

Spectral–Spatial Hyperspectral Classification via Structural-Kernel Collaborative Representation

Bing Tu¹, Member, IEEE, Chengle Zhou¹, Student Member, IEEE,
Xiaolong Liao, Guoyun Zhang¹, Member, IEEE, and Yishu Peng

Abstract—This letter introduces a novel spatial–spectral classification method for hyperspectral images (HSIs) based on a structural-kernel collaborative representation (SKCR), which considers one weak assumption of spatial neighborhood that of the pixels in a superpixel belong to the same class when exploiting contextual information in HSI. The proposed method consists of the following steps. First, a superpixel segmentation strategy is used to construct self-adaptive regions for the HSI. Then, the structural information within each superpixel block is extracted based on the density peak and K nearest neighbors. Next, dual kernels are separately utilized for the exploitation of the spectral and the spatial information. Finally, the dual kernels are combined and incorporated into a support-vector-machine classifier. Since the weak assumption of spatial neighborhood is well considered in the collaborative representation, the proposed method showed excellent classification performance for two widely used real hyperspectral data sets even when the number of training samples was relatively small.

Index Terms—Density peak (DP), dual kernel, hyperspectral image (HSI), spectral–spatial classification, superpixel, support vector machines (SVMs).

I. INTRODUCTION

OVER the past several decades, there has been considerable interest in hyperspectral imaging in the remote sensing community, due to the capability of this technology to achieve more detailed spectral characteristics than 3-D pictorial data. Hyperspectral imagery (HSI) has been widely used in many applications, such as classification [1]–[3], anomaly detection, and others.

In recent years, HSI classification has become a hot topic in research into remote sensing. Researchers have designed

extremely interesting solutions to the problem of the separability of high-dimensional spectral response curves [4]–[7]. In [4], a generalized composite kernel (GCK) is proposed by utilizing the incorporation of nonlinear transformations of spectral and contextual signatures combined with the support vector machines (SVMs) [1] for HSI classification. The resulting classifier is referred to as SVM-CK and models spatial information based on extended multiattribute profiles (EMPs). Experiments showed that the classification performance of SVM-CK can be significantly improved with respect to the SVM classifier. Incorporating spatial information into the kernel, Fang *et al.* [5] presented an approach to the effective use of the spectral–spatial information of superpixels via multiple kernels, termed superpixel-based classification via multiple kernels (SC-MK). The key idea behind the SC-MK is the construction of three separate kernels by exploiting spectral information, mean operations within superpixel blocks, and weight fusion among each block. In addition to the abovementioned works, Li *et al.* [6] introduced a kernel collaborative representation with Tikhonov regularization (KCRT), which aims to improve the separability of spectral information while incorporating spatial information at neighboring locations into kernel space. Ma *et al.* [7] designed a discriminative kernel collaborative representation and a Tikhonov regularization method (DKCRT), which can make the kernel collaborative representation of different classes more discriminative in the HSI classification.

Spatial information plays an important role in HSI processing in such as classification and can effectively improve the accuracy of identification and interpretation of ground covering. However, a weak assumption is generally ignored by traditional spatial information (fixed and self-adaptive shape) utilization methods; this is that all the pixels in a neighborhood belong to the same class. In this letter, we develop a new spatial–spectral classification method for HSI based on a structural-kernel collaborative representation (SKCR), which takes full advantage of the self-adaptive spatial contextual information and considers the weak assumption of spatial information. Specifically, the main innovative contributions of the proposed approach can be summarized as follows.

- 1) The density peak (DP) algorithm [8] is introduced into superpixels to obtain the reliability of local spatial information of pixels. It is found that the pixels in the homogeneous region are not all in the same class based on density.
- 2) Combining DP and K nearest neighbors (KNN) algorithms can redefine the spatial structure information

Manuscript received November 28, 2019; revised March 22, 2020; accepted April 13, 2020. This work was supported in part by the National Natural Science Foundation of China under Grant 61977022 and Grant 51704115, in part by the Natural Science Foundation of Hunan Province under Grant 2019JJ50212, in part by the Key Research and Development Program of Hunan Province under Grant 2019SK2102, in part by the Hunan Provincial Innovation Foundation for Postgraduate under Grant CX20190914, in part by the Engineering Research Center on 3D Reconstruction and Intelligent Application Technology of Hunan Province under Grant 2019-430602-73-03-006049, and in part by the Hunan Emergency Communication Engineering Technology Research Center under Grant 2018TP2022. (Bing Tu and Chengle Zhou contributed equally to this work.) (Corresponding authors: Guoyun Zhang; Yishu Peng.)

The authors are with the School of Information Science and Engineering, Hunan Institute of Science and Technology, Yueyang 414000, China (e-mail: tubing@hnist.edu.cn; chengle_zhou@foxmail.com; xiaolongliao.china@gmail.com; gy Zhang@hnist.edu.cn; lovepys@hnist.edu.cn).

Color versions of one or more of the figures in this letter are available online at <http://ieeexplore.ieee.org>.

Digital Object Identifier 10.1109/LGRS.2020.2988124

1545-598X © 2020 IEEE. Personal use is permitted, but republication/redistribution requires IEEE permission.

See <https://www.ieee.org/publications/rights/index.html> for more information.

of pixels. It is found that the structural information can improve the performance of spectral-spatial classification methods more than the traditional spatial information.

Experiments on the Indian Pines and Washington DC HSI public data sets are adopted to demonstrate the qualitative and quantitative superiority of the proposed SKCR method over several well-known kernel-based algorithms.

II. RELATED WORK

A. DP Clustering

The DP clustering algorithm is based on the assumption that the cluster center is surrounded by low-density data points and is farther away from another high-density data point. In fact, each sample is analyzed in terms of its local density ρ and its distance δ to other data points with higher local density. The DP clustering algorithm [8] can be summarized as follows:

$$d_{ij} = \|x_i - x_j\|_2^2 \quad (1)$$

where x_i and x_j are the points that belong to a set $\mathbb{S} = \{x_i\}_{i=1}^n$, n represents the number of points, and d_{ij} is the Euclidean distance between points x_i and x_j . Considering the abovementioned definition, the kernel-based local density ρ_i of data point x_i can be achieved as one of the following two methods:

$$\rho_i = \begin{cases} \sum_j \chi(d_{ij} - d_c), & \text{Cutoff kernel} \\ \sum_j e^{-\left(\frac{d_{ij}}{d_c}\right)^2}, & \text{Gaussian kernel} \end{cases} \quad (2)$$

where d_c is the cutoff distance. The Gaussian kernel has the advantage of decreasing the negative impact of the statistical errors. According to the literature, Gaussian kernel-based density has proved to be successful in HSI interpretation [9].

B. Superpixel Segmentation

Superpixel segmentation algorithms, such as simple linear iterative clustering (SLIC) and entropy rate superpixel (ERS) [10], have been widely used in HSI processing to exploit spatial-contextual information around pixels. Assume that S_n refers to a predefined superpixel block. The graph-based ERS algorithm first maps the image to a graph $G = (V, E)$, where V is a set of vectors and E is a set of edges. Then, a subset of edges Q is selected to segment the graph into S_n -related local regions. $R(\cdot)$ (the entropy rate term of the random walk) and $B(\cdot)$ (a balancing term that reduces small superpixels) are incorporated into the objective function to form the balanced superpixels as follows:

$$\max_Q \{R(Q) + \lambda B(Q)\} \quad \text{s. t. } Q \subseteq E \quad (3)$$

where $\lambda \geq 0$ is a weight that controls the contribution of the entropy rate term and the balancing term. The problem in (3) can be solved efficiently by a greedy algorithm. Finally, given a common label for each superpixel, the base image \mathbf{I} can be described as follows:

$$\mathbf{I} = \bigcup_{s_i=0}^{S_n} \Omega_{s_i} \quad \text{and} \quad \Omega_{s_i} \cap \Omega_{s_j} = \emptyset, \quad (s_i \neq s_j) \quad (4)$$

where Ω_{s_i} and Ω_{s_j} represent any two various superpixels in the base image \mathbf{I} .

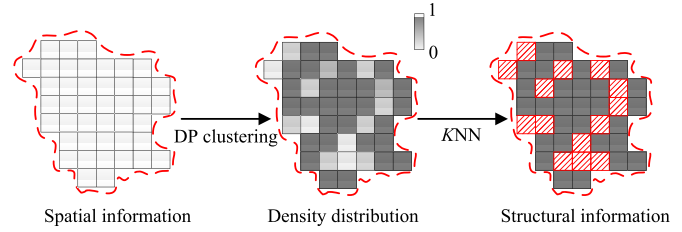


Fig. 1. Schematic of structural spatial information based on DP and KNN.

III. PROPOSED APPROACH

A. Motivation

In the traditional kernel-based HSI classification, the context information (fixed or adaptive) of a pixel is important feature information to identify its class attributes. Generally, researchers exploit spatial information based on an assumption that the pixels of the local region come from the same class. However, this is an assumption with insufficient arguments. Therefore, the motivation behind this letter is to combine DP and KNN algorithms to fully consider and overcome the weak assumption of local spatial information of pixels (see Fig. 1). Meanwhile, constructing a structural kernel utilizes spatial information more reliably.

B. SKCR Algorithm

In this letter, the structural-kernel-based collaborative representation is introduced to extract structural spatial information and improve the classification performance of the HSI supervised task. Specifically, the proposed SKCR method can effectively overcome the weak assumption of spatial information by adopting the DP and KNN approaches, which consists of three main steps. The details are given in the following.

1) *Construction of Shape-Adaptive Regions*: The principal component analysis (PCA) is first employed on the original HSI to obtain the first three principal components that are used as the base image for ERS segmentation. The HSI thus can be segmented into N nonoverlapping 2-D superpixel regions.

2) *Acquisition of Structure-Based Spatial Information*: Assume that a shape-adaptive region $\mathbf{R} = \{r_i^1, r_i^2, \dots, r_i^K\}$, where K refers to the number of pixels within a shape-adaptive region. The spectral angle can be calculated among different pixels for each shape-adaptive region as follows:

$$\text{dist} = \arccos\left(\frac{\langle r_i^u, r_i^v \rangle}{\|r_i^u\| \|r_i^v\|}\right). \quad (5)$$

With the aforementioned definitions in mind, the local density ρ^u of data point r_i^u can be calculated as follows:

$$\rho^u = \sum_u \exp\left\{-\left(\frac{\text{dist}}{d_c}\right)^2\right\}. \quad (6)$$

Next, we apply KNN to select pixels that have high density for each superpixel block. The formula is defined by

$$\hat{u} = \arg \max_{u \in k_0} \rho^u \quad (7)$$

where k_0 represents the range of pixels with higher density within a superpixel block. Meanwhile, a mean operation

is applied on the spectral pixels within each superpixel as follows:

$$\bar{r}_i = \frac{1}{N_{\hat{u}}} \sum_{k=1}^{N_{\hat{u}}} r_i^k \quad (8)$$

where $N_{\hat{u}}$ refers to the number of pixels with a relatively high density for each superpixel block.

3) *Spectral and Structural Kernel Collaboration*: In the training stage, a training set $(\mathbf{r}_1, \mathbf{r}_2, \dots, \mathbf{r}_n)$ is initially generated randomly from the original HSI. Then, the position indexes for selected pixels are utilized to extract pixels from the spectral feature image \mathbf{I}^{Spec} (constituted by all the spectral pixels) and the K -means feature image $\mathbf{I}^{\text{K-means}}$ (constituted by all the K NN-based filtering superpixels). The extracted pixels can constitute the spectral feature training data $(\mathbf{r}_1^{\text{S-T}}, \dots, \mathbf{r}_n^{\text{S-T}})$ and the K -means feature training data $(\mathbf{r}_1^{\text{K-T}}, \dots, \mathbf{r}_n^{\text{K-T}})$, respectively. Subsequently, the radial basis function (RBF) can be applied to the training samples to calculate a spectral kernel $K_{\text{spec}}^{\text{Train}}(\mathbf{r}_i^{\text{S-T}}, \mathbf{r}_j^{\text{S-T}})$ and a superpixel density-based structure kernel $K_{\text{struc}}^{\text{Train}}(\mathbf{r}_i^{\text{K-T}}, \mathbf{r}_j^{\text{K-T}})$ as follows:

$$K_{\text{spec}}^{\text{Train}}(\mathbf{r}_i^{\text{S-T}}, \mathbf{r}_j^{\text{S-T}}) = \exp\left(-\|\mathbf{r}_i^{\text{S-T}} - \mathbf{r}_j^{\text{S-T}}\|^2 / 2\theta^2\right)$$

$$K_{\text{struc}}^{\text{Train}}(\mathbf{r}_i^{\text{K-T}}, \mathbf{r}_j^{\text{K-T}}) = \exp\left(-\|\mathbf{r}_i^{\text{K-T}} - \mathbf{r}_j^{\text{K-T}}\|^2 / 2\theta^2\right) \quad (9)$$

where θ is half-peak width of the RBF kernel function. Here, the default setting in the LIBSVM library [11] is used. Next, the dual kernels are combined by using a weighted strategy

$$K_{\text{coll}}^{\text{Train}}(\mathbf{r}_i, \mathbf{r}_j) = \lambda \cdot K_{\text{spec}}^{\text{Train}} + (1 - \lambda) K_{\text{struc}}^{\text{Train}} \quad (10)$$

where λ are the tradeoff parameters. Finally, the collaborative kernel $K_{\text{coll}}^{\text{Train}}(\mathbf{r}_i, \mathbf{r}_j)$ can be incorporated to create a decision rule.

IV. EXPERIMENTAL RESULTS

A. Experimental Setup

1) *Data Sets*: To evaluate the effectiveness of the proposed SKCR method,¹ experiments are performed on two real HSI data set: the airborne visible/infrared imaging spectrometer (AVIRIS) Indian Pines data and the Hyperspectral Digital Image Collection Experiment (HYDICE) Washington DC. The first data set consists of 220 spectral bands across the spectral range of 0.4–2.5 μm and each band contains 145×145 pixels with a spatial resolution of 20 m per pixel. In our experiments, 20 water absorption bands (nos. 140–108, 150–163, and 220) of these data were removed. The false-color composition and reference data of the 16 classes of the Indian Pines data set are shown in Fig. 2(a) and (b). The second data set contains 280 scan lines and 307 pixels in each scan line. The false-color composite and reference data (six classes) of the Washington DC are shown in Fig. 2(c) and (d).

2) *Quality Indexes*: Three objective metrics, such as overall accuracy (OA), average accuracy (AA), and the Kappa coefficient, are adopted in our experiments to evaluate the quality of classification results. The OA measures the percentage of pixels that are correctly classified. The AA is the mean of the percentage of correctly classified pixels for each class. The Kappa coefficient measures the percentage of correctly

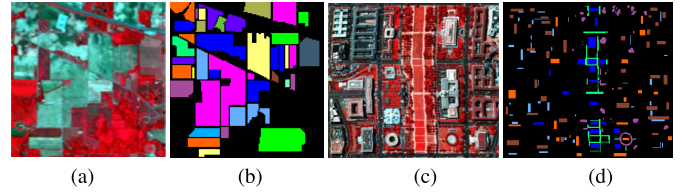


Fig. 2. Indian Pines and Washington DC data sets. Indian Pines data set: (a) three-band color composite and (b) reference data. Washington DC data set: (c) three-band color composite and (d) reference data.

TABLE I
DIFFERENT NUMBERS OF TRAINING AND TESTING SAMPLES OF SIXTEEN CLASSES IN THE INDIA PINES DATA SET AND SIX CLASSES IN THE WASHINGTON DC DATA SET

Indian Pines			Washington DC		
No	Class	Train/Test	No	Class	Samples
1	Alfalfa	3/43	1	Roof	3129
2	Corn-notill	20/1408	2	Grass	1790
3	Corn-min	13/817	3	Road	1402
4	Corn	4/233	4	Trail	1264
5	Grass/Pasture	8/475	5	Tree	1194
6	Grass/Trees	10/720	6	Shadow	1120
7	Grass/Pasture-mowed	3/25			
8	Hay-windrowed	8/470			
9	Oats	3/17			
10	Soybeans-notill	14/958			
11	Soybeans-min	31/2424			
12	Soybean-clean	9/584			
13	Wheat	4/201			
14	Woods	17/1248			
15	Building-Grass-T	10/376			
16	Stone-steel Towers	4/89			
Total		161/10088	Total		9899

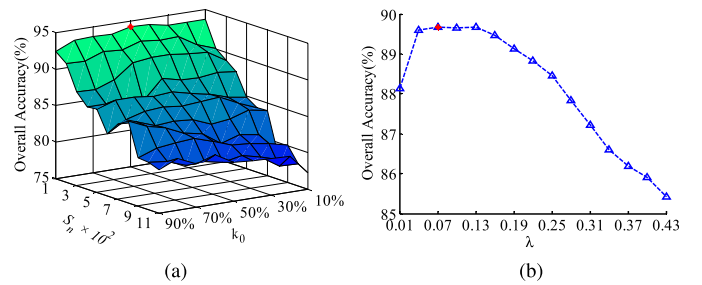


Fig. 3. (a) Superpixel blocks S_n and the nearest neighbor ratio k_0 for the analysis using the Indian Pines data set. (b) Influence of the tradeoff parameter λ on performance.

classified pixels corrected by the number of agreements that would be expected purely by chance.

B. Parameter Settings

In this section, the influence of the parameters relevant to the performance of the proposed SKCR method is analyzed. These parameters are the number of superpixel blocks S_n , the nearest neighbor ratio k_0 (the ratio of nearest neighbors to the total number of pixels in the corresponding superpixel block), and tradeoff parameter λ . The experiments are performed on the Indian Pines data set. As shown in Table I, the training samples used for classification are randomly selected, as 1.5% of the labeled data in the data set.

¹The MATLAB codes can be available at <https://github.com/chengle-zhou>

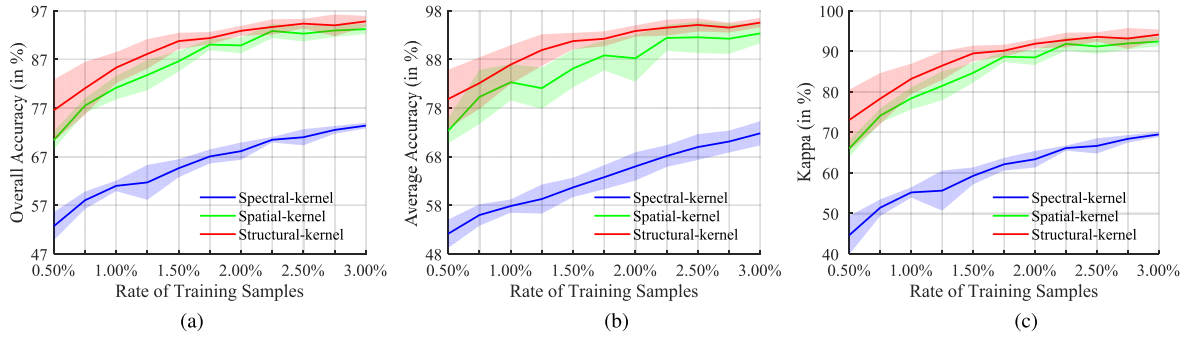


Fig. 4. Performance comparison of various kernel versions on the proposed method in the Indian Pine data set with different rates of labeled data. (a) OA. (b) AA. (c) Kappa coefficient.

In the first experiment, the impact of the superpixel blocks S_n and the nearest neighbor ratio k_0 on the performance of the proposed method on the classification of the hyperspectral data sets is tested. The ranges of S_n and k_0 are set to 1×10^2 to 11×10^2 and 10%–90%, respectively. As shown in Fig. 3(a), it can be found that the OAs of the classification results obtained using the proposed method are strongly affected by the changes to the parameter value. When k_0 is fixed, the spatial information of the pixels is not well extracted for the small superpixel size. Furthermore, the classification accuracy shows a trend of increasing first and then decreasing with the change in parameter k_0 when S_n is fixed. Therefore, the optimal parameter settings are set to $S_n = 100\%$ and $k_0 = 50\%$ in accordance with the best classification results produced by the proposed method.

In the second experiment, the influence of the tradeoff parameter λ on the performance of the proposed SKCR method is analyzed on the Indian Pine data set. As shown in Fig. 3(b), the classification accuracy of the SKCR is affected by λ and shows a trend of first rising and then decreasing. The reason is that the relative proportions of spectral and spatial information is controlled by λ in the SKCR. Therefore, $\lambda = 0.07$ is used as the default parameter setting for the SKCR.

C. More Analysis on Kernel Versions

To demonstrate the effectiveness of the proposed method, we further analyzed the effect of three kernel versions, i.e., spectral, spatial, and structural kernel, on classification performance with the different rates of training samples on the Indian Pine data set. The rate, referring to the randomly selected training samples that account for a given percentage of ground truth, is selected from the interval of $\{0.5\%, 1.0\%, \dots, 3.0\%\}$. As shown in Fig. 4, it can be observed that the structural-kernel-based SKCR method always achieves better classification results in terms of OAs, AAs, and Kappa. The reason is that: 1) for the spectral kernel, due to the small spatial resolution of the HSI data set, the differences among the classes are not significant; 2) for the spatial kernel, although the introduction of context information enhances variability, not all pixels in a neighborhood come from the same class (the existing spatial weak assumption); and 3) for the structural kernel, the advantage is that the spatial information of the pixel neighborhood is extracted according to local density and KNN to overcome the spatial weak assumption in 2).

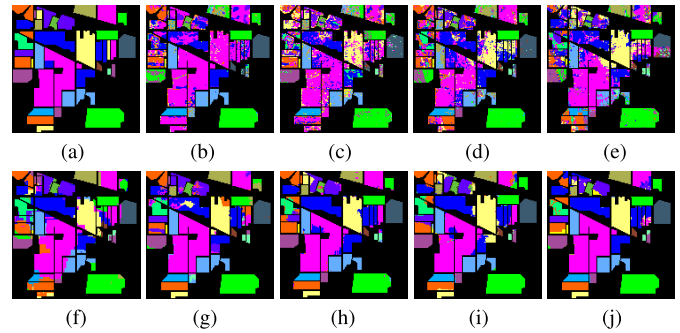


Fig. 5. Reference data and classification results (%) for the Indian Pine data set. (a) Reference data. (b)–(j) Classification maps generated by different methods: ELM (OA = 61.71%), SVM (OA = 66.23%), SVM-CK (OA = 68.91%), MLR-GCK (OA = 67.38%), JSRC (OA = 80.84%), SC-MK (OA = 87.74%), GFDN (89.77%), PCA-EPF (90.50%), and SKCR (OA = 93.33%).

D. Comparisons With Other Approaches

In this section, the proposed method is compared with the extreme learning machine (ELM) [12] method, SVM method [1], GCK (SVM-CK and multiple logistic regression (MLR)-GCK) [4] methods, the joint sparse representation classification (JSRC) method [13], the SC-MK method [5], Gabor filtering-based deep network (GFDN) [14], PCA based edge-preserving features (PCA-EPF) [15], and the proposed SKCR method. The SVM is implemented by applying the Gaussian kernel with fivefold cross validation. The other methods are implemented using the default parameters given by the authors.

This experiment is conducted on the Indian Pine data set using 1.5% of the labeled data as the training set (see Table I). The classification results is represented in Table II and Fig. 5. It can be observed from Table II that, for the proposed SKCR method, the classification results of all classes are above 85%, and the reports of nine classes achieve the highest accuracy with respect to other competitive methods. In addition to Table II, the classification results of the SKCR are more similar to the reference data compared with competitive methods (see Fig. 5). Specifically, it shows that the SKCR has a significant advantage in classification accuracy. Furthermore, we perform another experiment on a complex urban data set (Washington DC) to investigate the generalization of the proposed SKCR method. In the experiment, labeled data from 5 to 11 samples of each class is randomly selected as training samples. As shown in Fig. 6, it can be seen that the

TABLE II

CLASSIFICATION ACCURACY (IN PERCENT) OF THE INDIAN PINES IMAGE IN THE ELM, SVM, SVM-CK, MLR-GCK, JSRC, SC-MK, GFDN, PCA-EPF, AND SKCR METHODS. THE NUMBER IN THE PARENTHESIS IS THE STANDARD VARIATION OF THE ACCURACIES OBTAINED IN THE REPEATED EXPERIMENTS

The number of training samples is 1.5% of the reference data.									
Class	ELM	SVM	SVM-CK	MLR-GCK	JSRC	SC-MK	GFDN	PCA-EPF	SKCR
1	75.85(21.1)	39.18(12.5)	73.95(11.6)	59.53(18.4)	92.56(18.0)	100.0(0.00)	91.33(11.2)	100.0(0.00)	100.0(0.00)
2	59.63(3.85)	54.96(5.58)	62.83(3.88)	61.34(4.12)	77.27(6.65)	85.09(4.39)	83.76(4.45)	83.94(7.38)	85.55(5.46)
3	55.09(10.0)	51.30(7.07)	51.42(4.36)	48.08(7.14)	72.26(6.86)	81.82(10.1)	83.17(11.6)	82.89(8.33)	89.77(6.72)
4	53.16(30.3)	39.65(11.0)	34.76(8.78)	46.87(7.32)	78.28(11.4)	70.09(25.7)	84.72(11.8)	81.34(7.95)	82.32(16.7)
5	84.29(12.1)	77.93(12.6)	60.08(8.12)	51.62(9.14)	73.75(9.43)	78.72(4.29)	83.71(7.72)	85.41(4.66)	81.66(7.39)
6	77.74(2.38)	81.31(5.74)	83.24(8.87)	80.38(7.58)	69.81(10.4)	99.71(0.92)	91.54(6.72)	98.74(5.52)	98.64(2.13)
7	47.56(12.3)	47.22(23.9)	92.40(5.80)	86.80(14.4)	94.00(19.0)	100.0(0.00)	95.56(8.34)	88.02(26.0)	98.40(2.80)
8	92.41(2.68)	93.42(2.30)	93.85(5.23)	82.11(12.9)	97.45(4.22)	98.94(3.36)	93.79(4.93)	100.0(0.00)	100.0(0.00)
9	21.62(6.85)	31.62(16.7)	77.06(16.0)	80.00(17.4)	30.59(40.4)	100.0(0.00)	85.79(14.3)	100.0(0.00)	100.0(0.00)
10	64.24(10.1)	57.52(8.15)	62.47(6.21)	67.46(4.42)	74.48(7.24)	86.05(8.02)	84.11(5.85)	86.46(10.1)	86.27(5.88)
11	53.66(2.30)	65.55(3.61)	75.85(3.40)	76.09(4.14)	83.71(4.65)	86.61(4.38)	93.73(3.54)	89.09(2.45)	93.81(2.61)
12	52.70(14.5)	39.37(5.82)	36.83(7.94)	37.96(8.54)	61.82(12.3)	69.26(10.9)	83.05(9.94)	86.03(4.60)	86.25(8.65)
13	87.47(7.65)	86.92(5.16)	82.84(10.4)	61.74(18.6)	60.60(18.7)	93.73(13.0)	97.51(3.50)	100.0(0.00)	99.90(0.21)
14	88.51(3.26)	86.89(5.57)	92.58(2.52)	91.85(2.62)	95.62(3.41)	90.87(5.77)	97.78(2.91)	90.92(0.95)	98.25(2.64)
15	76.18(8.42)	44.31(5.33)	69.02(5.55)	67.02(8.50)	86.33(6.30)	91.28(8.95)	91.11(10.2)	98.79(1.85)	88.46(10.6)
16	95.44(8.69)	96.04(9.58)	75.51(11.9)	75.51(9.16)	64.83(29.9)	98.88(0.00)	83.74(15.2)	97.09(1.36)	97.53(1.16)
OA	65.33(1.39)	65.00(1.24)	70.18(1.59)	68.78(2.07)	79.71(1.32)	86.77(2.17)	89.49(1.45)	90.76(1.83)	91.52(1.44)
AA	67.85(2.91)	62.08(2.20)	70.29(2.04)	67.15(1.95)	75.83(3.47)	89.44(2.00)	89.02(2.70)	91.79(1.42)	92.93(1.45)
Kappa	59.51(1.73)	59.83(1.37)	65.81(1.85)	64.27(2.36)	76.93(1.50)	84.95(2.46)	88.00(1.65)	89.42(2.07)	90.34(1.62)

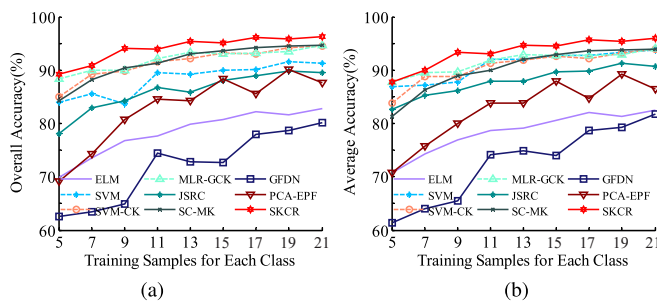


Fig. 6. Classification results for the Washington DC data set with a varying number of training samples by sparse representation classifier (SRC), ELM, SVM, SVM-CK, MLR-GCK, EPF, JSRC, SC-MK, and SKCR. (a) OA. (b) AA.

proposed SKCR method always obtains the best classification performance in terms of the highest OA and AA. These experimental results further verify the validity of the proposed SKCR method in hyperspectral classification.

V. CONCLUSION

In this letter, a new algorithm is proposed for HSI classification based on an SKCR. The basic idea behind this letter is that the proposed SKCR method takes full advantage of context information while considering the weak assumption of the spatial neighborhood. Exploiting DPs and KNN to extract the structural information of the neighborhood is the key contribution of the work described in this letter. The experimental results demonstrate that the SKCR produces excellent classification performance, exceeding that of traditional kernel-based classification approaches. In future work, we will explore the use of multipattern kernels to further improve the classification accuracy.

REFERENCES

- [1] 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.
- [2] B. Tu, C. Zhou, D. He, S. Huang, and A. Plaza, "Hyperspectral classification with noisy label detection via Superpixel-to-Pixel weighting distance," *IEEE Trans. Geosci. Remote Sens.*, early access, Jan. 20, 2020, doi: 10.1109/TGRS.2019.2961141.
- [3] 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.
- [4] J. Li, P. R. Marpu, A. Plaza, J. M. Bioucas-Dias, and J. A. Benediktsson, "Generalized composite Kernel framework for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 9, pp. 4816–4829, Sep. 2013.
- [5] 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.
- [6] W. Li, Q. Du, and M. Xiong, "Kernel collaborative representation with tikhonov regularization for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 1, pp. 48–52, Jan. 2015.
- [7] Y. Ma, C. Li, H. Li, X. Mei, and J. Ma, "Hyperspectral image classification with discriminative kernel collaborative representation and tikhonov regularization," *IEEE Geosci. Remote Sens. Lett.*, vol. 15, no. 4, pp. 587–591, Apr. 2018.
- [8] A. Rodriguez and A. Laio, "Clustering by fast search and find of density peaks," *Science*, vol. 344, no. 6191, pp. 1492–1496, Jun. 2014.
- [9] B. Tu, X. Zhang, X. Kang, J. Wang, and J. A. Benediktsson, "Spatial density peak clustering for hyperspectral image classification with noisy labels," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 5085–5097, Jul. 2019.
- [10] M.-Y. Liu, O. Tuzel, S. Ramalingam, and R. Chellappa, "Entropy-rate clustering: Cluster analysis via maximizing a submodular function subject to a matroid constraint," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 36, no. 1, pp. 99–112, Jan. 2014.
- [11] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 1–27, Apr. 2011.
- [12] G.-B. Huang, H. Zhou, X. Ding, and R. Zhang, "Extreme learning machine for regression and multiclass classification," *IEEE Trans. Syst. Man, Cybern. B, Cybern.*, vol. 42, no. 2, pp. 513–529, Apr. 2012.
- [13] Y. Chen, N. M. Nasrabadi, and T. D. Tran, "Hyperspectral image classification using dictionary-based sparse representation," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3973–3985, Oct. 2011.
- [14] X. Kang, C. Li, S. Li, and H. Lin, "Classification of hyperspectral images by Gabor filtering based deep network," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 11, no. 4, pp. 1166–1178, Apr. 2018.
- [15] X. Kang, X. Xiang, S. Li, and J. A. Benediktsson, "PCA-based edge-preserving features for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 12, pp. 7140–7151, Dec. 2017.