Hyperspectral Image Classification via Spatial Window-Based Multiview Intact Feature Learning

Yue Zhao, Yiu-ming Cheung, *Fellow, IEEE*, Xinge You, Qinmu Peng,
Jiangtao Peng, Peipei Yuan, and Yufeng Shi

Abstract—Due to the high dimensionality of hyperspectral images (HSIs), more training samples are needed in general for better classification performance. However, surface materials cannot always provide sufficient training samples in practice. HSI classification with small size training samples is still a challenging problem. Multiview learning is a feasible way to improve the classification accuracy in the case of small training samples by combining information from different views. This article proposes a new spatial window-based multiview intact feature learning method (SWMIFL) for HSI classification. In the proposed SWMIFL, multiple features that reflect different information of the original image are extracted and spatial windows are imposed on training samples to select unlabeled samples. Then, multiview intact feature learning is performed to learn the intact feature of the training and unlabeled samples. Considering that neighboring samples are likely to belong to the same class, labels of spatial neighboring samples are determined by two factors including the labels of training samples that locate in the spatial window and the labels learned from the intact feature. Finally, unlabeled samples that have same labels under these two factors are treated as new training samples. Experimental results demonstrate that the proposed SWMIFL-based classification method outperforms several well-known HSI classification methods on three real-world data sets.

Index Terms—Hyperspectral image (HSI) classification, multiview intact feature learning, small size training samples, spatial window.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) contain spectral information by reflecting electromagnetic waves in different bands. As different materials have different spectral reflectances, the spectral information can be used to distinguish pixels [1]. HSI classification aims to classify each pixel into one of land cover classes, which is an important part of hyperspectral image (HSI) analysis and has been applied in different fields, such as skin imaging, ground elements identifying, and mineral exploration [2].

For HSI classification, some pixel-based works that only utilize the spectral information have been proposed [2]. However, those methods often provide inconsistent salt-and-pepper classification results because they do not consider the spatial information [1]. In HSIs, pixels in the same homogeneous region are more likely to be in the same class [3]. Thus, in order to enhance the classification accuracy, some methods have been proposed to jointly exploit the spatial–spectral information of the HSI. For example, Fauvel *et al.* [4] classified HSIs with support vector machines (SVMs) using the spectral information and the spatial information derived by mathematical morphology. Li *et al.* [5] fused the spectral information with the intrinsic spatial information that contained in homogeneous regions of different sizes captured by the multiscale strategy. Shao *et al.* [6] incorporated the spatial information into sparse representation model via the graph Laplacian regularization, and obtained a more accurate coefficient matrix provided that spatial neighbors have similar representation coefficients. Pan *et al.* [7] proposed an ensemble learning-based HSI classification method that is composed of joint spectral–spatial features of different scales. Li *et al.* [8] exploited the spatial information by an iterative relaxation procedure which considers discontinuities existing in the data cube. Besides those, there are a number of work that utilized the spatial–spectral information for HSI classification [9]–[16]. Additionally, benefiting from the capacity of deep learning, some deep-learning-based methods have been proposed.

Liu *et al.* [17] used a deep residual 3-D convolutional neural network to extract the spectral–spatial features of HSIs to reduce the labeling uncertainty. Rao *et al.* [18] proposed a spatial–spectral relation network (SS-RN) for HSI classification. Ma *et al.* [19] proposed a network that takes the advantage of extra knowledge for information supplement and learns to compare rather than to classify for information exploration. Zhu *et al.* [20] proposed a well-designed 1-D-GAN as a spectral classifier and a robust 3-D-GAN as a spectral–spatial classifier. Pan *et al.* [21] utilized spatial information and spectral characteristics to construct the network and achieved higher accuracy when the number of training samples is limited. Zhong *et al.* [22] designed an end-to-end spectral–spatial residual network that takes raw 3-D cubes as input data without feature engineering for HSI classification. Pan *et al.* [23] proposed a small-scale data-based method to explore the application of deep learning approaches in HSI classification.

Except the pixel-based and the spatial–spectral-based methods, many other HSI classification approaches focused on extracting effective features. For example, Li *et al.* [24] extracted the local spatial texture information of HSIs by the local binary pattern operator. Kuo *et al.* [25] extracted the kernel nonparametric weighted feature of HSIs, which possesses the advantages of both linear and nonlinear transformation. Chang *et al.* [26] extracted low-dimensional features of HSIs via a nearest feature line embedding transformation. Shen *et al.* [27] selected informative and nonredundant Gabor features of HSIs through a symmetrical-uncertainty-based and Markov-blanket-based approach. Kuo *et al.* [28] extracted the kernel-based feature of HSIs with a criterion that contains the between-class and within-class information. Qian *et al.* [29] extracted 3-D discrete wavelet transform texture features of HSIs. Xia *et al.* [30] modeled the spatial-contextual information of the HSI by the extended multiattribute profiles. Tuia *et al.* [31] defined an active set feature learner to improve the HSI classification results. Zhao *et al.* [32] extracted the spectral feature and spatial-related features via balanced local discriminant embedding algorithm and convolutional neural network (CNN), and then stacked the spectral and spatial features together. Li *et al.* [33] learned more discriminative pixel-pair features of HSIs by utilizing deep CNN. Moreover, some works focused on extracting multiple features of HSIs. Li *et al.* [34] adaptively exploited information from both linear and nonlinear derived features of HSIs. Xu *et al.* [35] exploited multiple textural features of HSIs based on multiple morphological component analysis. Di *et al.* [36] generated multiple views by incorporating several hyperspectral data classification approaches with dynamic view updating and feature space bagging strategies. Volpi *et al.* [37] proposed a method for semi-supervised multiview feature extraction based on the multiset regularized kernel canonical correlation analysis. Zhang *et al.* [38] proposed a classification framework which can encode semantic context-aware representation to obtain promising features.

In the literature, multiview learning has a widespread application in machine learning community [39]. It aims to learn complementary information among multiview data. Multiview data is a set of data which is derived from different data sources that can describe objects from different aspects. If there is no natural source for multiview data, it is easy to manually generate multiview data by multiple feature extractors [40]. In this case, each feature is regarded as one view. In general, multiview data can improve the classification performance because they contain diversity and complementarity information of different single views. Some multiview learning-based HSI classification methods have been proposed recently. Fang *et al.* [2] proposed a multiple-feature-based adaptive sparse representation method for HSI classification. Xu *et al.* [40] proposed a new multiview active learning framework for HSI classification. Appice *et al.* [41] designed an application-specific co-training scheme to utilize both spectral information and spatial information for HSI classification. Di *et al.* [42] explored the intrinsic multiview information that is embedded in the hyperspectral data, then focused on samples with high uncertainty and built a contention pool to mitigate the computational cost, and they also proposed a multiview-based active learning method to optimally construct the training set for supervised HSI classification [43]. Zhou *et al.* [44] utilized a three-dimensional redundant wavelet transform to generate multiple views and integrated them in a multiview active learning framework. Chen *et al.* [45] treated different band sets as different views of land-covers and proposed a multiview graph embedding method to improve the performance of HSI classification.

Although the aforementioned methods applied multiview learning into HSI classification, none of them focused on learning multiview intact information of the HSI. In this article, we propose a spatial window-based multiview intact feature learning (SWMIFL) method for HSI classification. The proposed SWMIFL could extract multiview intact features from HSIs and also generate new training samples automatically. In detail, it imposes a spatial window on each training sample and predicts a label for each unlabeled neighboring pixel in the spatial window. On one side, due to the spatial homogeneous distribution of HSI, the unlabeled spatial neighboring pixels are likely to have the same label of the training sample located in the spatial window. On the other side, multiview intact feature is learned on both training samples and unlabeled spatial neighboring pixels. Unlabeled neighboring pixels can be labeled by classifying the intact feature using the nearest neighbor (NN) [46] classifier. For unlabeled sample x_i, we can assign label x_{l1} to it using NN classifier based on the intact feature, and meanwhile assign a label x_{l2} to it based on the labels of training samples in its spatial window. If x_{l1} and x_{l2} are the same, the unlabeled sample x_i will be regarded as new training samples otherwise the unlabeled sample is discarded. The above process will be repeated until all unlabeled samples are accessed. In summary, the main contributions of this article are threefold.

1) We are the first to attempt to utilize the multiview intact feature in HSI classification. The multiview intact feature contains intact information of the HSI, which can improve the classification performance.

TABLE I
FREQUENTLY USED NOTATION AND DESCRIPTIONS

notations	descriptions
z_i	A single sample vector.
$Z = \{z_1, z_2, ..., z_n\}$	A n sample set.
$Z^v = \{z_1^v, z_2^v, ..., z_n^v\}$	The v-th view of Z.
W_v	The view generation function of view v.

2) The spatial window is utilized in the intact feature learning process to utilize the spatial information of samples.

3) Taking into account the spatial neighborhood similarity and predicted labels on multiview intact features, the training set is gradually enlarged by automatically labeling pixels in the spatial neighborhood of training samples.

The rest of this article is organized as follows. Section II introduces the proposed method, including the introduction of multiview intact space learning, the description of the proposed SWMIFL method and its optimization process, the detailed description of the proposed method in steps. Section III discusses the experimental results and analysis, including the comparison of experiments between the proposed method and the state-of-the-art methods on three hyperspectral data sets. Section IV concludes the proposed method.

II. SPATIAL WINDOW-BASED MULTIVIEW INTACT FEATURE LEARNING FRAMEWORK

A. Notation

In this article, vectors are appeared in lowercase letters (e.g., z) and matrices are indicated by uppercase letters (e.g., Z). Specifically, I stands for the identity matrix. We represent a n sample set as $Z = \{z_1, z_2, \ldots, z_n\}$, where $z_i|_{i=1}^n$ denotes the ith column (sample) in Z. Moreover, Z^v represents the vth view of sample set Z: $Z^v = \{z_1^v, z_2^v, \ldots, z_n^v\}$. Furthermore, W_v and ε^v are the generation function and generation error for the vth view, respectively. For clarity, we summarize the frequently used notation and their corresponding descriptions in Table I.

B. Multiview Intact Space Learning Framework

Xu et al. [47] proposed the multiview intact space learning framework and verified the framework's superiority in face recognition, human motion recognition, and RGB-D object recognition. In the multiview intact space learning framework, it assumes that individual view captures partial information, and multiple views together possess redundant information of the object. The framework can address insufficiency in each individual view and integrate the encoded complementary information in multiple views to discover a latent intact representation of multiview data. For a n-sample and v-view training set: $Z^v|_{v=1}^m = \{z_1^v, z_2^v, \ldots, z_n^v\}$, the multiview intact space learning framework is defined as

$$\min_{W_v, X} \frac{1}{mn} \sum_{i=1}^n \sum_{v=1}^m \log\left(1 + \frac{\left\|z_i^v - W_v x_i\right\|^2}{c^2}\right) + C_1 \sum_{v=1}^m \|W_v\|_F^2 + C_2 \sum_{i=1}^n \|x_i\|^2 \quad (1)$$

where c is a constant scale parameter, and C_1 and C_2 are nonnegative constants that can be determined using cross validation.

Motivated by the multiview intact space learning framework, we will obtain the intact feature of HSI. As multiview intact space used an unsupervised learning process, we extend it to a semi-supervised method for HSI classification based on the characteristics of HSIs.

C. Spatial Window-Based Multiview Intact Feature Learning Framework

Adding information of unlabeled samples into learning process is a way to solve the small size training sample problem. In this process, unlabeled samples are automatically labeled by the proposed SWMIFL method and then treated as new training samples. In order to improve the classification performance, there are two factors needed to be considered: 1)the classification performance of the learned feature; 2) how to choose the new training samples. In the proposed spatial window-based multiview intact feature learning framework, we extract multiview intact feature of HSI to improve the classification performance. For the training samples, we use the objective function (1) to acquire the optimal intact feature generation function: $W_v^\star|_{v=1}^m$ and intact feature of training samples: X_{train}. Then, we establish a spatial window on each training sample. Unlabeled samples located in the spatial window are labeled in each iteration. In order to get the labels of these samples, we extract their intact features by

$$\min_{X_{\text{new}}} \frac{1}{mn} \sum_{i=1}^n \sum_{v=1}^m \log\left(1 + \frac{\left\|z_i^v - W_v^\star x_i\right\|^2}{c^2}\right) + C_2 \sum_{i=1}^n \|x_i\|^2 \quad (2)$$

where z_i^v is the ith unlabeled samples under view v, and x_i is the intact feature of ith unlabeled sample. After classifying unlabeled samples by the intact feature using NN classifier, we further consider the label of training samples located in the same spatial window. Only unlabeled samples that have consistent labels under these two factors will be regarded as new training samples.

D. Optimization of Spatial Window Based Multiview Intact Feature Learning Framework

According to the iteratively reweight residuals (IRR) optimization technique proposed in [47], we know that given fixed view generation functions $\{W_v\}_{v=1}^m$ or fix all data points $\{x_i\}_{i=1}^n$ in the intact feature space X, generation function W_v and intact feature X can be solved by two subproblems separately

$$\min_{W_v} \, g = \left\{ \frac{1}{n} \sum_{i=1}^n \log\left(1 + \frac{\|z_i^v - W_v x_i\|^2}{c^2}\right) + C_1 \|W_v\|_F^2 \right\} \quad (3)$$

$$\min_{x_i} \, h = \left\{ \frac{1}{m} \sum_{v=1}^m \log\left(1 + \frac{\|z_i^v - W_v x_i\|^2}{c^2}\right) + C_2 \|x_i\|^2 \right\}. \quad (4)$$

Setting the gradient of g with respect to $\{W_v\}$ to 0, we have

$$W_v = \sum_{i=1}^{n} z_i Q_i x_i^T \left(\sum_{i=1}^{n} x_i Q_i x_i^T + nC_1 \right)^{-1}. \quad (5)$$

Setting the gradient of h with respect to $\{x_i\}$ to 0, we have

$$x_i = \left(\sum_{v=1}^{m} W_v^T Q_v W_v + mC_2 \right)^{-1} \sum_{v=1}^{m} W_v^T Q_v z_i^v \quad (6)$$

where Q_v and Q_i are described in Algorithm I.

E. Detailed Description of Spatial Window-Based Multiview Intact Feature Learning in Steps

In this section, we will describe the proposed method in the following steps.

1) Step 1: Input the nonnegative parameters: c, C_1 and C_2 and extract multiview data $Z^v = \{z_1^v, \ldots, z_n^v\}|_{v=1}^{m}$ of the HSI. m is the number of views.
2) Step 2: Randomly select training samples.
3) Step 3: For the training samples, we randomly generate the original generation function $W_v^{(0)}|_{v=1}^{m}$ and intact feature $X^{(0)} = \{x_1^{(0)}, \ldots, x_n^{(0)}\}$. Generation function $W_v^{(0)}|_{v=1}^{m}$ can generate the v-th view data: $Z^v|_{v=1}^{m}$ from the intact feature $X^{(0)}$.
4) Step 4: According to the randomly generated generation function $W_v^{(0)}|_{v=1}^{m}$, we calculate $Q_v^{(0)}$ and the intact feature $X^{(1)} = \{x_1^{(1)}, \ldots, x_n^{(1)}\}$ by (6). Then, $Q_i^{(1)}$ and the generation function $W_v^{(1)}$ can be calculated by (5).
5) Step 5: If $\|x_i^{(1)} - x_i^{(0)}\|_{i=1}^{n}$ and $\|W_v^{(1)} - W_v^{(0)}\|_{v=1}^{m}$ satisfy the convergence conditions, we output the intact feature of training samples: $X_l = \{x_1^{(1)}, \ldots, x_n^{(1)}\}$ and generation function: $W_v^{\star} = W_v^{(1)}|_{v=1}^{m}$. Otherwise, we will repeat step 4 until $\|x_i^{(t)} - x_i^{(t-1)}\|_{i=1}^{n}$ and $\|W_v^{(t)} - W_v^{(t-1)}\|_{v=1}^{m}$ converge.
6) Step 6: For each training sample, we impose a 3×3 spatial-window on it to select unlabeled samples.
7) Step 7: We randomly generate the intact feature $X^{(0)} = \{x_1^{(0)}, \ldots, x_n^{(0)}\}$ of unlabeled data, then calculate $Q_v^{(0)}$ and the intact feature $X^{(1)} = \{x_1^{(1)}, \ldots, x_n^{(1)}\}$ by (2).
8) Step 8: If $\|x_i^{(1)} - x_i^{(0)}\|_{i=1}^{n}$ satisfies the convergence conditions, we output the intact feature of unlabeled samples: $X_u = \{x_1^{(1)}, \ldots, x_n^{(1)}\}$. Otherwise, we will repeat step 7 until $\|x_i^{(t)} - x_i^{(t-1)}\|_{i=1}^{n}$ converges.
9) Step 9: Classify the unlabeled samples by the intact feature.
10) Step 10: Classify the unlabeled samples by the class labels of training samples located in the spatial windows.
11) Step 11: Unlabeled samples which have same class labels by step 9 and step 10 are chosen as new training samples.
12) Step 12: Repeat Step 3 to Step 11 until all the unlabeled samples are classified or no sample satisfies Step 11.

According to the above descriptions, the spatial window based multiview intact feature learning framework is summarized in Algorithm I.

Algorithm 1

Input: non-negative parameters: c, C_1, C_2
  multi-view data: $Z^v = \{z_1^v, \ldots, z_n^v\}|_{v=1}^{m}$
Initialize: randomly   generated:
$W_v^{(t)}|_{t=0, v=1,\ldots,m}, x_i^{(t)}|_{t=0, i=1,\ldots,n}$

---

while(1)

$\begin{cases}
\textbf{for} \text{ training data, } i = 1, \ldots, n; v = 1, \ldots, m; \\
\quad t=0,1,2,\ldots, \textbf{do} \\
Q_v^{(t)} = \frac{1}{c^2 + \|z_i^v - W_v^{(t)} x_i^{(t)}\|^2} \\
x_i^{(t+1)} = \\
(\sum_{v=1}^{m}(W_v^{(t)})^T Q_v^{(t)} W_v^{(t)} + mC_2)^{-1} \sum_{v=1}^{m}(W_v^{(t)})^T Q_v^{(t)} z_i^v \\
Q_i^{(t+1)} = \frac{1}{c^2 + \|z_i^v - W_v^{(t)} x_i^{(t+1)}\|^2} \\
W_v^{(t+1)} = \\
\sum_{i=1}^{n} z_i Q_i^{(t+1)}(x_i^{(t+1)})^T (\sum_{i=1}^{n} x_i^{(t+1)} Q_i^{(t+1)}(x_i^{(t+1)})^T + nC_1)^{-1} \\
\textbf{if} \text{ the estimates of } x_i \text{ and } W_v \text{ converge then} \\
\textbf{break} \\
\textbf{end if} \\
\textbf{end for} \\
\textbf{Output:} \text{ intact feature: } X_l = \{x_1, \ldots, x_n\} \\
\quad \text{generation function: } W_v^{\star} = \{W_1^{\star}, \ldots, W_m^{\star}\}
\end{cases}$

Establish the spatial window on each training sample to extract testing samples.

$\begin{cases}
\textbf{for} \text{ testing data, } i = 1, \ldots, n; t = 0, 1, 2, \ldots, \textbf{do} \\
Q_v^{(t)} = \frac{1}{c^2 + \|z_i^v - W_v^{\star} x_i^{(t)}\|^2} \\
x_i^{(t+1)} = (\sum_{v=1}^{m}(W_v^{\star})^T Q_v^{(t)} W_v^{\star} + mC_2)^{-1} \sum_{v=1}^{m}(W_v^{\star})^T Q_v^{(t)} z_i^v \\
\textbf{if} \text{ the estimates of } x_i \text{ converge then} \\
\textbf{break} \\
\textbf{end if} \\
\textbf{end for} \\
\textbf{Output:} \text{ intact feature: } X_u = \{x_1, \ldots, x_n\}
\end{cases}$

Choose new training samples.
if meet the convergence conditions.
break
else repeat the above process.
end
end

III. EXPERIMENTAL RESULTS AND ANALYSIS

In this section, we will first introduce the data sets and experimental settings in this article. Then, we will describe the multiview feature extraction process. Furthermore, we will discuss the parameter setting. At last, the experimental results and analysis are listed.

A. Data Sets and Experimental Setting Description

Three real-world hyperspectral data sets are used to validate the performance of the newly proposed SWMIFL framework.

TABLE II

NUMBER OF TRAINING AND TESTING SAMPLES ON INDIAN PINES DATA SET

Class	Train	Test	Train	Test	Train	Test	Train	Test
Alfalfa	3	43	5	41	7	39	10	36
Corn-no till	3	1425	5	1423	7	1421	10	1418
Corn-min till	3	827	5	825	7	823	10	820
Corn	3	234	5	232	7	230	10	227
Grass/Trees	3	480	5	478	7	476	10	473
Grass/Pasture	3	727	5	725	7	723	10	720
Grass/Pasture-mowed	3	25	5	23	7	21	10	18
Hay-windrowed	3	475	5	473	7	471	10	468
Oats	3	17	5	15	7	13	10	10
Soybeans-notill	3	969	5	967	7	965	10	962
Soybeans-min till	3	2452	5	2450	7	2448	10	2445
Soybeans-clean till	3	590	5	588	7	586	10	583
Wheat	3	202	5	200	7	198	10	195
Woods	3	1262	5	1260	7	1258	10	1255
Bldg-grass-tree-drives	3	383	5	381	7	379	10	376
Stone-steel towers	3	90	5	88	7	86	10	83
Total	48	10201	80	10169	112	10137	160	10089

TABLE III

NUMBER OF TRAINING AND TESTING SAMPLES ON THE UNIVERSITY OF PAVIA DATA SET

Class	Train	Test	Train	Test	Train	Test	Train	Test
Asphalt	3	6628	5	6626	7	6624	10	6621
Meadows	3	18646	5	18644	7	18642	10	18639
Gravel	3	2096	5	2094	7	2092	10	2089
Trees	3	3061	5	3059	7	3057	10	3054
Metal sheets	3	1342	5	1340	7	1338	10	1335
Bare soil	3	5026	5	5024	7	5022	10	5019
Bitumen	3	1327	5	1325	7	1323	10	1320
Bricks	3	3679	5	3677	7	3675	10	3672
Shadows	3	944	5	942	7	940	10	937
Total	27	42749	45	42731	63	42713	90	42686

The first data set is gathered by the AVIRIS sensor over the Indian Pines test site in North-western Indiana. It consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range $0.4 \times 10^{(-6)} - 2.5 \times 10^{(-6)}$ m. The ground truth is designated into 16 classes and is not all mutually exclusive. By removing bands covering the region of water absorption, 200 bands are retained for the classification.

The second data set is acquired by the ROSIS sensor during a flight campaign over Pavia, northern Italy. It consists of 610×340 pixels. The number of spectral bands is 103 and the geometric resolution is 1.3 m. The ground truth is designated into nine classes.

The third data set is collected by the AVIRIS sensor over Salinas Valley, California, and is characterized by high spatial resolution (3.7-m pixels). It contains 512×217 pixels and 224 spectral reflectance bands. The ground truth is designated into 16 classes. The number of bands is reduced to 204 by removing bands covering the region of water absorption: [108–112], [154–167], 224.

In our experiments, we randomly choose 3, 5, 7, and 10 samples per class as training samples and the rest of the samples forms the testing set. The detailed sample number settings on three data sets are listed in Tables II–IV. The size of the spatial window used in this article is 3×3. Based on the extracted multiview intact features, the NN classifier is used for classification. The overall accuracy (OA), average accuracy (AA), κ coefficient, as well as classification accuracy of each class (CA) are used to evaluate different methods. All methods are implemented on MATLAB 2015b with 64 GB memory. All the results listed in this article are averaged on ten trials and the best results are marked in bold.

TABLE IV

NUMBER OF TRAINING AND TESTING SAMPLES ON SALINAS DATA SET

Class	Train	Test	Train	Test	Train	Test	Train	Test
Brocoli-green-weeds-1	3	2006	5	2004	7	2002	10	1999
Brocoli-green-weeds-2	3	3723	5	3721	7	3719	10	3716
Fallow	3	1973	5	1971	7	1969	10	1966
Fallow-rough-plow	3	1391	5	1389	7	1387	10	1384
Fallow-smooth	3	2675	5	2673	7	2671	10	2668
Stubble	3	3956	5	3954	7	3952	10	3949
Celery	3	3576	5	3574	7	3572	10	3569
Grapes-untrained	3	11268	5	11266	7	11264	10	11261
Soil-vinyard-develop	3	6200	5	6198	7	6196	10	6193
Corn-senesced-green-weeds	3	3275	5	3273	7	3271	10	3268
Lettuce-romaine-4wk	3	1065	5	1063	7	1061	10	1058
Lettuce-romaine-5wk	3	1924	5	1922	7	1920	10	1917
Lettuce-romaine-6wk	3	913	5	911	7	909	10	906
Lettuce-romaine-7wk	3	1067	5	1065	7	1063	10	1060
Vinyard-untrained	3	7265	5	7263	7	7261	10	7258
Vinyard-vertical-trellis	3	1804	5	1802	7	1800	10	1797
Total	48	54081	80	54049	112	54017	160	53969

B. Multiview Feature Extraction

In order to generate multiview feature with diversity and complementarity, we use multiple morphological component analysis (MMCA) method to extract multiple features of HSI [35]. MMCA method can adaptively exploit information from both linear- and nonlinear-derived features. It separates the image into multiple pairs of morphological components: enhanced attribute component x_s and weakened attribute component x_t. Each pair of components represents a linear combination of the original image x

$$x = x_s + x_t + n \qquad (7)$$

where n is the residual. Five attribute features, namely content, coarseness, contrast, horizontal, and vertical, are extracted by the MMCA method. The content feature includes the standard cartoon and texture components. The coarseness feature represents the edge intensity information in a local region. The contrast feature represents the variance rate of the pixel intensity. The horizontal and vertical ones are two directionality features that describe the orientation attribute of the local texture. Previous results have demonstrated that each view contains some information and multiple views contain complementary information of different views [40].

In this article, we use the above five views and also consider the spectral feature (original image) and maximum noise fraction transform (MNF) [48] feature. That is, seven views are used.

C. Visualization of the Multiview Intact Feature Learning Process

We use the "swiss roll" data to show the intermediate results of multiview intact feature learning process. The "swiss roll" data were widely used in machine learning. The data were created by randomly sampling from a Gaussian mixture model with centers/means at (7.5, 7.5), (7.5, 12.5), (12.5, 7.5), and (12.5, 12.5). The covariance for each gaussian was the 2×2 identity matrix. We uniformly sample 2000 3-D data points from the benchmark "swiss roll" data as shown in Fig. 1(a) and project these points into xy-plane, xz-plane, and yz plane to construct three different views [see Fig. 1(b)], two of them are flats, and the remaining one is a surface. We then utilize these three views to learn the intact feature. The original "swiss roll" data are the intact space of three views as it is
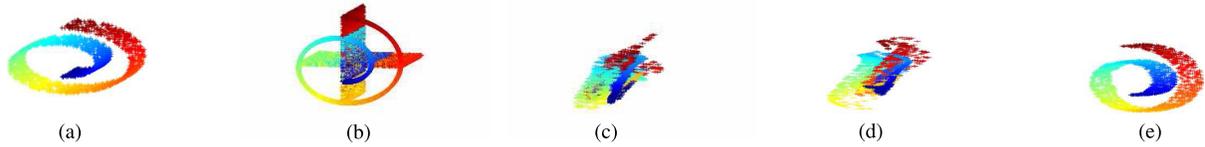
Fig. 1. Visualization of the multiview intact feature learning process. (a) Original "swiss roll" data. (b) Three views of the original data. (c) Feature learned in the $1^{\#}$ iteration. (d) Feature learned in the $2^{\#}$ iteration. (e) Learned multiview intact feature.

Fig. 2. Classification results with different values of parameter c with five training samples per-class.

Fig. 3. Classification results with different values of C_1 and C_2 with five training samples per-class. (The axis is transformed through $\log_{10}$.) (a) OA values. (b) AA values. (c) κ coefficients.
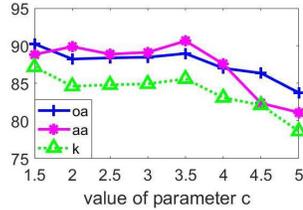

Fig. 4. Classification results with different values of window width with five training samples per-class.

the source domain of them. If the learned multiview intact feature contains intact information of three views, it should appear like a "swiss roll." In the learning process, it needs three iterations to get the final result. The intermediate results are shown in Fig. 1(c) and (d) and the final learned multiview intact feature is shown in Fig. 1(e). It can be seen that the proposed method can gradually learn intact feature within three iterations and the final learned intact feature is indeed a "swiss roll."

D. Parameter Setting

There are three nonnegative parameters in the proposed SWMIFL framework: c, C_1, and C_2. We first discuss parameter c in the proposed algorithm. The classification OA, AA, and κ coefficient with different values of parameter c on the University of Pavia data set are shown in Fig. 2. We can see that the SWMIFL algorithm shows relatively stable results when c is smaller than 3.5. In the following experiments, c is set as 2. C_1 and C_2 values used in SWMIFL are tuned in the range $\{10^{-10}, 10^{-9}, 10^{-8}, 10^{-6}, 10^{-4}\}$ and $\{10^{-8}, 10^{-7}, 10^{-6}, 10^{-5}\}$ respectively. We select one C1 and one C2 in the range each time, and obtain the classification accuracy with the selected values. After traversing all values of C1 and C2, the value with higher classification accuracy is used in comparison experiments. The results on the Salinas data set are showed in Fig. 3. Based on the results in Fig. 3, we fix $C_1 = 10^{-8}$ and $C_2 = 10^{-7}$ in all experiments. We then discuss the size of the spatial window in the proposed algorithm. The classification OA, AA, and κ coefficient with different widths of spatial window on the University of Pavia data set are shown in Fig. 4. We can conclude from the results that it is better to choose the 3×3 spatial window. This is because one of the factors we need to consider is that unlabeled samples may have the same label of training sample in the spatial window. The closer the samples locate, the more likely they belong to the same class. Choosing the 3×3 spatial window is to select the closest neighboring samples.

Additionally, we discuss parameter m (number of views) of the proposed method. Multiview data contain more information than single view data, this is why multiview data can be used to improve the classification results. As we have said in Section III-B, seven views are used in the proposed method. In order to analysis the influence of m, we show the classification results of the proposed method by using two views (spectral and MNF features), three views (spectral, MNF, and coarseness features), four views (spectral, MNF, coarseness, and horizontal features), five views (spectral, MNF, coarseness, horizontal, and contrast features), six views (spectral, MNF, coarseness, horizontal, contrast, and content features) and all seven views on University of Pavia and Salinas data set. The results are shown in Fig. 5(a) and (b) respectively. It can be seen that the results are bad when only two views are used. As the number of views increases, the performance of the proposed method is dramatically improved at first and remains stable when the number of views is larger than four. From the results in Fig. 5, we can conclude that the proposed method shows relatively stable results when the number of views is between four and seven. In the experiments, seven views are used.
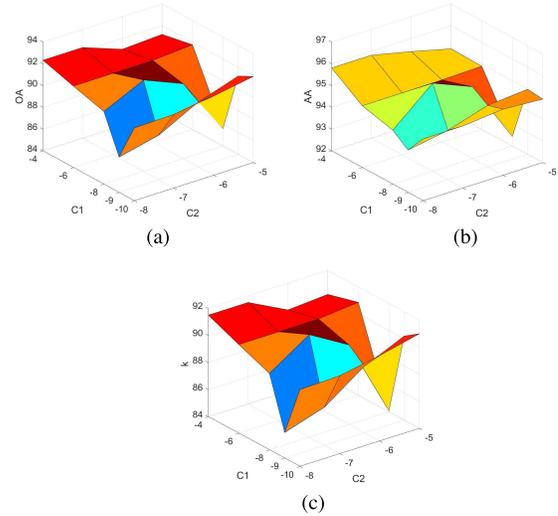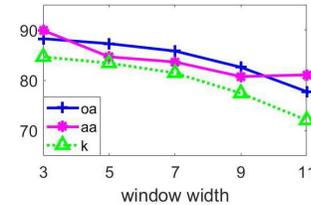
TABLE V

CLASSIFICATION RESULTS OF SWMIFL WITH NN CLASSIFIER AND SVM CLASSIFIER ON INDIAN PINES DATA SET

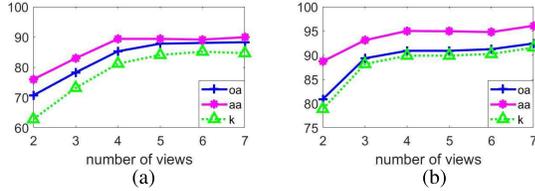|	3			5			7			10		
	OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
SWMIFL+NN	77.48	87.72	74.71	85.51	91.71	83.60	90.14	94.74	88.81	93.09	96.32	92.16
SWMIFL+SVM	78.28	89.17	75.71	86.31	92.33	84.50	92.17	95.38	91.05	94.95	97.46	94.24

Fig. 5. Classification results with different numbers of views with five training samples per-class. (a) Results on PaviaU data set. (b) Results on Salinas data set.

Fig. 6. Number of classes that each method outperforms other methods on three data sets.

As the proposed method focuses on the feature extraction via multiview intact feature learning, the simplest classifier, i.e., NN classifier, is used. To show the effect of classifier, here we test the classification performance of the proposed SWMIFL with NN classifier and SVM classifier [49] on Indian Pines data set. In the SVM, radial basis function (RBF) kernel is used. There are two parameters in the RBF kernel: γ and C. They are tuned in {0.001, 0.01, 0.1} and {10, 100, 1000, 10000}, respectively. The classification results with $C = 1000$, $\gamma = 0.01$ are listed in Table V. From the results, we can see that the classification results with SVM classifier is slightly higher than these with NN classifier. No matter which classifier is used, the results are excellent. That is, the performance of the proposed method is less dependent on the classifier. Considering that the NN classifier is free of parameters, we use it in the experiments.

E. Experimental Results Analysis

In this section, we will compare the proposed SWMIFL method with eight related methods, namely multiple-feature-based adaptive sparse representation (MFASR) [2], multiple logistic regression-based discontinuity preserving relaxation with spatial preprocessing (ppMLRpr) [8], local binary patterns and extreme learning machine (LBP-ELM) [24], hierarchical guidance filtering-based ensemble classification (HiFi-we) [7], multiple morphological component analysis (MMCA) [35], multiview intact space learning (MISL) [47], spatial and class structure regularized sparse representation graph (SCSSR) [6], semi-supervised partial label learning (SSPL) [50], and rolling guidance filter and vertex component analysis network (RVCANet) [21]. Among these methods, MFASR and MISL are multiview learning-based methods, HiFi-we is very similar to multiview learning-based methods, SSPL and SCSSR are semi-supervised methods, RVCANet is deep learning-based method and reveals better performance when the training samples available is limited. Specially, we compare SWMIFL framework with SCSSR by choosing 15 samples per-class as training samples on Indian Pines data set. The comparison experimental results and analysis are listed below.
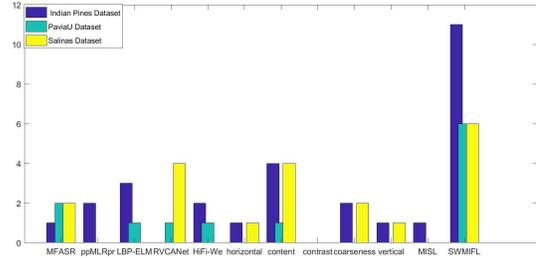
1) Experimental Results on the Indian Pines Data Set: On the Indian Pines data set, the dimensionality of intact feature learned by SWMIFL framework is 20. Table VI shows the classification results of the proposed SWMIFL framework and the comparison methods. We can see from the results that the OA, AA, and κ coefficient of the proposed SWMIFL framework are better than other methods, including both multiview learning-based methods and other outstanding classification methods. In detail, the OA of SWMIFL is 10% higher than the second-best method in the case of three and five training samples per class. To show the specific results on each class, we show the classification accuracy of each class in the case of 5 training samples per class in Table VII and also list the number of classes that each method outperforms other methods in Fig. 6. It is clear that the proposed SWMIFL method obtains the highest classification accuracy on 11 out of 16 classes. On the classes "Corn-no till," "Corn-min till," "Soybeans-notill," and "Soybeans-min till," all the comparison methods show bad results while our SWMIFL provides acceptable results and improves other methods nearly 10% in accuracy. In particular, the proposed SWMIFL shows excellent performance (100% accuracy) on the classes with extremely limited samples (i.e., "Grass/Pasture-mowed" and "Oat" with 28 and 20 samples). It is demonstrated that the proposed method is effective for the classification in the case of small training samples. Fig. 7 shows the classification maps of all the methods when the training sample number of each class is 5. Compared with other methods, the classification map of our method is more similar to the classification map of the ground truth. The classification results with 15 training samples per-class on Indian Pines data set are shown in Table VIII. We can see from Table VIII that the proposed method also has the highest OA, AA, and κ coefficient and outperforms other methods on most classes.

Here, we show the effect of the number of training samples on each methods and present the OA, AA, and κ coefficient results as the change of training samples in Fig. 8. As can be seen, the classification accuracy gaps between methods

TABLE VI

CLASSIFICATION RESULTS WITH DIFFERENT TRAINING SAMPLE NUMBERS ON INDIAN PINES DATA SET

Number (per-class)		3			5			7			10		
Accuracy		OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
MFASR		63.52	75.51	58.86	71.87	81.82	68.23	74.84	85.82	71.74	82.50	90.05	80.22
ppMLRpr		61.92	73.50	57.18	69.58	80.50	65.96	76.27	84.60	73.21	82.76	88.45	80.51
LBP-ELM		64.67	78.48	60.87	69.02	81.43	65.55	72.60	83.94	69.39	80.92	88.93	78.58
RVCANet		52.08	65.73	46.05	52.74	64.09	46.95	63.06	74.13	58.10	66.51	78.86	62.16
HiFi-We		66.03	78.31	61.89	75.14	85.60	72.08	82.43	89.85	80.17	86.27	92.51	84.51
MMCA	horizontal	66.52	80.37	62.59	72.52	84.93	69.22	76.33	87.02	73.44	80.00	89.66	77.58
	content	66.61	80.61	62.72	73.55	85.44	70.33	77.19	87.75	74.38	80.88	90.16	78.55
	contrast	66.39	79.62	62.53	71.30	83.44	67.99	75.29	86.40	72.36	78.44	88.83	75.87
	coarseness	66.67	79.20	62.98	72.11	83.62	68.97	75.66	86.15	72.86	78.84	88.60	76.34
	vertical	66.96	80.53	63.10	73.30	84.86	70.07	76.78	87.15	73.94	80.61	89.75	78.26
SSPL		42.82	56.28	36.51	47.99	61.62	42.20	50.09	64.67	44.64	53.68	67.51	48.53
MISL		63.07	78.47	59.02	65.72	79.40	61.62	70.10	82.99	66.59	65.56	75.97	61.63
SWMIFL		**77.48**	**87.72**	**74.71**	**85.51**	**91.71**	**83.60**	**90.14**	**94.74**	**88.81**	**93.09**	**96.32**	**92.16**

TABLE VII

CLASSIFICATION RESULTS OF EACH CLASS WITH FIVE TRAINING SAMPLES PER-CLASS ON INDIAN PINES DATA SET

| Class	MFASR	ppMLRpr	LBP-ELM	RVCANet	HiFi-We	MMCA					SSPL	MISL	SWMIFL
						horizontal	content	contrast	coarseness	vertical			
Alfalfa	97.56	87.32	**100.00**	90.70	98.78	94.57	93.48	97.46	97.83	94.93	84.42	96.01	98.91
Corn-no till	48.09	44.81	50.61	18.04	61.03	63.26	67.09	64.29	66.48	66.74	31.13	62.89	**76.07**
Corn-min till	63.70	50.96	53.57	44.14	71.15	63.26	68.01	59.56	59.78	64.90	39.86	53.23	**78.19**
Corn	69.40	86.15	94.37	51.71	93.75	90.93	88.96	92.83	99.09	88.33	48.38	90.65	**100.00**
Grass/Trees	69.42	82.13	73.33	43.13	76.67	92.31	**93.93**	82.82	81.95	88.23	64.91	64.56	93.20
Grass/Pasture	77.70	96.51	76.23	83.36	96.07	98.93	99.13	95.37	88.24	98.61	77.01	92.51	**99.77**
Grass/Pasture-mowed	99.28	99.40	**100.00**	96.00	97.83	100.00	100.00	99.40	99.40	**100.00**	89.88	98.81	**100.00**
Hay-windrowed	96.86	94.18	98.92	59.58	82.56	92.47	93.55	92.85	99.34	91.11	65.83	96.30	**100.00**
Oats	**100.00**	**100.00**	**100.00**	94.12	**100.00**	100.00	100.00	100.00	100.00	**100.00**	90.00	99.17	**100.00**
Soybeans-notill	70.44	73.29	74.76	42.41	70.27	69.19	68.86	64.51	67.87	72.57	50.74	59.86	**81.81**
Soybeans-min till	67.39	60.02	60.33	54.36	59.41	52.95	53.39	55.93	60.42	54.11	35.38	60.85	**82.15**
Soybeans-clean till	72.99	45.67	56.16	36.10	**80.10**	61.02	63.01	63.55	55.12	62.42	25.94	56.46	75.18
Wheat	99.50	99.76	97.72	99.50	99.50	99.59	99.19	98.94	98.29	99.11	93.33	99.43	**100.00**
Woods	88.44	**93.12**	82.28	88.03	96.31	85.69	85.06	82.12	80.01	85.59	67.51	44.81	89.06
Bldg-grass-tree-drives	90.68	76.08	85.79	27.68	86.75	95.08	94.30	85.84	84.63	92.40	28.11	**95.38**	92.96
Stone-steel towers	97.73	98.57	98.75	96.67	99.43	98.92	99.10	99.64	99.46	98.75	93.55	99.46	**100.00**
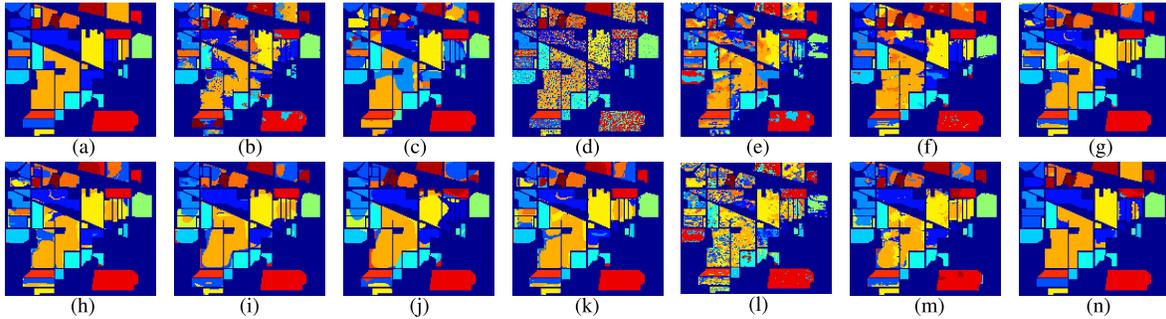
Fig. 7. Classification maps of all the compared methods with five training samples per-class on Indian Pines data set. (a) Groundtruth. (b) MFASR. (c) ppMLRpr. (d) LBP-ELM. (e) RVCANet. (f) HiFi-We. (g) Horizontal. (h) Content. (i) Contrast. (j) Coarseness. (k) Vertical. (l) SSPL. (m) MISL. (n) SWMIFL.

TABLE VIII

CLASSIFICATION RESULTS WITH 15 TRAINING SAMPLES PER-CLASS ON INDIAN PINES DATA SET

Class	MFASR	ppMLRpr	LBP-ELM	RVCANet	HiFi-We	MMCA					SCSSR	SSPL	MISL	SWMIFL
						horizontal	content	contrast	coarseness	vertical				
Alfalfa	97.31	87.68	**100.00**	96.77	99.46	98.19	97.83	99.64	99.64	98.19	71.70	88.04	97.46	99.46
Corn-no till	72.56	70.63	72.00	44.44	69.19	77.03	79.98	75.37	75.97	79.31	**90.00**	38.39	79.12	89.92
Corn-min till	86.93	81.99	82.27	59.51	85.38	84.44	82.07	86.57	81.12	82.75	**100.00**	47.59	75.20	99.97
Corn	91.82	89.45	99.58	83.78	93.99	96.55	93.81	99.02	99.09	96.13	97.60	63.78	93.74	**100.00**
Grass/Trees	88.18	92.96	88.72	86.54	85.33	96.55	96.69	90.89	89.58	94.41	78.30	81.02	52.83	**98.14**
Grass/Pasture	96.34	98.74	93.72	98.18	98.58	99.63	99.89	98.45	96.37	99.73	95.60	84.36	73.72	**100.00**
Grass/Pasture-mowed	**100.00**	98.81	**100.00**	**100.00**	96.15	100.00	100.00	100.00	100.00	**100.00**	38.50	96.43	92.26	**100.00**
Hay-windrowed	**100.00**	95.29	**100.00**	97.41	98.74	98.54	99.16	99.41	**100.00**	97.45	**100.00**	82.32	99.62	**100.00**
Oats	**100.00**	**100.00**	**100.00**	**100.00**	**100.00**	100.00	100.00	100.00	100.00	**100.00**	20.80	98.33	90.00	**100.00**
Soybeans-notill	88.77	84.84	84.38	82.24	85.44	88.41	90.02	83.93	83.76	88.96	53.50	60.73	79.05	**93.06**
Soybeans-min till	77.75	73.11	71.80	54.02	76.86	72.08	72.32	73.51	76.15	74.16	73.60	43.10	73.17	**93.81**
Soybeans-clean till	91.87	81.03	81.17	71.80	90.77	85.64	88.56	81.08	75.83	85.41	95.90	44.07	84.82	**97.93**
Wheat	99.56	99.51	98.21	**100.00**	99.51	100.00	99.76	99.07	98.13	99.84	100.00	92.93	68.92	**100.00**
Woods	**98.81**	95.96	95.42	91.60	94.72	92.74	93.10	90.25	89.06	93.02	48.70	72.92	25.06	93.28
Bldg-grass-tree-drives	93.40	96.20	97.41	48.79	95.24	98.53	99.35	97.80	97.50	99.31	94.60	40.33	72.02	**100.00**
Stone-steel towers	98.72	97.67	99.46	**100.00**	99.36	99.10	**100.00**	99.28	98.92	99.82	77.70	95.70	**100.00**	**100.00**
OA	86.68	83.86	83.64	70.13	85.06	85.62	86.28	84.63	84.21	86.25	87.59	57.07	70.56	**95.41**
AA	92.63	90.24	91.51	82.19	91.8	92.96	93.28	92.18	91.32	93.03	77.28	70.63	78.56	**97.85**
κ	84.87	81.82	81.65	66.53	83.07	83.80	84.55	82.70	82.24	84.50	86.23	52.18	67.07	**94.78**

are gradually reduced and the classification accuracies tend to be stable when the number of training samples increases. In particular, the classification accuracies of SWMIFL are always higher than other methods, and can achieve acceptable classification results when the number of training samples is small. It means the proposed SWMIFL method is more robust to the change of training sample number.

Moreover, the run time of all methods on the Indian Pines data set with five training samples per-class is recorded. The results are shown in Table IX. It can be seen that the run
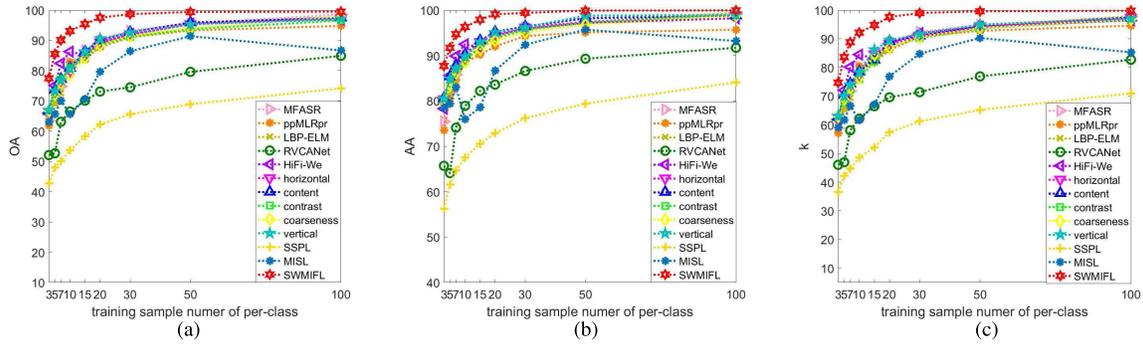
Fig. 8. Classification results with different training sample numbers on Indian Pines data set. (a) OA values. (b) AA values. (c) κ coefficients.

TABLE IX
RUN TIME (SECONDS) FOR THE CLASSIFICATION OF THE INDIAN PINES DATA SET WITH FIVE TRAINING SAMPLES PER-CLASS

	MFASR	ppMLRpr	LBP-ELM	RVCANet	HiFi-We	MMCA	SSPL	MISL	SWMIFL
Run Time	232.1	27.1	5.7	5053.1	334.9	8521.4	5.4	3.5	602.5

TABLE X
CLASSIFICATION RESULTS WITH DIFFERENT TRAINING SAMPLE NUMBERS ON UNIVERSITY OF PAVIA DATA SET

Number(per-class)		3			5			7			10		
Accuracy		OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
MFASR		69.82	83.81	63.28	76.95	88.21	71.36	79.73	89.44	74.72	86.67	94.08	83.14
ppMLRpr		68.12	77.08	61.11	77.39	81.13	71.36	78.62	82.11	73.07	80.02	84.13	74.75
LBP-ELM		48.43	52.69	37.73	55.92	57.58	45.17	61.18	63.66	52.23	69.30	69.88	61.19
RVCANet		56.28	71.46	47.11	62.16	79.39	54.34	66.20	79.50	58.44	84.39	85.25	79.09
HiFi-We		59.71	73.48	50.79	76.78	83.40	70.94	81.99	84.74	76.94	84.72	89.66	80.48
MMCA	horizontal	69.62	78.04	62.70	74.87	81.25	68.78	82.48	84.69	77.53	86.20	87.99	82.18
	content	69.83	79.19	62.87	76.60	83.22	70.90	82.91	85.86	78.12	86.53	89.03	82.64
	contrast	69.27	72.97	61.99	75.88	77.78	69.68	81.30	81.18	75.99	83.44	83.92	78.69
	coarseness	66.19	68.47	58.32	73.81	73.72	67.15	77.93	75.95	71.84	79.18	78.01	73.49
	vertical	67.27	77.96	62.29	75.28	81.59	69.31	82.49	84.92	77.56	86.19	88.17	82.18
SSPL		58.09	66.64	48.26	61.24	70.10	52.46	64.03	73.36	55.47	65.86	75.23	57.57
MISL		63.75	75.47	55.96	72.72	81.07	66.42	74.66	82.38	68.69	83.38	87.44	78.71
SWMIFL		**84.75**	**85.31**	**79.94**	**88.25**	**89.90**	**84.63**	**96.63**	**94.47**	**95.51**	**97.21**	**94.80**	**96.27**

TABLE XI
CLASSIFICATION RESULTS OF EACH CLASS WITH FIVE TRAINING SAMPLES PER-CLASS ON UNIVERSITY OF PAVIA DATA SET

class	MFASR	ppMLRpr	LBP-ELM	RVCANet	HiFi-We	MMCA					SSPL	MISL	SWMIFL
						horizontal	content	contrast	coarseness	vertical			
Asphalt	**81.58**	76.33	33.68	59.89	60.91	67.02	67.66	60.94	53.34	66.71	48.28	59.37	73.05
Meadows	63.82	75.26	59.51	47.31	72.91	70.50	71.75	76.10	70.92	60.33	66.67	**91.05**
Gravel	85.44	77.00	53.31	82.46	90.31	79.97	80.79	75.90	81.14	78.04	59.80	75.66	90.71
Trees	**93.57**	68.27	49.25	98.30	90.52	76.41	83.78	67.22	56.36	76.77	90.48	72.04	56.79
Metal sheets	99.91	92.63	100.00	100.00	72.46	99.75	**100.00**	98.71	86.79	99.98	99.22	99.78	100.00
Bare soil	78.62	82.06	73.53	51.95	90.37	97.40	99.18	99.71	99.55	98.68	54.89	95.17	**100.00**
Bitumen	99.84	90.16	74.89	96.60	97.66	94.45	93.08	93.86	91.77	93.70	77.98	95.91	**100.00**
Bricks	91.87	77.90	44.70	78.15	75.44	56.32	59.27	64.92	70.04	58.31	45.97	72.14	**99.05**
Shadows	99.24	90.53	29.36	99.89	**100.00**	89.42	93.47	62.67	48.52	91.18	93.93	92.89	98.42

time of SWMIFL is not the longest when compared with comparison methods, and the best classification accuracy of SWMIFL can be achieved in an acceptable time.

2) Experimental Results on the University of Pavia Data Set: On the University of Pavia data set, the dimensionality of intact feature learned by the SWMIFL framework is 12. Table X lists the OA, AA, and κ coefficient of different methods when the training sample number of each class is 5. It can be seen that the proposed SWMIFL provides consistent better results than other methods in different numbers of training samples. In particular, in the case of extremely limited training samples (i.e., three training samples per class), the OA of our SWMIFL is about 15% higher than methods. We can also se that the OA and κ of MFASR are far lower than the SWMIFL method, but its AA is closer to the proposed method. The reason is that the classification accuracy of classes with fewer samples (i.e., "Gravel," "Trees," "Metal sheets," "Bitumen,"

and "Shadows") are high, and the classification accuracies of classes with larger samples (i.e., "Meadows" and "Bare soil") are low. This can be from Table XI, where the accuracy of each class is listed. Fig. 6 counts the number of classes that each method outperforms other methods, where our method outperforms other methods on six out of nine classes. In general, by exploiting multiview intact information, our SWMIFL method provides the highest OA, AA, and κ coefficient. Fig. 9 shows the classification maps of all the methods when the training sample number of each class is 5. It can be seen that the SWMIFL method shows relatively better results than other methods in terms of consistent classification results with little salt and pepper noise. In particular, the classification map of the proposed method has overwhelming advantage on class "Meadows."

3) Experimental Results on the Salinas Data Set: On the Salinas data set, the dimensionality of intact feature learned
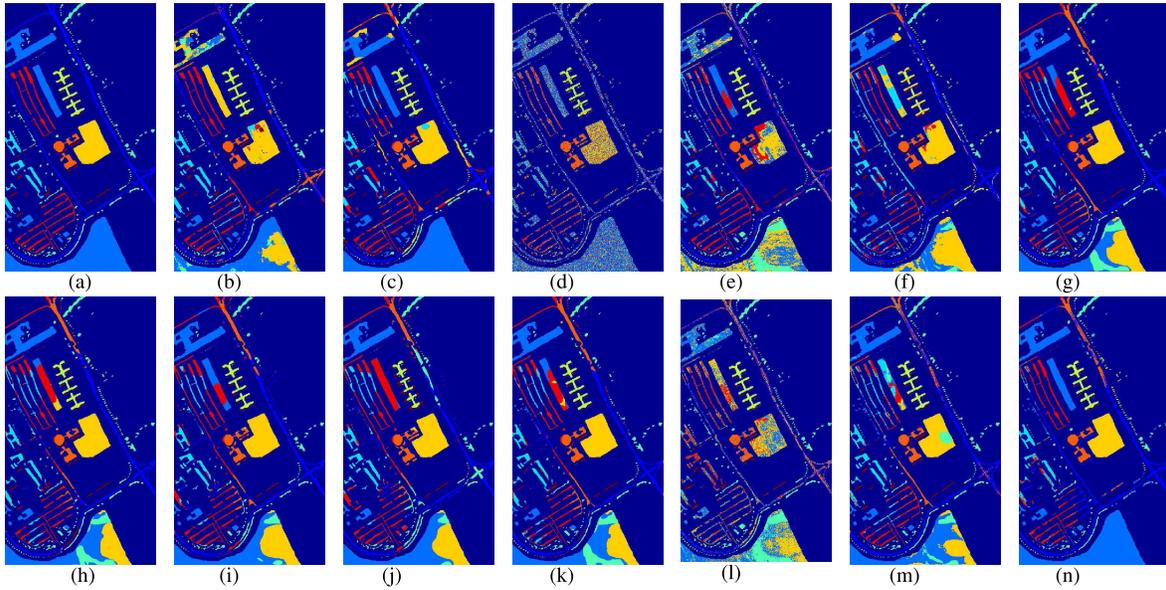
Fig. 9. Classification maps of all the compared methods with five training samples per-class on University of Pavia data set. (a) Groundtruth. (b) MFASR. (c) ppMLRpr. (d) LBP-ELM. (e) RVCANet. (f) HiFi-We. (g) Horizontal. (h) Content. (i) Contrast. (j) Coarseness. (k) Vertical. (l) SSPL. (m) MISL. (n) SWMIFL.

TABLE XII

CLASSIFICATION RESULTS WITH DIFFERENT TRAINING SAMPLES ON SALINAS DATA SET

Number (per-class)		3			5			7			10		
Accuracy		OA	AA	κ	OA	AA	κ	OA	AA	κ	OA	AA	κ
MFASR		84.63	90.81	82.93	88.47	93.27	87.19	88.69	93.91	87.42	91.29	95.58	90.32
ppMLRpr		83.16	86.36	81.36	86.10	91.31	84.61	89.72	94.24	88.61	90.15	92.49	89.07
LBP-ELM		71.95	77.03	69.06	79.08	83.68	76.92	83.70	87.14	81.97	88.76	91.09	87.55
RVCANet		84.13	90.95	82.30	85.38	90.69	83.64	86.09	93.97	84.58	86.49	93.58	85.01
HiFi-We		82.70	89.18	80.85	88.16	92.93	86.87	89.15	93.52	87.97	89.86	94.41	88.75
MMCA	horizontal	86.05	93.55	84.61	89.58	94.99	88.46	92.06	96.05	91.20	94.09	97.14	93.44
	content	86.46	93.69	85.06	89.74	95.07	88.65	92.20	96.12	91.35	94.19	97.12	93.55
	contrast	85.15	91.91	83.59	88.48	93.33	87.24	90.77	94.64	89.77	92.89	95.66	92.10
	coarseness	85.27	91.59	83.75	88.99	93.35	87.82	91.22	94.75	90.27	93.78	96.29	93.10
	vertical	85.26	92.79	83.73	88.65	94.37	87.44	91.25	95.49	90.30	93.12	96.32	92.36
SSPL		81.41	87.91	79.33	81.21	88.78	79.17	82.63	90.10	80.72	83.76	91.18	81.97
MISL		78.52	88.23	76.32	89.22	93.98	88.00	90.16	94.50	89.12	89.52	93.69	88.28
SWMIFL		**88.40**	**94.28**	**87.17**	**92.39**	**96.05**	**91.56**	**93.24**	**97.24**	**92.49**	**96.50**	**98.43**	**96.11**

TABLE XIII

CLASSIFICATION RESULTS OF EACH CLASS WITH FIVE TRAINING SAMPLES PER-CLASS ON SALINAS DATA SET

class	MFASR	ppMLRpr	LBP-ELM	RVCANet	HiFi-We	MMCA					SSPL	MISL	SWMIFL
						horizontal	content	contrast	coarseness	vertical			
Brocoli-green-weeds-1	99.48	97.11	97.09	**99.95**	98.87	97.88	97.90	95.84	97.35	96.41	99.05	96.70	98.43
Brocoli-green-weeds-2	96.37	98.94	85.59	95.08	99.48	99.79	99.70	99.80	99.45	99.72	96.64	99.15	**99.91**
Fallow	94.07	74.51	88.44	84.85	95.88	98.54	98.63	98.50	99.03	98.68	73.76	98.28	**99.07**
Fallow-rough-plow	**99.70**	97.79	95.91	98.71	98.13	98.52	98.82	96.14	94.95	98.43	98.58	98.77	97.99
Fallow-smooth	98.51	90.50	69.59	**99.07**	98.37	95.70	96.10	90.43	90.83	95.44	96.07	97.96	96.40
Stubble	99.65	97.63	85.77	99.95	98.50	99.87	99.88	99.46	99.32	99.74	99.54	99.75	**99.98**
Celery	99.24	96.74	94.75	99.75	98.57	99.86	**99.94**	98.77	98.49	**99.94**	99.35	99.25	99.84
Grapes-untrained	72.30	65.16	64.42	84.58	66.75	62.02	62.44	61.67	62.40	59.13	56.17	73.63	**75.57**
Soil-vinyard-develop	99.64	98.38	80.11	97.45	99.66	100.00	100.00	99.80	100.00	99.92	95.93	99.95	99.99
Corn-senesced-green-weeds	84.37	89.12	80.10	85.89	81.37	91.58	91.32	91.59	91.74	90.54	79.28	83.67	**92.68**
Lettuce-romaine-4wk	97.07	98.14	87.16	96.34	94.31	98.17	96.61	97.69	98.75	96.07	93.80	97.57	**99.28**
Lettuce-romaine-5wk	**100.00**	97.15	83.87	**100.00**	99.13	99.91	99.95	99.18	94.61	99.93	98.19	99.28	99.59
Lettuce-romaine-6wk	97.49	93.47	79.68	**99.34**	96.38	94.51	94.34	87.50	88.06	93.21	97.67	94.30	94.14
Lettuce-romaine-7wk	96.01	93.24	85.83	91.75	91.36	95.78	**97.12**	90.90	88.33	97.10	92.18	94.74	96.90
Vinyard-untrained	74.66	75.28	72.86	37.66	82.08	90.44	**90.92**	88.49	92.08	89.74	58.68	74.23	88.97
Vinyard-vertical-trellis	83.77	97.73	87.81	80.60	88.07	97.32	97.39	97.59	**98.13**	95.88	85.57	96.46	97.97

by SWMIFL framework is 12. Table XII lists the OA, AA, and κ coefficients of all methods with different numbers of training samples. The classification results with five training samples per class are shown in Table XIII. Fig. 6 lists the number of classes that each method outperforms other methods. Due to the learned multiview intact feature, the proposed method can obtain the best OA, AA, and κ coefficients in all experiments. SWMIFL acquires the best classification results on six classes. It is worth noting that, the improved classification accuracy of SWMIFL on Salinas data set is lower than its on Indian Pines data set and University of Pavia

data set. One possible reason is that the classification edges between different classes on Salinas data set are smoother and flatter, which is easier to classify than complex edges. In this case, although the improvement of SWMIFL is not as obvious as its on other two data sets, it still achieves the best classification performance and can achieve OA more than 88% with only three training samples each class. These can prove the classification performance of SWMIFL sufficiently. Fig. 10 shows the classification map of different methods. It can be seen that most of methods produce poor results on the large classes "Grapes-untrained" and "Vinyard-untrained."
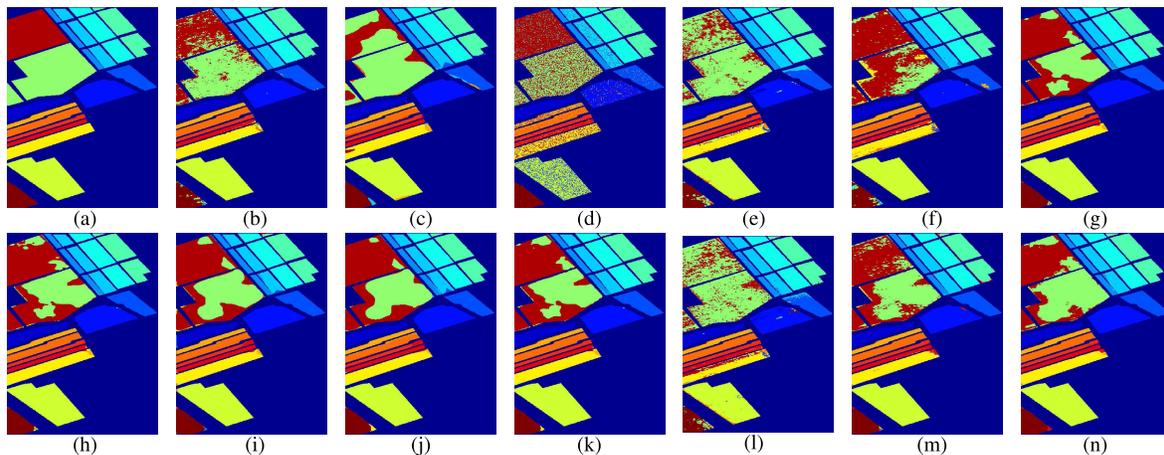
Fig. 10. Classification maps of all the compared methods with five training samples per-class on Salinas data set. (a) Groundtruth. (b) MFASR. (c) ppMLRpr. (d) LBP-ELM. (e) RVCANet. (f) HiFi-We. (g) Horizontal. (h) Content. (i) Contrast. (j) Coarseness. (k) Vertical. (l) SSPL. (m) MISL. (n) SWMIFL.

Because these two classes are spatially adjacent, their spectral characteristics show certain similarity. It is very difficult to classify them. Notwithstanding, our SWMIFL produces relatively better results on these two classes even if only five training samples per class is used.

IV. CONCLUSION

In this article, a spatial window-based multiview intact feature learning model is proposed for HSI classification. The multiview intact features that contain diversity information of HSIs are learned. Experimental results on three real-world data sets demonstrate that the proposed method reveals better classification when the number of training sample is small. However, the proposed method can be further improved. For example, the new training samples are unlabeled samples in the spatial windows whose labels classified by the intact feature are consistent with labels of training samples in the same windows. How to optimize this process to improve the reliability of the selected new training samples is a problem that needs to be solved.

REFERENCES

[1] G. Cheng, F. Zhu, S. Xiang, Y. Wang, and C. Pan, "Semisupervised hyperspectral image classification via discriminant analysis and robust regression," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 595–608, Feb. 2016.

[2] L. Fang, C. Wang, S. Li, and J. A. Benediktsson, "Hyperspectral image classification via multiple-feature-based adaptive sparse representation," *IEEE Trans. Instrum. Meas.*, vol. 66, no. 7, pp. 1646–1657, Jul. 2017.

[3] B. Wu, Y. Zhu, X. Huang, and J. Li, "Generalization of spectral fidelity with flexible measures for the sparse representation classification of hyperspectral images," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 52, pp. 275–283, Oct. 2016.

[4] M. Fauvel, Y. Tarabalka, J. A. Benediktsson, J. Chanussot, and J. C. Tilton, "Advances in spectral-spatial classification of hyperspectral images," *Proc. IEEE*, vol. 101, no. 3, pp. 652–675, Mar. 2013.

[5] H. Li, Y. Song, and C. L. P. Chen, "Hyperspectral image classification based on multiscale spatial information fusion," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 9, pp. 5302–5312, Sep. 2017.

[6] Y. Shao, N. Sang, C. Gao, and L. Ma, "Spatial and class structure regularized sparse representation graph for semi-supervised hyperspectral image classification," *Pattern Recognit.*, vol. 81, pp. 81–94, Sep. 2018.

[7] B. Pan, Z. Shi, and X. Xu, "Hierarchical guidance filtering-based ensemble classification for hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 7, pp. 4177–4189, Jul. 2017.

[8] J. Li, M. Khodadadzadeh, A. Plaza, X. Jia, and J. M. Bioucas-Dias, "A discontinuity preserving relaxation scheme for spectral-spatial hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 625–639, Feb. 2016.

[9] Y.-Q. Zhao, L. Zhang, and S. G. Kong, "Band-subset-based clustering and fusion for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 2, pp. 747–756, Feb. 2011.

[10] R. Ji, Y. Gao, R. Hong, Q. Liu, D. Tao, and X. Li, "Spectral–spatial constraint hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 3, pp. 1811–1824, Mar. 2014.

[11] C. Chen, W. Li, H. Su, and K. Liu, "Spectral-spatial classification of hyperspectral image based on kernel extreme learning machine," *Remote Sens.*, vol. 6, no. 6, pp. 5795–5814, Jun. 2014.

[12] W. Fu, S. Li, and L. Fang, "Spectral-spatial hyperspectral image classification via superpixel merging and sparse representation," in *Proc. IEEE Int. Geosci. Remote Sens. Symp. (IGARSS)*, Jul. 2015, pp. 4971–4974.

[13] H. Huang and X. L. Zheng, "Hyperspectral image classification with combination of weighted spatial-spectral and KNN," *Opt. Precis. Eng.*, vol. 24, no. 4, pp. 873–881, 2016.

[14] O. Rajadell, P. Garcia-Sevilla, and F. Pla, "Spectral–spatial pixel characterization using Gabor filters for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 4, pp. 860–864, Jul. 2013.

[15] M. Cui, S. Prasad, W. Li, and L. M. Bruce, "Locality preserving genetic algorithms for spatial–spectral hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 6, no. 3, pp. 1688–1697, Jun. 2013.

[16] B. B. Damodaran, R. R. Nidamanuri, and Y. Tarabalka, "Dynamic ensemble selection approach for hyperspectral image classification with joint spectral and spatial information," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 8, no. 6, pp. 2405–2417, Jun. 2015.

[17] B. Liu, X. Yu, A. Yu, P. Zhang, G. Wan, and R. Wang, "Deep few-shot learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 4, pp. 2290–2304, Oct. 2018.

[18] A. Kianisarkaleh and H. Ghassemian, "Spatial-spectral locality preserving projection for hyperspectral image classification with limited training samples," *Int. J. Remote Sens.*, vol. 37, no. 21, pp. 5045–5059, Nov. 2016.

[19] X. Ma, S. Ji, J. Wang, J. Geng, and H. Wang, "Hyperspectral image classification based on two-phase relation learning network," *IEEE Trans. Geoence Remote Sens.*, vol. 57, no. 12, pp. 10398–10409, Dec. 2019.

[20] L. Zhu, Y. Chen, P. Ghamisi, and J. A. Benediktsson, "Generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 9, pp. 5046–5063, Sep. 2018.

[21] B. Pan, Z. Shi, and X. Xu, "R-VCANet: A new deep-learning-based hyperspectral image classification method," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 5, pp. 1975–1986, May 2017.

[22] Z. Zhong, J. Li, Z. Luo, and M. Chapman, "Spectral–spatial residual network for hyperspectral image classification: A 3-D deep learning framework," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 847–858, Feb. 2018.

[23] B. Pan, Z. Shi, and X. Xu, "MugNet: Deep learning for hyperspectral image classification using limited samples," *ISPRS J. Photogramm. Remote Sens.*, vol. 145, pp. 108–119, Nov. 2018.

[24] 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.

[25] B.-C. Kuo, C.-H. Li, and J.-M. Yang, "Kernel nonparametric weighted feature extraction for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 47, no. 4, pp. 1139–1155, Apr. 2009.

[26] Y.-L. Chang, J.-N. Liu, C.-C. Han, and Y.-N. Chen, "Hyperspectral image classification using nearest feature line embedding approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 1, pp. 278–287, Jan. 2014.

[27] L. Shen, Z. Zhu, S. Jia, J. Zhu, and Y. Sun, "Discriminative Gabor feature selection for hyperspectral image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 10, no. 1, pp. 29–33, Jan. 2013.

[28] B.-C. Kuo, H.-H. Ho, C.-H. Li, C.-C. Hung, and J.-S. Taur, "A kernel-based feature selection method for SVM with RBF kernel for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 1, pp. 317–326, Jan. 2014.

[29] Y. Qian, M. Ye, and J. Zhou, "Hyperspectral image classification based on structured sparse logistic regression and three-dimensional wavelet texture features," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 4, pp. 2276–2291, Apr. 2013.

[30] J. Xia, M. Dalla Mura, J. Chanussot, P. Du, and X. He, "Random subspace ensembles for hyperspectral image classification with extended morphological attribute profiles," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 4768–4786, Sep. 2015.

[31] D. Tuia, R. Flamary, and N. Courty, "Multiclass feature learning for hyperspectral image classification: Sparse and hierarchical solutions," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 272–285, Jul. 2015.

[32] W. Zhao and S. Du, "Spectral–spatial feature extraction for hyperspectral image classification: A dimension reduction and deep learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 8, pp. 4544–4554, Aug. 2016.

[33] W. Li, G. Wu, F. Zhang, and Q. Du, "Hyperspectral image classification using deep pixel-pair features," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 844–853, Feb. 2017.

[34] 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.

[35] X. Xu, J. Li, X. Huang, M. Dalla Mura, and A. Plaza, "Multiple morphological component analysis based decomposition for remote sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 5, pp. 3083–3102, May 2016.

[36] W. Di and M. M. Crawford, "View generation for multiview maximum disagreement based active learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 50, no. 5, pp. 1942–1954, May 2012.

[37] M. Volpi, G. Matasci, M. Kanevski, and D. Tuia, "Semi-supervised multiview embedding for hyperspectral data classification," *Neurocomputing*, vol. 145, pp. 427–437, Dec. 2014.

[38] M. Zhang, W. Li, and Q. Du, "Diverse region-based CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 27, no. 6, pp. 2623–2634, Jun. 2018.

[39] Y. Zhao et al., "Multi-view manifold learning with locality alignment," *Pattern Recognit.*, vol. 78, pp. 154–166, Jun. 2018.

[40] X. Xu, J. Li, and S. Li, "Multiview intensity-based active learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 2, pp. 669–680, Feb. 2018.

[41] A. Appice, P. Guccione, and D. Malerba, "A novel spectral-spatial co-training algorithm for the transductive classification of hyperspectral imagery data," *Pattern Recognit.*, vol. 63, pp. 229–245, Mar. 2017.

[42] W. Di and M. M. Crawford, "Active learning via multi-view and local proximity co-regularization for hyperspectral image classification," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 618–628, Jun. 2011.

[43] W. Di and M. M. Crawford, "Multi-view adaptive disagreement based active learning for hyperspectral image classification," in *Proc. IEEE Int. Geosci. Remote Sens. Symp.*, Jul. 2010, pp. 1374–1377.

[44] X. Zhou, S. Prasad, and M. M. Crawford, "Wavelet-domain multiview active learning for spatial–spectral hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 9, pp. 4047–4059, Sep. 2016.

[45] P. Chen, L. Jiao, F. Liu, J. Zhao, and Z. Zhao, "Dimensionality reduction for hyperspectral image classification based on multiview graphs ensemble," *J. Appl. Remote Sens.*, vol. 10, no. 3, Jul. 2016, Art. no. 030501.

[46] R. J. Samworth, "Optimal weighted nearest neighbour classifiers," *Ann. Statist.*, vol. 40, no. 5, pp. 2733–2763, Oct. 2012.

[47] C. Xu, D. Tao, and C. Xu, "Multiview intact space learning," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 37, no. 12, pp. 2531–2544, Dec. 2015.

[48] 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.

[49] C.-C. Chang and C.-J. Lin, "LIBSVM: A library for support vector machines," *ACM Trans. Intell. Syst. Technol.*, vol. 2, no. 3, pp. 27:1–27:27, 2011.

[50] Q.-W. Wang, Y.-F. Li, and Z.-H. Zhou, "Partial label learning with unlabeled data," in *Proc. 28th Int. Joint Conf. Artif. Intell.*, Aug. 2019, pp. 3755–3761.

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 an Assistant Professor with the School of Computer Science and Information Engineering, Hubei University, Wuhan. Her research interests include computer vision, pattern recognition, machine learning.

Yiu-ming Cheung (Fellow, IEEE) received the Ph.D. degree from the Department of Computer Science and Engineering, The Chinese University of Hong Kong, Hong Kong.
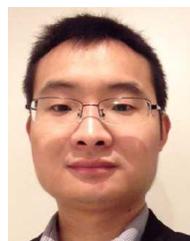
He is currently a Full Professor with the Department of Computer Science, Hong Kong Baptist University, Hong Kong. His research interests include machine learning, pattern recognition, visual computing, and optimization.

Prof. Cheung is a fellow of IET, BCS, and RSA, and a Distinguished Fellow of IETI. He is the Founding Chair of the Computational Intelligence Chapter of the IEEE Hong Kong Section, and the Chair of the Technical Committee on Intelligent Informatics of the IEEE Computer Society. He serves as an Associate Editor for the IEEE TRANSACTIONS ON NEURAL NETWORKS AND LEARNING SYSTEMS, the IEEE TRANSACTIONS ON CYBERNETICS, PATTERN RECOGNITION, KNOWLEDGE AND INFORMATION SYSTEMS, NEUROCOMPUTING, and so on.

Xinge You received the Ph.D. degree from the Department of Computer Science, Hong Kong Baptist University, Hong Kong, in 2004.

He is a Professor with the School of Electronics Information and Communications, Huazhong University of Science and Technology, Wuhan. His research interests include data mining, pattern recognition, machine earning, and computer vision.

Qinmu Peng received the Ph.D. degree from the Department of Computer Science, Hong Kong Baptist University, Hong Kong, in 2015.

He is an Assistant Professor with the School of Electronic Information and Communications, Shenzhen Research Institute of Huazhong University of Science and Technology, and Huazhong University of Science and Technology, Wuhan, China. His research interests include medical image processing, pattern recognition, machine earning, and computer vision.

Jiangtao Peng received the B.S. and M.S. degrees from Hubei University, Wuhan, China, in 2005 and 2008, and the Ph.D. degree from the Institute of Automation, Chinese Academy of Sciences, Beijing, China, in 2011.

He is a Professor with the Faculty of Mathematics and Statistics, Hubei University. His research interests include machine learning and hyperspectral image processing.

Yufeng Shi received the B.S. degree from the School of Optoelectronic Science and Engineering, University of Electronic Science and Technology of China, in 2017. He is pursuing the Ph.D. degree with the School of Electronic Information and Communications, Huazhong University of Science and Technology.

His research interests include multimedia information retrieval and machine learning.

Peipei Yuan received the M.S. degree in information and computation science from Huazhong Agricultural University, Wuhan, China, in 2014. She is pursuing the Ph.D. degree with the School of Electronic Information and Communications, Huazhong University of Science and Technology, Wuhan.

Her research interests include computer vision, pattern recognition, machine learning and data mining.