

Research Article

Spectral-Spatial Hyperspectral Image Semisupervised Classification by Fusing Maximum Noise Fraction and Adaptive Random Multigraphs

Eryang Chen ^{1,2,3,4} Ruichun Chang ^{4,5} Kaibo Shi ² Ansheng Ye,^{1,2} Fang Miao,³ Jianghong Yuan,⁶ Ke Guo,⁴ Youhua Wei,⁴ and Yiping Li^{1,4}

¹College of Geophysics, Chengdu University of Technology, Chengdu 610059, China

²School of Electronic Information and Electrical Engineering, Chengdu University, Chengdu 610106, China

³Key Laboratory of Pattern Recognition and Intelligent Information Processing of Sichuan, Chengdu University, Chengdu 610106, China

⁴Geomathematics Key Laboratory of Sichuan Province, Chengdu University of Technology, Chengdu 610059, China

⁵Digital Hu Line Research Institute, Chengdu University of Technology, Chengdu 610059, China

⁶School of Intelligent Engineering, Sichuan Changjiang Vocational College, Chengdu 610106, China

Correspondence should be addressed to Ruichun Chang; 2293721870@qq.com and Kaibo Shi; skbs111@163.com

Received 12 March 2021; Accepted 15 June 2021; Published 24 June 2021

Academic Editor: Zi-Peng Wang

Copyright © 2021 Eryang Chen et al. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

Hyperspectral images (HSIs) contain large amounts of spectral and spatial information, and this provides the possibility for ground object classification. However, when using the traditional method, achieving a satisfactory classification result is difficult because of the insufficient labeling of samples in the training set. In addition, parameter adjustment during HSI classification is time-consuming. This paper proposes a novel fusion method based on the maximum noise fraction (MNF) and adaptive random multigraphs for HSI classification. Considering the overall spectrum of the object and the correlation of adjacent bands, the MNF was utilized to reduce the spectral dimension. Next, a multiscale local binary pattern (LBP) analysis was performed on the MNF dimension-reduced data to extract the spatial features of different scales. The obtained multiscale spatial features were then stacked with the MNF dimension-reduced spectral features to form multiscale spectral-spatial features (SSFs), which were sent into the RMG for HSI classification. Optimal performance was obtained by fusion. For all three real datasets, our method achieved competitive results with only 10 training samples. More importantly, the classification parameters corresponding to different hyperspectral data can be automatically optimized using our method.

1. Introduction

Due to the advancements in remote sensing technology, hyperspectral images (HSIs) are containing increasing spectral and spatial information (SSI), resulting in their extensive use in domains, such as forest inventory [1], urban area monitoring [2], road extraction [3], geological surveys [4], precision agriculture [5], environmental protection [6], military applications [7], hydrocarbon detection [8], oil reservoir exploration [9], and lake sediment analysis [10]. HSI classification is a crucial research topic related to these applications. In contrast to those of Synthetic Aperture Radar

(SAR) [11] or RGB images [12], the two main challenges associated with HSI classification are the high dimensionality of the dataset and the redundancy of spectral information.

Many approaches for HSI dimension reduction have been proposed. The most frequently used methods are principal component analysis (PCA) [13, 14], independent component analysis (ICA) [15], maximum noise fraction (MNF) [16], linear discriminant analysis [17], and the deep learning approach [18]. Uddin et al. [13] proposed an information-theoretic normalized-based PCA for HSI classification. Fu et al. [14] proposed a segmented PCA-based algorithm for HSI classification.

However, because the noise in hyperspectral data can easily mask subtle hyperspectral features, careful noise removal is required to extract useful information. This phenomenon is problematic for PCA, which maximizes the variance of the orthogonal set of the projection sample and vector. Unlike PCA, MNF aims to maximize the signal-to-noise ratio (SNR) rather than the variance. MNF is worth considering in HSI classification because removing the noise is effective during dimensionality reduction.

HSI classification can be significantly improved by employing an appropriate object-based [19], spectral-spatial fusion-based approach [20], decision fusion-based method [21], or deep learning-based method [22]. However, two challenges are unavoidable for most researchers. On the one hand, the traditional depth learning-based approaches rely heavily on a large amount of labeled data to achieve competitive results, whereas HSI annotation is expensive and time-consuming because it requires expert knowledge and skills. On the other hand, many methods require the use of a large number of parameter settings during the experiment, which involves expert experience [23]. HSI classification has been widely studied, and various classification methods have been adopted, namely, support vector machine (SVM) [24], random forests [25], neural networks [26], low-rank representation [27], sparse representation [28], deep learning methods [29, 30], and meta-learning methods [31]. Xu et al. [29] directly used the random patches extracted from the HSI image as the convolution kernel, without any training, to improve the classification efficiency. Yin et al. [30] proposed a CapsNet-based alternative data-driven HSI classification model. The aforementioned results demonstrate that the proposed method can achieve ideal results when the training samples are sufficient.

To integrate the spatial information, researchers have proposed many spectral-spatial feature extraction (FE) techniques, such as the extinction profile (EP) [32] and local binary patterns (LBP) [33]. In contrast with the EP features, LBP features facilitate the mining of the HSI texture information, such as global contrast information and texture depth [34]. In this study, we adopted the LBP features as spatial features.

However, when the labeled samples in the training set are insufficient, the classification accuracy achieved by the traditional methods significantly reduces [35] because of the so-called Hughes effect or the curse of dimensionality [36]. Moreover, many methods require a series of manual parameter settings. For instance, to extract spatial features, researchers must select an appropriate window size for the neighborhood. The selection is time-consuming [37] and requires expertise [38]. Since 2010, ensemble learning methods for HSI classification have received significant attention because of their dependence on limited training samples [39]. Many methods have been developed, such as support vector machines [40], boosting [41], segmentation-based methods [19], unsupervised methods [42], and semisupervised methods [43]. However, the graph-based semisupervised method is rarely considered in HSI classification [44, 45].

Motivated by the aforementioned discussions, by combining the SSI, we proposed a new spectral-spatial HSI

semisupervised classifier based on MNF and adaptive random multigraphs (SS-MNF-ARMG). The primary contributions of this study are as follows:

- (1) A novel spectral-spatial HSI semisupervised classification framework was developed. Because of the adaptive properties, the optimal parameters can be determined without artificial auxiliaries.
- (2) By introducing the MNF, the noise in the HSI can be removed more efficiently during dimensionality reduction. On the basis of dimension-reduced HSI, the SSI is combined, which can degrade the curse of dimensionality.
- (3) In contrast with several studies, SS-MNF-ARMG can achieve competitive performance for three real HSI datasets while leveraging tiny labeled samples, which is further improved by introducing RMG in a new mode.

2. Relevant Work

2.1. MNF. Let X be the HSI data, and S and N are the signal and noise part of X , respectively. The goal of MNF is to seek out a linear transformation matrix W to maximize the SNR of the transformed data. Assume that S and N are uncorrelated; then X can be represented as [43]

$$X = S + N. \quad (1)$$

And, the covariance matrix (CM) of X can be obtained by

$$\text{cov}(X) = \sum_X = \sum_S + \sum_N, \quad (2)$$

where \sum_S and \sum_N are the CMs of S and N , respectively. The MNF transform can be expressed as

$$Y = W^T X, \quad (3)$$

where Y is the MNF result of X , W are the eigenvectors associated with the L largest eigenvalues of $\sum_N^{-1} \sum_X$, and L is the number of MNF principal components.

Then the SNR of each $y_i \in Y$ can be described as

$$\frac{\text{cov}\{w_i^T S\}}{\text{cov}\{w_i^T N\}} = \frac{w_i^T \sum_S w_i}{w_i^T \sum_X w_i}, \quad (4)$$

where $\text{cov}\{\cdot\}$ computes the variance and w_i is the i th component in W . Then we can obtain W by solving the following problem:

$$\arg \max_W \frac{W^T \sum_S W}{W^T \sum_N W} = \arg \max_W \frac{W^T \sum_X W}{W^T \sum_N W} - 1. \quad (5)$$

2.2. LBP. For each pixel in HSI, a P -bit binary code and a new matrix with the new value (binary to decimal value) are generated by thresholding adjacent pixel values. The LBP operator can be calculated by

$$\text{LBP}_{i,r}(N_c) = \sum_{i=0}^{P-1} g(N_i - N_c)2^i, \quad (6)$$

where P is the number of neighbors represented on a circle with the radius r , and N_c and N_i , respectively, represent the gray level intensity values of the center and the i th neighbour. The binary threshold function $g(x)$ is described as

$$g(x) = \begin{cases} 0, & x < 0, \\ 1, & x \geq 0. \end{cases} \quad (7)$$

Take the 10th band of Indian Pines HSI as an example; the procedure of LBP is shown in Figure 1. As shown in Figure 1, for a given center pixel in a 3×3 window, binary labels (“0” or “1”) are assigned to adjacent pixels according to whether the gray value of the center pixel is large. Starting at the top left corner, all binary codes are joined clockwise to produce an 8-bit binary number. The resulting binary number is called LBP code. The results show that LBP algorithm has significant rotation invariance and gray invariance [46], and it can be effectively applied to HSI classification [47].

2.3. RMG. Given the HSI dataset $D = X_{lab} \cup X_{ulab}$ is comprised of l labeled data $X_{lab} = \{(x_i)\}_{i=1}^l$ and u unlabeled data $X_{ulab} = \{(x_i)\}_{i=l+1}^{l+u}$, then a weighted graph can be obtained. The vertices in the graph consist of X_{lab} and X_{ulab} . Weighted edges, which can be defined as a matrix $W = \mathbb{R}^{(l+u) \times (l+u)}$, represent the similarities between associated nodes. For a c -class classification problem, it can be defined as a quadratic optimization problem [43]:

$$\min_F \text{tr}((F - Y)^T C (F - Y)) + \text{tr}(F^T L F), \quad (8)$$

where $\text{tr}(\cdot)$ is the trace function and $F \in \mathbb{R}^{(l+u) \times c}$ is the prediction matrix to be learned. The indicator vector $y_i \in Y = (y_1, y_2, \dots, y_l, 0, \dots, 0)^T \in \mathbb{R}^{(l+u) \times c}$ is the label vector corresponding to x_i and 0 is the 0 vector. In addition, we can obtain y_{ij} by the following equation:

$$y_{ij} = \begin{cases} 1, & \text{if } x_i \text{ belongs to the } j\text{th class,} \\ 0, & \text{otherwise.} \end{cases} \quad (9)$$

Then each x_i can be classified to the j th class if F_{ij} is the largest one in the i th row of F which can be described as $y_i = \arg \max_j F_{ij}$, $j = 1, \dots, c$. $C \in \mathbb{R}^{(l+u) \times (l+u)}$ is a diagonal matrix, and its element c_i can be calculated as

$$c_i = \begin{cases} \alpha_l, & 1 \leq i \leq l, \alpha_l > 0, \\ \alpha_u, & l + 1 \leq i \leq l + u, \alpha_u \geq 0, \end{cases} \quad (10)$$

where α_l and α_u are two parameters.

The popular choice for L is the graph Laplacian [48], which is defined as

$$L = D - W, \quad (11)$$

where W is the weight matrix of the graph, which can be formulated by Gaussian kernel as

$$w_{ij} = e^{-\left(\|x_i - y_i\|^2 / 2\sigma^2\right)} \quad (12)$$

where σ is the kernel width parameter to be adjusted. And the diagonal matrix D is the row sum of W .

However, it is difficult to discover the neighborhood structure inherent in the graph and learn the proper compact representation automatically. To solve this problem, researchers have proposed the anchor graph algorithm (AGA) [49] and multiview anchor graph algorithm [50]. The AGA extrapolates the Laplacian eigenvector of the graph to the eigenfunction, allowing constant time hashing of new data points. Then the hierarchical threshold learning method is used to make each feature function generate more than 1 bit to improve the search accuracy. And the label prediction function in the AGA can be expressed as

$$f(x_i) = \sum_{j=1}^m P_{ij} f(a_j), \quad (13)$$

where P_{ij} is the data-adaptive weight, and each a_j in $A = \{a_j\}_{j=1}^m$ is an anchor point. By this formula, the solution space of the unknown labels can be reduced from a larger space to a smaller one. The centers of the K-means cluster are selected as anchor points because these centers have a powerful representation that covers the entire dataset. In this paper, we use Local Anchor Embedding (LAE) [49] to calculate the anchor points. Figure 2 shows the flowchart of RMG.

3. Proposed Method

The SS-MNF-ARMG framework is shown in Figure 3. Because of their adaptive properties, the optimal parameters can be determined without artificial auxiliaries. The framework comprises three main modules:

- (1) Preprocessing the HSI image by applying MNF, the noise in the HSI can be removed effectively during dimension reduction. This result avoids the problem of the dimension.
- (2) Through LBP, spatial vectors corresponding to different neighborhood regions are obtained. These spatial vectors are stacked with the spectral vector, respectively, and a series of spectral-spatial stacked feature information can be obtained.
- (3) For classification and decision fusion, a set of necessary parameters for the RMG is established and the spectral-spatial feature information is integrated into the RMG. Next, a set of classification results with different accuracies are obtained, based on which the optimal classification results are obtained through decision fusion. In addition, by injecting randomness into the graph in the RMG, overfitting due to the limited training sample can be avoided.

The proposed SS-MNF-ARMG algorithm is summarized in Algorithm 1.

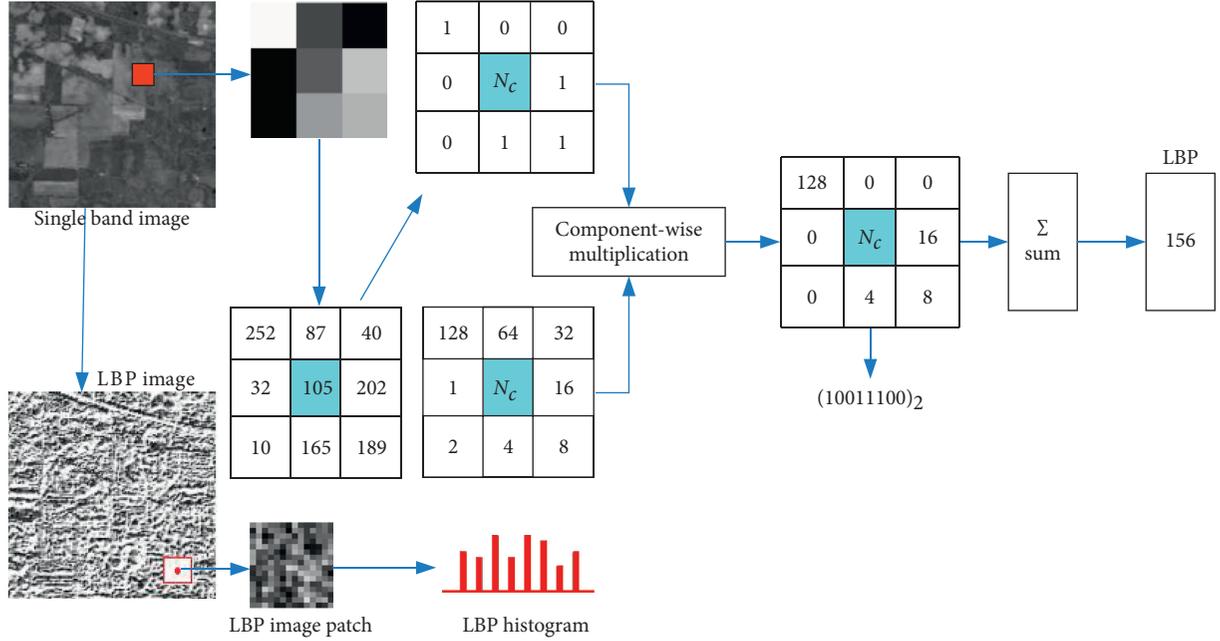


FIGURE 1: Implementation of LBP FE.

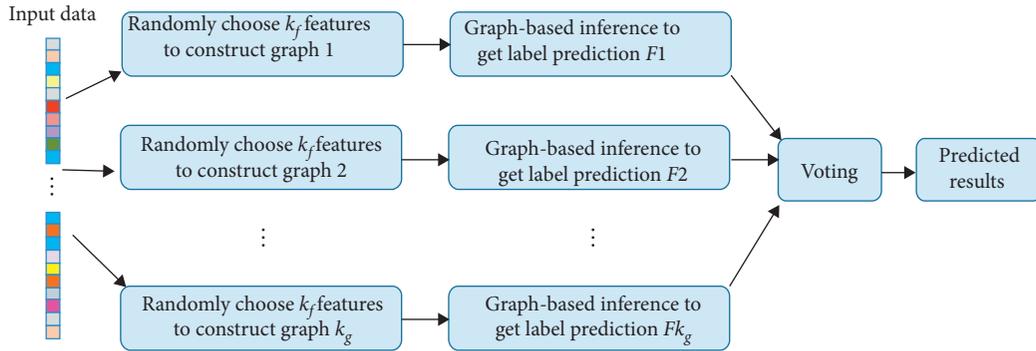


FIGURE 2: Flowchart of the RMG.

4. Experimental Results and Analysis

4.1. Experimental Datasets. Three hyperspectral datasets were employed to evaluate the performance of the SS-MNF-ARMG.

Indian Pines: this scene was gathered by the AVIRIS sensor over the Indian Pines test site in northwestern Indiana and consists of 145×145 pixels and 224 spectral reflectance bands in the wavelength range 0.4 to $2.5 \mu\text{m}$. This scene, which includes 16 different ground truths, contains two-thirds of agriculture and one-third of forest or other natural perennial vegetation. The number of bands was reduced to 200 by removing the 24 water absorption bands.

Pavia University: this scene was acquired by the ROSIS sensor during a flight campaign over Pavia, northern Italy. The number of spectral bands was 103 at Pavia University. It is a 610×340 pixel image containing nine different ground objects with a geometric resolution of 1.3 meters.

Salinas: this scene was collected by the 224-band AVIRIS sensor over the agricultural area of Salinas Valley, California,

and is characterized by high spatial resolution (3.7 m pixels). After discarding 20 water absorption bands, the size of this data image was 512×217 , with 204 bands. Salinas ground truth contains 16 classes, including vegetables, bare soils, and vineyard fields.

4.2. Analysis of Experimental Parameters. Parameters can affect the classification accuracy, such as the number of spectral bands ban , dN , um , the patch size W , the number of sampling points nr , the number of graphs kg , and the number of features kf . The adaptive properties of the proposed SS-MNF-ARMG are such that most optimal parameters for W , nr , kg , and kf can be determined without artificial auxiliaries.

Based on the existing result [44], we let $W \in \{15, 19, 23\}$, $nr \in \{15, 18, 21\}$, $kg \in \{19, 25, 31\}$, and $kf \in \{2\sqrt{\dim(V_{spe_spa})}, 4\sqrt{\dim(V_{spe_spa})}, 8\sqrt{\dim(V_{spe_spa})}\}$, where $\dim(V_{spe_spa})$ represents the dimensionality of the spectral-spatial vector V_{spe_spa} .

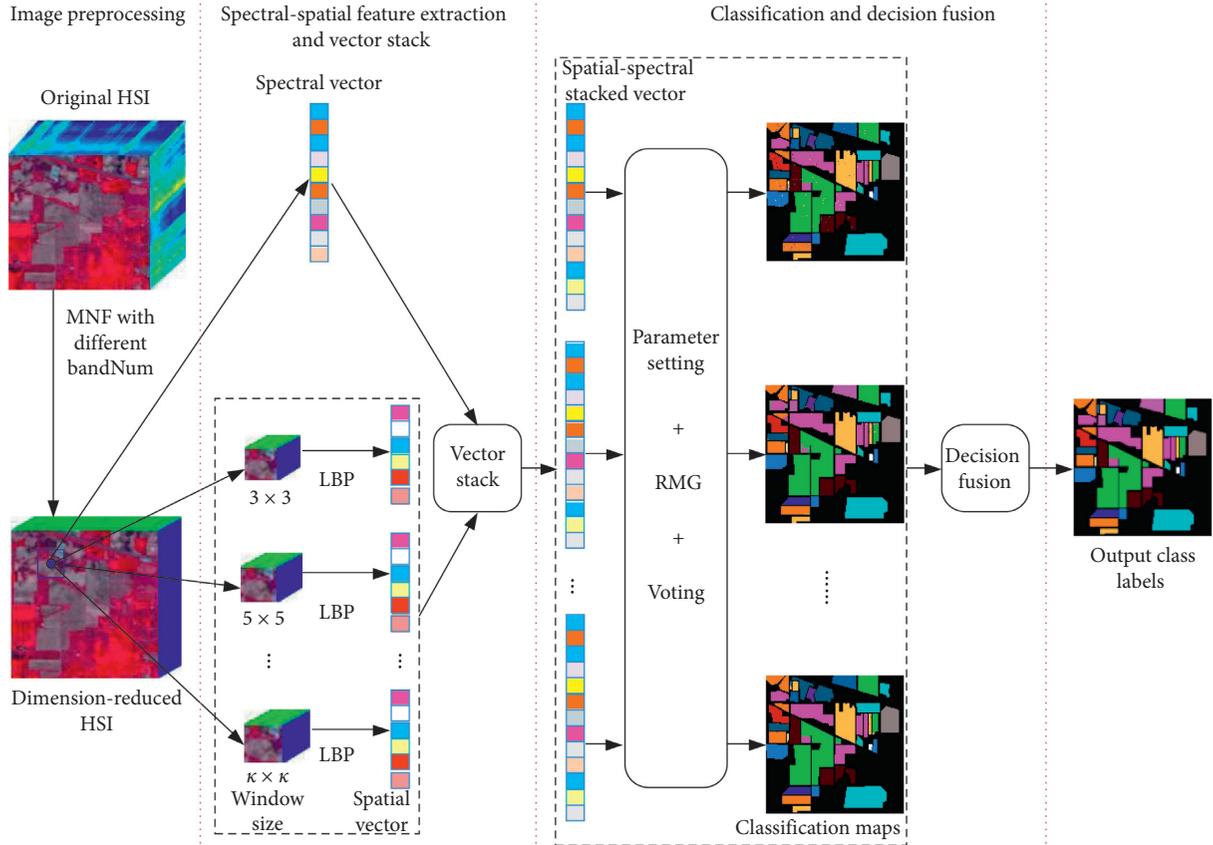


FIGURE 3: Schematic of the SS-MNF-ARMG algorithm.

Input: the original HSI set is $D \in \mathbb{R}^{x \times y \times h}$; the training set is tra_s ; the test set is tes_s ; the number of spectra after dimension reduction is $bandNum = [h'_1, h'_2, \dots, h'_k]$; the patch size for FE is $W = [w_1, w_2, \dots, w_t]$, where $\{w_i\}_{i=1}^t$ is odd and $w_i \geq 3$;

Output: the best classification result of all test samples;

- (1) **for** each $h'_i \in bandNum$ **do**
- (2) Obtain the dimension-reduced HSI $D'_i \in \mathbb{R}^{x \times y \times h'_i}$ by using MNF;
- (3) Extract the spectral vector Spe_vec of D'_i ;
- (4) **for** each $w_j \in W$ **do**
- (5) Calculate the spatial vector Spa_vec_j by using LBP;
- (6) Obtain the spectral-spatial vector by stacking the Spe_vec and the Spa_vec_j ;
- (7) Obtain the best $OA_{i,j}$ (overall accuracy) by voting, and the corresponding confusion matrix $ConfMat_{i,j}$;
- (8) **end for**
- (9) Obtain the best OA' from $OA_{i,j}$ by decision fusion, and the corresponding confusion matrix $ConfMat'$;
- (10) **end for**

ALGORITHM 1: SS-MNF-ARMG.

In this study, we set the number of bands from 3 to 35 to evaluate the impact on the three HSI datasets. We conducted the experiment five times, and the optimal experimental results are shown in Figure 4.

It can be observed that, on the Indian Pines dataset and the Salinas scene dataset, the value of overall classification accuracy (OA) shows a trend of steady growth. In other words, the increase in the number of spectral bands number of training has a positive promotion on the classification performance. In particular, on the Pavia University dataset, there is a sharp growth when the number of bands is above

15. In general, the OA of the proposed method is above 78% for all three HSI datasets. Especially in the Salinas scene dataset, the OA of the method is not less than 92% even with a small number of spectral bands, demonstrating the robustness of our method.

4.3. Comparison and Analysis. The proposed SS-MNF-ARMG method was compared with several state-of-the-art spectral-spatial fusion methods, such as Pixion-based classifier [19], PCA-SPCA-2D-SSA [14], R-VCANet [20], RN-

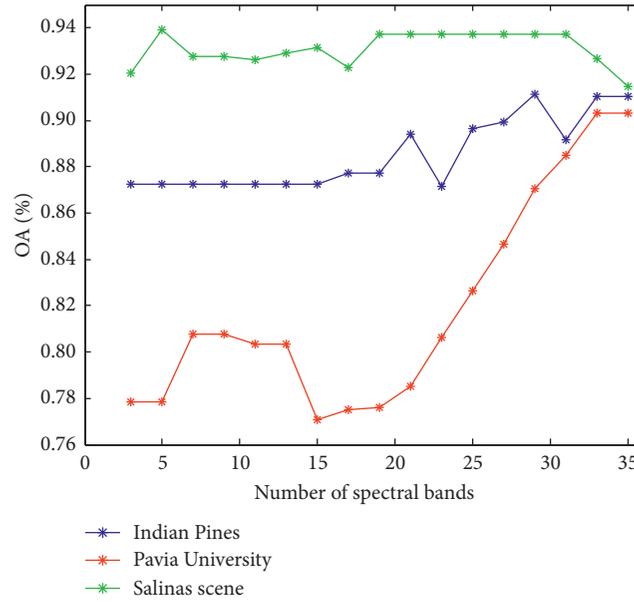


FIGURE 4: Influence of the spectral band numbers.

TABLE 1: A comparison of different methods for hyperspectral image classification.

Group	Subgroup	Method	Advantages	Difficulties
Segmentation-based methods	Object-based classification	Pixon-based classifier	The noise pixels in the classification map were removed and the suitable land cover smoothing map was obtained	The spectral characteristics of abnormal pixels are often similar to the background, resulting in the loss of information when the abnormal pixels are removed
Feature fusion	Features stacking	PCA-SPCA-2D-SSA	Combined with appropriate spatial features, the classification efficiency of the algorithm will be higher	Complex structure
	Joint spectral-spatial FE	R-VCANet	The use of high correlation among SSI	Complex structure
	Meta-learning-based classifiers	RN-FSC	Less sensitive to the number of training samples	Limited generalization ability for large-scale hyperspectral datasets
	Deep learning-based classifiers	iCapsNet RPNNet	Well-initialized shallow layers FE and classification are carried out under a unified framework	Complex structure Insufficient training samples will lead to overfitting
Decision fusion	Decision fusion	MBFSDA	Combination of supplemental information and several advanced classifiers	Selection of suitable feature extractor
		Our method	(i) Self-adaptability of parameters (ii) Leveraging tiny labeled samples	Complex structure

FSC [31], iCapsNet [30], RPNNet [29], and MBFSDA [21]. A comparison of the above algorithms can be seen in Table 1.

To quantitatively compare the classification performance of the methods shown in Table 1, we used the average classification accuracy (AA), overall classification accuracy (OA), and the kappa coefficient (kappa) to assess the classification performance. To demonstrate the superior performance of the SS-MNF-ARMG with a limited number of training samples, we randomly selected 10 samples from each class as training samples. Tables 2–4 show the ground truth classes and their respective training and testing numbers for the three HSI datasets. Tables 5–7 summarize the experimental results for the three HSI datasets, from which the following conclusions can be drawn:

- (1) The results on the Indian Pines dataset show that almost all algorithms are effective. We can observe from Table 5 that the proposed SS-MNF-ARMG achieves a 3.78–28.6% advantage over the other methods in OA. In addition, for the classes that other methods do not recognize accurately, SS-MNF-ARMG can obtain better results, such as objects 1#, 2#, 3#, 5#, 6#, and 7# in Table 5.
- (2) The results on the Pavia University dataset demonstrate that our method has advantages over some state-of-the-art methods. As a decision fusion-based algorithm, the proposed SS-MNF-ARMG surpasses MBFSDA by 3.87% in OA. In addition, for the classes that other methods do not recognize accurately, SS-

TABLE 2: The ground truth classes and their respective training and testing numbers of the Indian Pines datasets (10 classes).

No.	Name of class	Training	Testing
1	Corn-notill	10	1418
2	Corn-mintill	10	820
3	Grass-pasture	10	473
4	Grass-trees	10	720
5	Hay-windrowed	10	468
6	Soybean-notill	10	962
7	Soybean-mintill	10	2445
8	Soybean-clean	10	583
9	Woods	10	1255
10	Bldg-Grass-Tree-Drives	10	376

TABLE 3: The ground truth classes and their respective training and testing numbers of the Pavia University datasets.

No.	Name of class	Training	Testing
1	Asphalt	10	6621
2	Meadows	10	18639
3	Gravel	10	2089
4	Trees	10	3054
5	Painted metal sheets	10	1335
6	Bare Soil	10	5019
7	Bitumen	10	1320
8	Self-Blocking Bricks	10	3672
9	Shadows	10	937

TABLE 4: The ground truth classes and their respective training and testing numbers of the Salinas datasets.

No.	Name of class	Training	Testing
1	Brocoli_green_weeds_1	10	1999
2	Brocoli_green_weeds_2	10	3716
3	Fallow	10	1966
4	Fallow_rough_plow	10	1384
5	Fallow_smooth	10	2668
6	Stubble	10	3949
7	Celery	10	3569
8	Grapes_untrained	10	11261
9	Soil_vineyard_develop	10	6193
10	Corn_senesced_green_weeds	10	3268
11	Lettuce_romaine_4weeks	10	1058
12	Lettuce_romaine_5weeks	10	1917
13	Lettuce_romaine_6weeks	10	906
14	Lettuce_romaine_7weeks	10	1060
15	Vineyard_untrained	10	7258
16	Vineyard_vertical_trellis	10	1797

MNF-ARMG can obtain better results, such as objects 2#, 3#, 6#, and 8# in Table 6.

- (3) The results on the Salinas dataset show that all methods show close results in AA, but SS-MNF-ARMG has a competitive advantage in OA and kappa. From Table 7, we can observe that, compared with R-VCANet, SS-MNF-ARMG obtains a 0.1% improvement in OA, a 3.13% improvement in AA, and a 7.31% improvement in kappa. Furthermore, the proposed SS-MNF-ARMG surpasses iCapsNet

by 11.01% in OA. In addition, for the classes that other methods do not recognize accurately, SS-MNF-ARMG can obtain better results, such as objects 3#, 7#, 8#, 9#, 15#, and 16# in Table 7.

- (4) In general, the decision fusion-based methods outperform the segmentation-based methods and feature fusion-based methods. First and foremost, the LBP features improve the performance of the SS-MNF-ARMG, and the application of MNF reduces the HSI data dimension, controlling the noise in the

TABLE 5: Classification results of different techniques for the Indian Pines dataset (10 classes).

No.	Pixon-based	PCA-SPCA-2D-SSA	R-VCANet	RN-FSC	iCapsNet	RpNet	MBFSDA	SS-MNF-ARMG
1	41.42	61.55	54.94	72.29	64.71	60.32	70.57	75.38
2	33.93	86.49	74.09	78.97	95.56	73.38	81.06	95.6
3	71.43	99.32	83.54	90.59	82.09	93.56	81.09	100.00
4	23.16	96.05	94.78	89.82	99.06	84.2	82.06	93.2
5	87.73	99.37	96.13	99.79	99.59	96.73	99.39	99.79
6	65.81	87.97	75.55	86.61	73.35	86.47	77.48	93.53
7	76.09	86.28	61.43	89.94	88.01	81.48	66.37	92.43
8	22.96	97.01	80.64	98.25	74.59	60.1	73.94	97.69
9	95.44	98.85	88.15	82.61	98.84	58.73	90.8	97.63
10	91.58	93.52	82.63	94.82	96.84	95.79	59.21	94.82
AA	60.96	90.64	79.19	88.37	87.27	79.08	78.2	93.01
OA	62.45	87.27	86.28	86.16	85.83	75.94	76.42	91.05
Kappa	56.7	85.3	72.05	83.97	83.63	72.32	72.95	89.61

Bold values represent the best results among these methods.

TABLE 6: Classification results of different techniques for the Pavia University dataset (9 classes).

No.	Pixon-based	PCA-SPCA-2D-SSA	R-VCANet	RN-FSC	iCapsNet	RpNet	MBFSDA	SS-MNF-ARMG
1	45.20	94.83	79.09	87.28	65.84	83.52	86.68	70.83
2	82.00	81.91	71.76	84.33	84.55	54.19	89.22	96.78
3	90.23	85.47	85.45	90.42	29.91	70.70	87.28	96.24
4	74.05	87.99	94.28	78.09	82.70	91.64	79.34	66.42
5	85.13	97.55	99.98	99.56	100.00	100.00	99.33	96.51
6	18.55	71.55	58.44	63.25	40.22	65.30	89.52	100.00
7	92.48	98.77	94.52	52.09	21.40	92.41	88.12	95.49
8	84.17	79.90	73.70	84.81	57.19	74.93	68.88	95.11
9	97.57	96.30	99.82	95.94	0.00	93.24	82.89	77.51
AA	74.38	88.25	84.12	81.94	66.63	80.66	85.70	88.32
OA	69.63	84.50	83.23	81.75	53.54	68.81	86.45	90.32
Kappa	57.72	79.86	67.97	75.84	56.07	61.78	82.31	87.26

Bold values represent the best results among these methods.

TABLE 7: Classification results of different techniques for the Salinas dataset (16 classes).

No.	Pixon-based	PCA-SPCA-2D-SSA	R-VCANet	RN-FSC	iCapsNet	RpNet	MBFSDA	SS-MNF-ARMG
1	99.90	96.30	96.95	99.26	93.41	98.01	96.95	97.11
2	99.38	88.94	99.06	93.21	67.16	99.19	97.65	98.98
3	96.20	92.09	97.31	97.87	86.49	96.91	99.60	100.00
4	91.61	90.23	99.46	85.50	90.36	99.21	100.00	86.21
5	86.89	88.94	97.35	84.81	90.39	98.36	96.30	88.67
6	88.79	96.08	99.87	99.35	99.02	99.60	99.14	89.47
7	97.32	91.36	95.26	96.65	95.07	96.42	93.55	97.65
8	87.20	96.44	64.01	66.24	82.12	76.01	78.33	96.66
9	99.84	85.97	97.79	97.34	98.63	96.58	97.81	100.00
10	46.34	86.38	81.38	82.66	93.13	84.66	94.23	87.22
11	52.90	94.17	93.17	73.96	60.09	96.25	97.38	88.30
12	30.10	80.49	96.48	96.84	94.78	97.87	99.27	97.18
13	89.74	96.17	98.70	85.39	49.51	93.56	98.69	86.42
14	84.21	99.30	90.35	90.39	76.76	97.76	75.70	95.79
15	94.79	86.68	56.92	65.85	49.76	66.26	80.38	95.36
16	92.09	98.17	90.79	86.89	96.89	97.57	99.89	100.00
AA	83.58	91.43	90.93	87.64	77.26	93.39	94.05	94.06
OA	87.15	91.73	93.56	93.12	82.72	88.16	90.96	93.73
Kappa	85.63	90.90	85.72	85.44	74.72	86.83	89.95	93.03

Bold values represent the best results among these methods.

HSI data. In addition, the randomness in the SS-MNF-ARMG can be regarded as a regularization technique, which alleviates the phenomenon of overfitting. Finally, benefitting from the semi-supervised deep learning framework, the classification accuracy is relatively satisfactory even with only a few labeled samples.

Our experiments were performed using MATLAB 2018b on a computer with an IntelCore(TM) i5-4300M 2.60 GHz CPU, 16 GB memory, and a 64-bit Windows 7 system. For the three real HSI datasets, the duration to execute our algorithm was several minutes to several hours; notably, the duration to execute the other methods that we reviewed in this study was shorter than that of our method.

5. Conclusions and Further Research

In this study, we developed a novel decision fusion method for HSI data classification. In the proposed SS-MNF-ARMG, MNF and multiscale LBP were integrated to extract local SSFs. On the one hand, MNF helped reduce the dimension of the HSI, remove the noise in the HSI, and extract the spectral features from the HSI data. On the other hand, multiscale LBP was applied to the MNF dimension-reduced images to derive spatial features at different scales. These spatial features were further fused with the MNF spectral feature to form the SSFs. Compared with some state-of-the-art spectral-spatial classification methods, our experimental results have demonstrated that SS-MNF-ARMG can achieve higher classification accuracy with limited training samples. This method is effective for distinguishing different land cover types. In addition, a set of optimal parameters for different hyperspectral data can be obtained automatically.

Although the SS-MNF-ARMG algorithm has provided promising results, the classification accuracies of various datasets remain different. Further research could attempt to further improve the generalization ability of our method. Due to human ecological destruction or natural disasters (e.g., earthquakes), the vegetation in some areas has significantly changed. Monitoring vegetation restoration in these areas is of substantial importance. Therefore, we plan to research the application of the HSI classification in vegetation restoration monitoring.

Data Availability

The data used to support the findings of this study are available at GIC (http://www.ehu.us/ccwintco/index.php?title=Hyperspectral_Remote_Sensing_Scenes).

Conflicts of Interest

The authors declare no conflicts of interest.

Acknowledgments

This work was supported by the National Natural Science Foundation of China under Grant nos. 61703060, 61802036, 61701048, and 61873305, the Sichuan Science and

Technology Program under Grant no. 21YYJC0469, the Project funded by China Postdoctoral Science Foundation under Grant no. 2020M683274, the Fundamental Research Funds of the Central Universities, Southwest Minzu University (2019NQN07), the Opening Fund of Geomathematics Key Laboratory of Sichuan Province (scsxdz2020zd01 and csxdz201710), and the Key Laboratory of Pattern Recognition and Intelligent Information Processing of Sichuan (MSSB-2018-0 and MSSB-2020-9).

References

- [1] T. Matsuki, N. Yokoya, and A. Iwasaki, "Hyperspectral tree species classification of Japanese complex forest with the aid of LiDAR data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 5, pp. 2177–2187, 2015.
- [2] P. Hardin and A. Hardin, "Hyperspectral remote sensing of urban areas," *Geography Compass*, vol. 7, no. 1, pp. 7–21, 2013.
- [3] Y. Xie, F. Miao, K. Zhou et al., "HsgNet: a road extraction network based on global perception of high-order spatial information," *ISPRS International Journal of Geo-Information*, vol. 8, no. 12, p. 571, 2019.
- [4] Z. Pan, J. Liu, L. Ma et al., "Research on hyperspectral identification of altered minerals in Yemaquan West Gold Field, Xinjiang," *Sustainability*, vol. 11, no. 2, p. 428, 2019.
- [5] X. Kang, X. Xiang, S. Li et al., "PCA-based edge-preserving features for hyperspectral image classification," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 55, no. 12, pp. 7140–7151, 2017.
- [6] M. B. Stuart, A. J. S. McGonigle, and J. R. Willmott, "Hyperspectral imaging in environmental monitoring: a review of recent developments and technological advances in compact field deployable systems," *Sensors*, vol. 19, no. 14, p. 3071, 2019.
- [7] X. Briottet, Y. Boucher, A. Dimmeler et al., "Military applications of hyperspectral imagery, Targets and backgrounds XII: characterization and representation," *International Society for Optics and Photonics*, vol. 6239, p. 62390B, 2006.
- [8] B. Hörig, F. Kühn, F. Oschütz et al., "HyMap hyperspectral remote sensing to detect hydrocarbons," *International Journal of Remote Sensing*, vol. 22, no. 8, pp. 1413–1422, 2001.
- [9] C. Butz, M. Grosjean, D. Fischer et al., "Hyperspectral imaging spectroscopy: a promising method for the biogeochemical analysis of lake sediments," *Journal of Applied Remote Sensing*, vol. 9, no. 1, Article ID 096031, 2015.
- [10] K. Jacq, Y. Perrette, B. Fanget et al., "Hyperspectral imaging for lake sediment cores analysis," in *Proceedings of the IPA-IAL Stockholm 2018 Conference*, pp. 735–743, Stockholm, Sweden, June 2018.
- [11] E. C. Koeniguer, F. Janez, and J. M. Nicolas, "Change detection in SAR time-series based on the coefficient of variation," 2019, <https://arxiv.org/abs/1904.11335>.
- [12] M. Zhang, J. Zhou, K. A. Sudduth et al., "Estimation of maize yield and effects of variable-rate nitrogen application using UAV-based RGB imagery," *Biosystems Engineering*, vol. 189, pp. 24–35, 2020.
- [13] M. P. Uddin, M. A. Mamun, M. I. Afjal et al., "Information-theoretic feature selection with segmentation-based folded principal component analysis (PCA) for hyperspectral image classification," *International Journal of Remote Sensing*, vol. 42, no. 1, pp. 286–321, 2021.

- [14] H. Fu, G. Sun, J. Ren et al., "Fusion of PCA and segmented-PCA domain multiscale 2-D-SSA for effective spectral-spatial feature extraction and data classification in hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, 2020.
- [15] C. Ruichun, W. Lu, and W. Maozhi, "Application of FastICA in mineral information extraction from hyperspectral remote sensing data," *Remote Sensing for Land & Resources*, vol. 25, no. 4, pp. 129–132, 2013.
- [16] A. A. Green, M. Berman, P. Switzer et al., "A transformation for ordering multispectral data in terms of image quality with implications for noise removal," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 26, no. 1, pp. 65–74, 1988.
- [17] W. Li, S. Prasad, J. E. Fowler et al., "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 50, no. 4, pp. 1185–1198, 2011.
- [18] W. Zhao and S. Du, "Spectral-spatial feature extraction for hyperspectral image classification: a dimension reduction and deep learning approach," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 8, pp. 4544–4554, 2016.
- [19] A. Zehtabian and H. Ghassemian, "Automatic object-based hyperspectral image classification using complex diffusions and a new distance metric," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 54, no. 7, pp. 4106–4114, 2016.
- [20] B. Pan, Z. Shi, X. Xu, and R-VCANet, "A new deep-learning-based hyperspectral image classification method," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 10, no. 5, pp. 1975–1986, 2017.
- [21] M. Imani and H. Ghassemian, "Discriminant analysis in morphological feature space for high-dimensional image spatial-spectral classification," *Journal of Applied Remote Sensing*, vol. 12, no. 1, Article ID 016024, 2018.
- [22] A. Sellami, A. B. Abbes, V. Barra et al., "Fused 3-D spectral-spatial deep neural networks and spectral clustering for hyperspectral image classification," *Pattern Recognition Letters*, vol. 138, pp. 594–600, 2020.
- [23] M. Imani and H. Ghassemian, "An overview on spectral and spatial information fusion for hyperspectral image classification: current trends and challenges," *Information Fusion*, vol. 59, pp. 59–83, 2020.
- [24] X. Huang and L. Zhang, "An SVM ensemble approach combining spectral, structural, and semantic features for the classification of high-resolution remotely sensed imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 51, no. 1, pp. 257–272, 2012.
- [25] K. Y. Peerbhaya, O. Mutanga, and R. Ismail, "Random forests unsupervised classification: the detection and mapping of solanum mauritianum infestations in plantation forestry using hyperspectral data," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 8, no. 6, pp. 3107–3122, 2015.
- [26] K. Shi, J. Wang, Y. Tang et al., "Reliable asynchronous sampled-data filtering of T-S fuzzy uncertain delayed neural networks with stochastic switched topologies," *Fuzzy Sets and Systems*, vol. 381, pp. 1–25, 2020.
- [27] H. Han, G. Wang, M. Wang et al., "Hyperspectral Unmixing via nonconvex sparse and low-rank constraint," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 13, pp. 5704–5718, 2020.
- [28] W. Dong, F. Fu, G. Shi et al., "Hyperspectral image super-resolution via non-negative structured sparse representation," *IEEE Transactions on Image Processing*, vol. 25, no. 5, pp. 2337–2352, 2016.
- [29] Y. Xu, B. Du, F. Zhang et al., "Hyperspectral image classification via a random patches network," *ISPRS Journal of Photogrammetry and Remote Sensing*, vol. 142, pp. 344–357, 2018.
- [30] J. Yin, S. Li, H. Zhu et al., "Hyperspectral image classification using CapsNet with well-initialized shallow layers," *IEEE Geoscience and Remote Sensing Letters*, vol. 16, no. 7, pp. 1095–1099, 2019.
- [31] K. Gao, B. Liu, X. Yu et al., "Deep relation network for hyperspectral image few-shot classification," *Remote Sensing*, vol. 12, no. 6, p. 923, 2020.
- [32] P. Ghamisi, R. Souza, J. A. Benediktsson et al., "Hyperspectral data classification using extended extinction profiles," *IEEE Geoscience and Remote Sensing Letters*, vol. 13, no. 11, pp. 1641–1645, 2016.
- [33] L. Fang, N. He, S. Li et al., "A new spatial-spectral feature extraction method for hyperspectral images using local covariance matrix representation," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 6, pp. 3534–3546, 2018.
- [34] C. Ge, Q. Du, W. Li et al., "Hyperspectral and LiDAR data classification using kernel collaborative representation based residual fusion," *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, vol. 12, no. 6, pp. 1963–1973, 2019.
- [35] W. N. Khotimah, M. Bennamoun, F. Boussaid et al., "A high-performance spectral-spatial residual network for hyperspectral image classification with small training data," *Remote Sensing*, vol. 12, no. 19, p. 3137, 2020.
- [36] G. Hughes, "On the mean accuracy of statistical pattern recognizers," *IEEE Transactions on Information Theory*, vol. 14, no. 1, pp. 55–63, 1968.
- [37] R. Girshick, J. Donahue, T. Darrell et al., "Rich feature hierarchies for accurate object detection and semantic segmentation," in *Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition*, pp. 580–587, Columbus, OH, USA, June 2014.
- [38] K. Simonyan and A. Zisserman, "Very deep convolutional networks for large-scale image recognition," 2014, <https://arxiv.org/abs/1409.1556>.
- [39] C. Qi, Z. Zhou, Y. Sun et al., "Feature selection and multiple kernel boosting framework based on PSO with mutation mechanism for hyperspectral classification," *Neurocomputing*, vol. 220, pp. 181–190, 2017.
- [40] X. Ceamanos, B. Waske, J. A. Benediktsson et al., "A classifier ensemble based on fusion of support vector machines for classifying hyperspectral data," *International Journal of Image and Data Fusion*, vol. 1, no. 4, pp. 293–307, 2010.
- [41] Y. Gu and H. Liu, "Sample-screening MKL method via boosting strategy for hyperspectral image classification," *Neurocomputing*, vol. 173, pp. 1630–1639, 2016.
- [42] J. Jiang, J. Ma, C. Chen et al., "SuperPCA: a superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery," *IEEE Transactions on Geoscience and Remote Sensing*, vol. 56, no. 8, pp. 4581–4593, 2018.
- [43] Q. Zhang, J. Sun, G. Zhong et al., "Random multi-graphs: a semi-supervised learning framework for classification of high dimensional data," *Image and Vision Computing*, vol. 60, pp. 30–37, 2017.
- [44] F. Gao, Q. Wang, J. Dong et al., "Spectral and spatial classification of hyperspectral images based on random multi-graphs," *Remote Sensing*, vol. 10, no. 8, p. 1271, 2018.
- [45] H. Chen, F. Miao, and X. Shen, "Hyperspectral remote sensing image classification with CNN based on quantum genetic-

- optimized sparse representation,” *IEEE Access*, vol. 8, pp. 99900–99909, 2020.
- [46] F. Bianconi, R. Bello-Cerezo, and P. Napoletano, “Improved opponent color local binary patterns: an effective local image descriptor for color texture classification,” *Journal of Electronic Imaging*, vol. 27, no. 1, Article ID 011002, 2017.
- [47] B. Tu, W. Kuang, G. Zhao et al., “Hyperspectral image classification by combining local binary pattern and joint sparse representation,” *International Journal of Remote Sensing*, vol. 40, no. 24, pp. 9484–9500, 2019.
- [48] F. R. K. Chung, “Lectures on spectral graph theory,” *CBMS Lectures, Fresno*, vol. 6, no. 92, pp. 17–21, 1996.
- [49] W. Liu, J. Wang, S. Kumar et al., “Hashing with graphs,” in *Proceedings of the 28th International Conference on Machine Learning, ICML 2011*, pp. 1–8, Bellevue, WA, USA, June-July 2011.
- [50] S. Kim and S. Choi, “Multi-view anchor graph hashing,” in *Proceedings of the 2013 IEEE International Conference on Acoustics, Speech and Signal Processing*, pp. 3123–3127, IEEE, Vancouver, Canada, May 2013.