning, and multimedia security.

Yue Zhao received the Ph.D. degree from the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan, China, in 2019.

She is currently a Lecturer with the School of Computer Science, Hubei University, Wuhan, China. Her research interests include computer vision, pattern recognition, and machine learning.