IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received 3 August 2023, accepted 3 September 2023, date of publication 12 September 2023,
date of current version 15 September 2023.

Digital Object Identifier 10.1109/ACCESS.2023.3314694

RESEARCH ARTICLE

A Novel Semi-Supervised Dynamic Classifier Selection Method for HSI Classification Based on SP Segmentation

XIANG GE[1], XUEXIANG YU[2], AND XU YANG[2]
[1]Huainan City Construction and Development Planning and Design Institute, Huainan 232001, China
[2]School of Spatial Information and Surveying Engineering, Anhui University of Science and Technology, Huainan 232001, China

Corresponding author: Xiang Ge (543658314@qq.com)

This work was supported in part by the Key Research and Development Plan of Anhui Province, in 2021: Construction of public service platform for coordinated monitoring, analysis and decision-making of mining subsidence disasters under Grant 202104a07020014; and in part by the 2021 Anhui Province Science and Technology Major Project: "Beidou +" cloud platform collaborative monitoring and rapid warning of typical geological disasters research and development and demonstration application under Grant 202103a05020026.

ABSTRACT This paper proposes a novel hyperspectral image classification method that combines dynamic semi-supervised multiple-kernel collaborative representation ensemble selection with superpixel (SP) consistency constraints. The method is based on the consistency principle of labels within SP blocks, where the hyperspectral image is divided into different SP blocks, and each block is treated as an independent classification task. It applies a dynamic ensemble selection strategy to select high-confidence samples from the unlabeled data and assigns pseudo-labels to expand the available training sample set. Additionally, it employs a multiple-kernel collaborative representation classifier as the base classifier to better capture sample similarities and correlations, thereby improving the classification performance. Experimental results demonstrate that the proposed method achieves superior classification accuracy on various datasets such as Indian Pines, Purdue, and KSC, outperforming the traditional Meta-DES method significantly.

INDEX TERMS Hyperspectral image classification, superpixel segmentation, semi-supervised, dynamic classifier selection, ensemble learning.

I. INTRODUCTION

Hyperspectral remote sensing is considered an advanced remote sensing technology, holding significant importance [1]. Through the collection of continuous spectral information of surface objects, detailed information regarding the intrinsic attributes and conditions of these objects can be obtained [2]. Its application is wide-ranging and includes geological exploration, agricultural monitoring, environmental protection [3], and wetland [4] research. Hyperspectral image (HSI) classification is a technique used to finely distinguish surface features based on spectral information, which is deemed crucial for various applications such as geological exploration, environmental monitoring, urban planning, and agricultural management [5], [6], [7]. Important implications are brought about by this technique as it allows for a better understanding and monitoring of various changes on the ground, enabling effective resource optimization and environmental protection. However, in practical applications, an important problem is often faced by HSI classification, known as the "small sample problem" [7], [8], [9]. In large-scale remote sensing images, certain types of land cover (e.g., specific minerals or rare vegetation) may only be represented by a few pixels, leading to a scarcity of training samples. As a consequence, traditional supervised learning methods may not perform well on this type of problem, as they typically require a large number of samples for effective training. Thus, the key challenge that HSI classification needs to overcome is how to be modeled with limited labeled samples.

In HSI classification tasks, the limitations of traditional supervised learning methods in adequately exploiting classification performance arise due to the high dimensionality of data and the limited number of samples available.

The associate editor coordinating the review of this manuscript and approving it for publication was Wenming Cao.

As a resolution, semi-supervised learning provides an effective approach, amalgamating labeled and unlabeled data for learning purposes. Within the realm of HSI classification, semi-supervised learning can be broadly categorized into four methodological approaches: graph-based methods [10], [11], [12], generative adversarial networks (GANs) [13], [14], [15] and variational autoencoders (VAEs) [16], transfer learning and domain adaptation, and active learning. Graph-based semi-supervised learning methods employ data similarity for the construction of graph representations [17], [18], where nodes represent data samples, and edges indicate relationships between data instances. In the context of HSI classification, these methods leverage the similarity between labeled and unlabeled data to facilitate label propagation across the graph. Commonly used techniques for graph construction encompass k-nearest neighbors and graph-cut-based approaches. Subsequently, graph propagation [19] algorithms are employed to iteratively propagate labels on the graph, thereby endowing unlabeled data with labels. The utilization of graph-based methods not only enhances the training sample size for the classification model but also fosters the development of more accurate models of classification boundaries through the exploitation of unlabeled data, leading to improved classification performance. GANs and VAEs [20] constitute two potent generative models, capable of producing high-quality samples, which are harnessed in semi-supervised learning for the augmentation of labeled data. In the context of HSI classification, GANs generate realistic unlabeled data akin to genuine samples, thereby enriching the dataset. VAEs [21], on the other hand, learn latent representations of samples and perform interpolation in the latent space, yielding new samples. These generated samples can be amalgamated with labeled data to train the classification model more effectively, thereby resulting in enhanced performance. The realm of HSI classification may be confronted with domain differences among various datasets, wherein the data distributions vary, leading to diminished generalization performance on novel data. Addressing the issue of limited sample size, transfer learning and domain adaptation methods encompass the acquisition of knowledge from the source domain and its application to the target domain. Transfer learning ameliorates the performance of the target task by facilitating the sharing of features or knowledge from the source domain. In contrast, domain adaptation endeavors to mitigate the domain disparities between the source and target domains, thereby fostering the model's improved adaptation to new unlabeled data. Active learning, as a semi-supervised learning approach, empowers the model to actively select the most informative samples for labeling, thus reducing the dependency on labeled samples. In the context of HSI classification, active learning facilitates the selection of the most representative and discriminative samples to optimize model training. Active learning algorithms often accomplish the active labeling of samples based on measures of uncertainty or boundary samples. In summary, within the domain of HSI classification, semi-supervised learning encompasses a range of methodological approaches, including graph-based techniques exploiting data similarity for label propagation; GANs [22], [23] and VAEs for data augmentation; transfer learning and domain adaptation to address domain disparities; and active learning, which reduces the reliance on labeled samples through the active selection of informative data points.

In the aforementioned semi-supervised learning methods, data augmentation is a commonly employed technique, which enables the enlargement of training datasets through the generation of novel samples, thereby enhancing the model's performance. Nonetheless, several scientific issues demand careful attention during the data augmentation process. Firstly, in graph-based semi-supervised learning approaches, the critical concern lies in the selection of appropriate similarity metrics and the determination of the number of neighboring samples for constructing the graph. Such choices may lead to distinct graph structures and varying outcomes in label propagation. Secondly, when employing GANs and VAEs for sample generation, striking a balance between sample diversity and realism is crucial. Overemphasizing realism may result in excessively conservative samples, lacking coverage of the entire data distribution. Thirdly, the effectiveness of transfer learning and domain adaptation methods hinges upon resolving domain discrepancy issues. However, during the process of data augmentation, the introduction of further domain disparities may impinge upon the model's generalization capabilities. Finally, the success of active learning relies on the judicious selection of representative and discriminative samples for annotation. Yet, the sample selection strategy may be influenced by factors such as data distribution imbalance and label noise, leading to non-representative choices.

To address the issue of low confidence in augmented samples in the above semi-supervised methods, this paper proposes a dynamic semi-supervised multi-kernel collaborative representation ensemble selection with superpixel (SP) [24], [25], [26] consistency constraints for HSI classification. The method is based on the principle of label homogeneity within SP blocks, where the HSI is divided into different SP blocks. Each SP block is treated as an independent classification task, and constraints are applied using these SP blocks. Furthermore, a multi-kernel collaborative representation classifier (CRC) is used as the base classifier to better capture the similarities and correlations between samples. The classifier selection idea is introduced to select high-confidence samples from unlabeled data and assign pseudo-labels to them for augmenting the available training sample set. The main innovations of this paper are as follows:

A. INTEGRATION OF DYNAMIC ENSEMBLE SELECTION WITH SP SEGMENTATION

The paper combines the concept of dynamic ensemble selection with SP segmentation. Through SP segmentation, the

HSI is divided into different SP blocks, each treated as a separate classification task. Then, the dynamic ensemble selection strategy is applied to select high-confidence samples from the unlabeled data and assign pseudo-labels to augment the available training sample set.

B. DYNAMIC MULTI-KERNEL CRC CLASSIFIER SELECTION

To better capture the similarity and correlation between samples, the paper adopts the multi-kernel CRC as the base classifier. This classifier effectively utilizes multiple feature representations of samples and collaboratively expresses them during the classification process, thereby improving classification performance.

C. SP BLOCK LABEL CONSISTENCY

The paper considers the consistency of sample labels within SP blocks and further enhances classifier accuracy through dynamic multi-kernel CRC classifier selection. This aids in better utilizing the information within SP blocks and strengthening classifier robustness.

D. JOINT MULTI-SCALE SP SEGMENTATION FOR SAMPLE AUGMENTATION

By utilizing SP segmentation at different scales for sample augmentation, and the diversity and generalization ability of sample augmentation is enhanced. Finally, the ensemble learning method is used to combine the augmentation results at different scales to obtain the final classification result.

The remainder of this paper is organized as follows. Section II introduces related works. Section III proposes the dynamic semi-supervised multi-kernel collaborative representation ensemble selection algorithms. In Section IV, experiments and analysis with three real hyperspectral data are presented. Finally, Section V draws the conclusion and perspectives.

II. RELATED WORKS

A. COLLABORATIVE REPRESENTATION CLASSIFIER (CRC)

Collaborative Representation-based Classification (CRC) [27], [28] is a powerful technique used in face recognition and hyperspectral image (HSI) classification. The method aims to find the optimal linear representation of testing samples based on a training set and a testing set [29]. By minimizing the reconstruction error and a regularization term, CRC calculates collaborative representation coefficients that act as weights to reconstruct the testing samples using the training data. This approach effectively leverages the collaborative information among the training samples to enhance the classification performance, making it particularly suitable for scenarios where the data may be linearly inseparable [30], [31]. CRC has shown outstanding performance in various image classification tasks due to its ability to capture the underlying relationships and achieve accurate and robust classification results.

CRC is described in detail as follows. Give a training set X and a testing set Y The collaborative representation coefficients $\boldsymbol{\alpha}$ is obtained as the following formula,

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}^*} ||\boldsymbol{Y} - \boldsymbol{X}\boldsymbol{\alpha}^*||_2^2 + \lambda||\boldsymbol{\alpha}^*||_2^2 \tag{1}$$

where $||Y - X\boldsymbol{\alpha}^*||_2^2$ is the reconstruction error, λ is the regularization parameter, and $\lambda||\boldsymbol{\alpha}^*||_2^2$ controls the complexity of $\boldsymbol{\alpha}^*$. The analytical solution can be easily obtained,

$$\boldsymbol{\alpha} = \left(\boldsymbol{X}^T\boldsymbol{X} + \lambda\boldsymbol{I}\right)^{-1}\boldsymbol{X}^T\boldsymbol{Y} \tag{2}$$

where $\boldsymbol{I}$ denotes the identity matrix.

1) SP SEGMENTATION (SP)

In the field of hyperspectral image classification, SP segmentation [24], [25], [26] plays a crucial role, particularly in preserving spatial information. Hyperspectral images contain rich spectral information, but the spatial relationships between pixels are also essential for accurate classification. Traditional pixel-level classification may overlook the spatial correlations among pixels, leading to less precise classification results. SP segmentation effectively preserves the spatial locality by grouping neighboring pixels into compact and correlated SP regions. This approach maintains the spectral information of the hyperspectral image while utilizing the spatial correlations to improve classification accuracy. By clustering similar pixels together, the SP regions enable classifiers to more easily discern the spatial distribution of different categories during classification. Through SP segmentation, complex textures and spatial variations in hyperspectral images are better captured and represented, enhancing the classifier's ability to recognize various land cover categories in the image. Consequently, in the domain of hyperspectral image classification, SP segmentation proves to be a beneficial and effective preprocessing technique, significantly improving classification accuracy and providing a more reliable foundation for tasks such as remote sensing image analysis and land cover identification.

B. DYNAMIC ENSEMBLE SELECTION

Dynamic Ensemble Selection (DES) [32], [33] is an advanced approach in ensemble learning, aiming to discover an optimal combination of base classifiers to outperform using the entire ensemble. A crucial aspect of DES is defining a reliable classifier confidence index, which falls into two main categories: supervised and unsupervised methods. In supervised methods, DES leverages techniques such as K-Nearest Neighbors (KNN) [34] and clustering to establish a region of competence (RoC) for each test instance. The RoC comprises training instances that are most similar or closest to the test instance. Base classifiers exhibiting strong performance within each RoC are selected to make predictions, thereby enhancing the ensemble model's accuracy and generalization ability. The confidence of each base classifier is determined based on its classification performance within the RoC, and

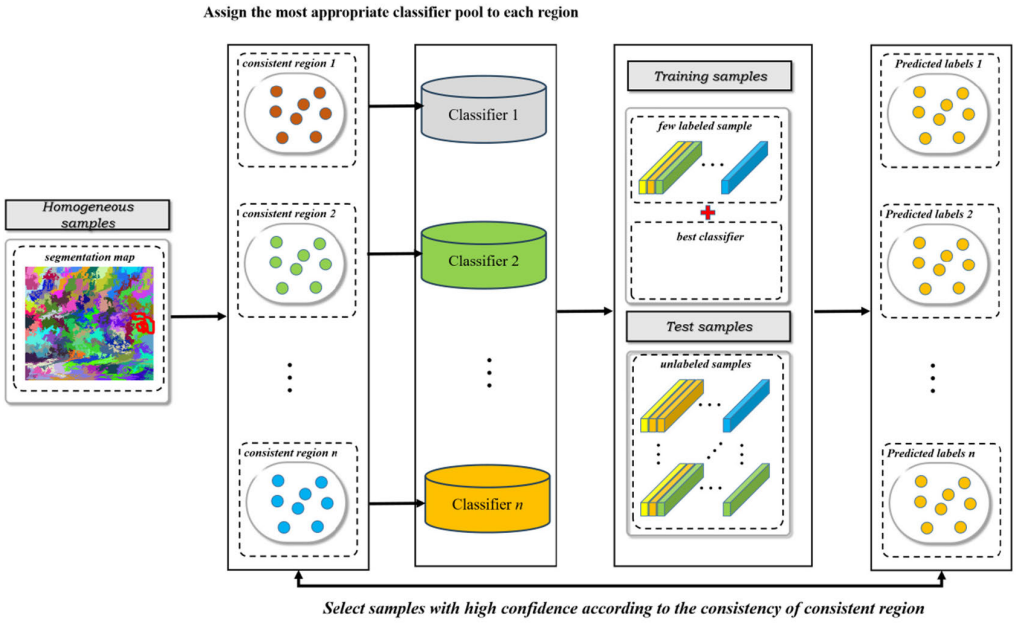

FIGURE 1. Flow chart of proposed methods.

then an optimal subset of base classifiers is chosen according to their confidence. Conversely, unsupervised methods approach DES as an optimization problem. The selective ensemble is formulated as an optimization challenge, and the classifier confidence index is computed by solving this problem. These techniques utilize optimization-based approaches to dynamically select the most competent base classifiers for each test instance.

III. PROPOSED METHOD

A. SP SEGMENTATION FOR HSI

SP segmentation [24], [25] is an important technique in the field of image processing, which divides an image into contiguous, compact, and similar regions. In this paper, the Simple Linear Iterative Clustering (SLIC) method is employed to perform SP segmentation on the HS. Additionally, the homogeneous property of SP regions is utilized for sample augmentation. The specific process is as follows.

1) DATA PREPARATION

In SP segmentation, the HSI is first transformed into a 2D spatial representation, integrating spectral information with spatial information into a two-dimensional matrix. This step allows each pixel in the image to be represented as a vector with spatial coordinates.

2) INITIALIZATION OF SP CENTERS

To initiate the SP segmentation process, K initial SP centers are uniformly selected on the image, where K represents the number of SPs. The choice of K should consider the image size and the desired SP density. These initial centers serve as seeds for pixel assignment.

3) SP ASSIGNMENT

For each pixel, its closest SP center is determined by computing the distance between the pixel and all the SP centers. Typically, the Euclidean distance or other distance metrics are used to measure the similarity between pixels. The pixel is then assigned to the nearest SP center, forming the initial SP regions. The formula for calculating the distance (Euclidean distance) from a pixel to a SP center is as follows.

$$D = [(L-L_c)^2 + (a-a_c)^2 + (b-b_c)^2 + \frac{m^2}{S^2}(x-x_c)^2 + \frac{m^2}{S^2}(y-y_c)^2]^{\frac{1}{2}} \quad (3)$$

where, L, a, b represent the high spectral resolution values of the pixel, (x, y) denotes the pixel's position, L_c, a_c, b_c are the high spectral resolution values of the SP center, (x_c, y_c) represents the position of the SP center, S is the preset SP size (used to control compactness), and m is the weight parameter between spatial distance and color distance.

4) UPDATE OF SP CENTERS

For each SP, the center position and spectral values are re-computed by calculating the average position and spectral values of all pixels within that SP region. This step aims to adjust the SP center to the mean feature values of its constituent pixels, providing a better representation of local image features. The formula for updating the SP center is as follows.

$$X_c = \frac{1}{N}\sum_{i=1}^{N} x_i, Y_c = \frac{1}{N}\sum_{i=1}^{N} y_i, L_c = \frac{1}{N}\sum_{i=1}^{N} L_i, a_c = \frac{1}{N}\sum_{i=1}^{N} a_i, b_c = \frac{1}{N}\sum_{i=1}^{N} b_i, \quad (4)$$

where, (X_c, Y_c) represents the updated position of the SP center, N is the number of pixels assigned to the SP center, and L_i, a_i, b_i are the high spectral resolution values of the pixels assigned to the SP center.

5) ITERATIVE OPTIMIZATION

The process of pixel assignment and updating of SP centers is repeated iteratively until the positions of the SP centers no longer exhibit significant changes or the predefined maximum iteration count is reached. The objective of iteration is to progressively optimize the shapes and positions of the SPs, achieving more accurate SP segmentation results.

6) POST-PROCESSING

Upon completing the SP segmentation, post-processing steps can be performed to further optimize the segmentation results. For example, similar SP regions can be merged to reduce over-segmentation, or smoothing can be applied to the boundaries to eliminate discontinuities. Post-processing improves the quality and continuity of the SP segmentation results, making them more suitable for practical applications.

B. DYNAMIC MULTI-KERNEL CRC CLASSIFIER SELECTION (DKCRCS)

1) MULTI-KERNEL CRC CONSTRUCTED

Based on CRC, a kernel function is introduced to project the data into a high-dimensional feature space, and the so-called kernel CR (KCRC) is formulated as

$$\boldsymbol{\alpha} = \arg\min_{\boldsymbol{\alpha}^*} ||\boldsymbol{\Phi}(\boldsymbol{X}) - \boldsymbol{\Phi}(\boldsymbol{Y})\boldsymbol{\alpha}^*||_2^2 + \lambda||\boldsymbol{\alpha}^*||_2^2 \quad (5)$$

where $\boldsymbol{\Phi}()$ projects $\boldsymbol{X}$ and $\boldsymbol{Y}$ to a high-dimensional space, respectively.

The coefcient α can be calculated as

$$\boldsymbol{\alpha} = (\boldsymbol{\Phi}(\boldsymbol{X})^T \boldsymbol{\Phi}(\boldsymbol{X}) + \lambda \boldsymbol{I})^{-1} \boldsymbol{\Phi}(\boldsymbol{X})^T \boldsymbol{\Phi}(\boldsymbol{Y}) \quad (6)$$

This paper introduces three different kernel methods to improve the CRC algorithm and employs them as alternative classifiers for multi-classifier selection. The first kernel method is the most common Gaussian kernel, which can be represented as follows,

$$\boldsymbol{K}_{\text{Gaussian}}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{exp}(-\frac{\|\boldsymbol{x_i} - \boldsymbol{x_j}\|^2}{2\sigma^2}) \quad (7)$$

The second kernel method is the Laplacian kernel, which can be represented as follows,

$$\boldsymbol{K}_{\boldsymbol{Laplacian}}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{exp}(-\frac{\|\boldsymbol{x_i} - \boldsymbol{x_j}\|}{\sigma}) \quad (8)$$

The third kernel method is the linear kernel, which can be represented as follows,

$$\boldsymbol{K}_{\boldsymbol{Linear}}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \boldsymbol{x_i^T x_j} \quad (9)$$

2) ROC DEFINED BASED ON K-NN

K-Nearest Neighbors (K-NN) methods are commonly used for ROC definition in DES. The primary concept behind these methods is to leverage the correlation between validation and test data. The dataset is partitioned into regions with similar characteristics using clustering techniques. Subsequently, the classification performance of each classifier is assessed within these homogeneous regions. Specifically, the K-NN method identifies k samples around each test point by computing their distances from the samples in the dataset. These selected samples form a validation set, which provides essential prior information for training each classifier. Firstly, calculate the distance between each test sample and all the training samples. The distance metric commonly used is Euclidean distance, given by.

$$\boldsymbol{d}(\boldsymbol{x_i}, \boldsymbol{x_j}) = \sqrt{\sum_{k=1}^{n} (\boldsymbol{x_{ik}} - \boldsymbol{x_{jk}})^2} \quad (10)$$

where $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$ are two feature vectors from the dataset, n is the number of features, $\boldsymbol{x_{ik}}$ and $\boldsymbol{x_{jk}}$ are the k-th features of $\boldsymbol{x_i}$ and $\boldsymbol{x_j}$, respectively.

Then select the k-nearest neighbors to each test sample based on the calculated distances. These k-nearest neighbors form the validation set for that test sample. By adopting this approach, K-NN enables the evaluation of classifier performance within local regions, capturing localized patterns and nuances in the data. This methodology offers valuable insights into how classifiers perform under varying conditions, ultimately aiding in the comprehensive analysis of classification results in DES scenarios.

3) CLASSIFIER SELECTION

Based on the RoC defined in *2)*, the optimal classifier is assigned to each unknown region through certain evaluation indicators and selection methods.

C. CONSISTENCY CONSTRAINED SAMPLE AUGMENTATION (SP-SA)

Firstly, in the DKCRCS method, a most suitable classifier is selected for each RoC. Secondly, based on the principle of SP segmentation, the unknown labeled samples within each RoC are expected to be similar. Therefore, *SP-SA* utilizes this similarity assumption and estimates the labels for each RoC based on the results of the dynamically selected classifiers. Specifically, it aggregates the classification results for each RoC and selects the class with the highest occurrence as the more confident classification result. Finally, to complete the semi-supervised learning process, SP-SA propagates the highly confident classification results as labels to the unknown samples within each RoC, assigning them to the corresponding classes. This is equivalent to using a portion of the RoC samples for supervised learning to improve the classifier's performance. After the label propagation, the newly labeled RoCs are added to the original training samples, forming an augmented training set. These

additional RoC samples will help the classifier better learn specific features and decision boundaries, thereby enhancing the classification performance.

D. DYNAMIC MULTI-KERNEL CRC CLASSIFIER SELECTION BASED ON AUGMENTED SAMPLES (DKCRCS-SA)

1) MULTI-KERNEL CRC CONSTRUCTED

The process of classifier pool construction is the same as the DKCRCS algorithm.

2) ROC DEFINED BASED ON AUGMENTED SAMPLES

The difference here compared to DKCRCS is that we use the augmented samples as new training data to construct the Receiver Operating Characteristic (RoC), thereby increasing the reliability of RoC.

3) CLASSIFIER SELECTION

The process of classifier selection is the same as the DKCRCS algorithm.

E. MULTI-SCALE SP SEGMENTATION SEMI-SUPERVISED ENSEMBLE (MS-SP-SSE)

While the combined approach of SP segmentation and dynamic classifier selection for data augmentation can partially address the small sample problem in hyperspectral image classification, different numbers of SP segments can lead to some negative effects on data augmentation, mainly in the following aspects:

1) IMBALANCED SAMPLE PROBLEM

If the number of SP segments is too small, it may result in an imbalanced sample problem. Some classes of SP blocks may be highly prevalent in the images, while others may be very rare. This could lead to an unequal distribution of samples during data augmentation, with some classes having too many samples and others having too few, potentially affecting the training and performance of the model.

2) INFORMATION LOSS

A low number of SP segments may cause information loss. If the SP blocks are too large, data augmentation may lead to the loss of fine-grained details and local information. This could result in the model performing poorly on tasks that require fine-grained classification.

The paper proposes a multi-scale SP segmentation ensemble method to further improve data augmentation and enhance the performance of the model. This approach overcomes the limitations of a single SP segmentation result while utilizing ensemble techniques to merge multiple scales of SP segmentation results, thereby improving the stability and generalization ability of the classification model. The specific steps are as follows:

3) CONSTRUCTING DIFFERENT-SCALE SP SEGMENTATION MODELS

Firstly, multiple SP segmentation models are built using different SP segmentation parameters, such as SP block size and the number of SPs. Each model generates a set of SP segmentation results at different scales, reflecting various levels of features in the image.

4) DATA AUGMENTATION

For each scale of SP segmentation result, corresponding augmentation methods are employed to increase the number of samples in the dataset. Various data augmentation techniques, such as translation, rotation, mirroring, and color transformations, can be used to generate more training samples.

5) ENSEMBLE OF MULTI-SCALE SP SEGMENTATION RESULTS

The SP segmentation results from different scales are integrated. This can be achieved through a simple voting scheme or a more complex weighted voting method to determine the final SP segmentation result. The integrated result will comprehensively capture the image features and may alleviate the deficiencies of a single scale.

By constructing a multi-scale SP segmentation ensemble method, it becomes possible to leverage different scale feature information and increase dataset diversity through data augmentation. This helps to improve the issue of imbalanced data, reduce information loss, and enhance the model's performance in high-spectral image classification tasks. Additionally, the ensemble method contributes to improving the model's robustness and stability, further enhancing the reliability of the classification results.

IV. EXPERIMENTS AND ANALYSIS

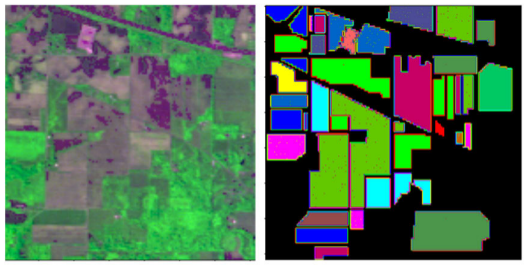

FIGURE 2. (a) False-color image and (b) ground truth of the Indian Pines data set.

A. DATASET DESCRIPTION

The Indian Pines dataset comprises 224 spectral bands collected by the AVIRIS sensor situated in northwest Indiana. These spectral bands cover wavelengths ranging from 0.4 to 2.5 um. After eliminating water absorption bands, the dataset is left with 200 effective bands. The image encompasses

TABLE 1. Sixteen classes of the indian pines data set.

Class	Name	Samples
1	Alfalfa	46
2	Corn-no till	1428
3	Corn-min till	830
4	Corn	237
5	Grass/pasture	483
6	Grass/trees	730
7	Grass/pasture-mowed	28
8	Hay-windrowed	478
9	Oats	20
10	Soybean-no till	972
11	Soybean-min till	2455
12	Soybeans-clean	593
13	Wheat	205
14	Woods	1265
15	Buildings-grass	386
16	Stone-steel-towers	93

TABLE 2. Sixteen classes of the purdue data set.

Class	NAME	Class
1	Road	1287
2	Grass	1114
3	Shadow	219
4	Soil	379
5	Tree	1351
6	Roof	1285

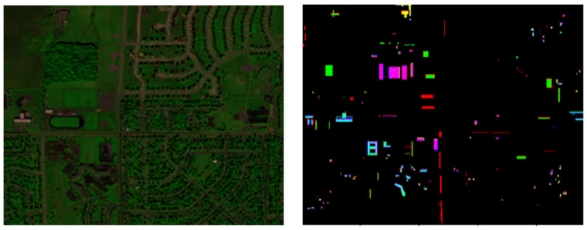

FIGURE 3. (a) False-color image and (b) ground truth of the Purdue data set.

16 classes, and each pixel is represented in a 145 × 145 format with a spatial resolution of 20 m. A detailed description of each class is provided in Table 1, and visual representations of the dataset can be observed in FIGURE 2.

TABLE 3. Sixteen classes of the KSC set.

Class	Name	Samples
1	scrub	761
2	Willow swamp	243
3	Cabbage palm hammock	256
4	Cabbage palm/oak hammock	252
5	slashpine	161
6	oak/Broadleaf hammock	229
7	Hardwood swamp	105
8	Graminoid marsh	431
9	Spartina marsh	520
10	Cattail marsh	404
11	Salt marsh	419
12	Mudflats	503
13	water	927

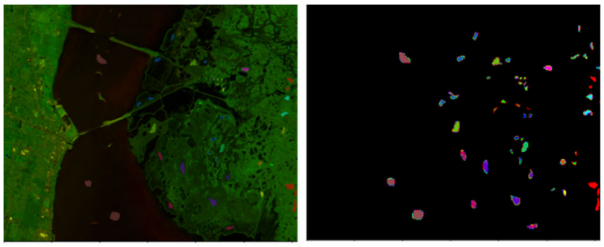

FIGURE 4. (a) False-color image and (b) ground truth of the KSC data set.

The second set of experimental data is the hyperspectral remote sensing data of Purdue University's West Lafayette campus. This data was collected on September 30, 1999, using the Hyperspectral Mapper (HYMAP) sensor aboard an airborne platform. The spectral range of this data spans from 450 nanometers to 2480 nanometers, with a relatively high spatial resolution of 3.6 meters. After excluding water-absorbing bands, a total of 126 spectral bands are provided, and the spatial dimensions are 377 pixels × 512 pixels. The main land cover classes include Road, Grass, and Shadow, among others. The corresponding false-color image of this hyperspectral remote sensing data is shown in Figure 2.6 (a), while the ground truth data is presented in FIGURE 3. The specific land cover types, their corresponding labels, and sample quantities are listed in Table 2.

The third image employed is the KSC data set, obtained using the AVIRIS sensor. It comprises 224 spectral bands covering wavelengths from 0.4 to 2.5 um. Following the elimination of the absorbent band, the image encompasses 13 distinct classes and 176 bands, with a resolution of 512 × 614 pixels and a spatial resolution of 18 m. Table 3 provides a detailed description of these classes, while FIGURE 4. displays the corresponding images.

TABLE 4. Classification accuracy of different methods for indian pines (25 training samples were randomly selected for each class).

class	Train	SVM	RF	CRC	ProCRC	GBDT	LightGBM	Meta-DES	DKCRCS	DMCRCS-SA	MS-SP-SSE
1	25	88.24	88.24	64.71	100.00	70.59	94.12	88.24	100.00	100.00	94.12
2	25	41.26	30.46	44.38	30.38	34.16	38.14	53.01	47.21	47.21	48.01
3	25	48.91	2.94	41.61	45.20	53.65	56.59	46.99	52.75	52.75	51.73
4	25	53.85	38.97	40.51	37.44	38.97	52.82	51.28	65.64	65.64	74.87
5	25	75.06	40.05	79.41	83.98	65.90	69.79	81.24	81.24	81.24	88.33
6	25	79.85	57.65	80.74	85.88	75.59	83.68	83.97	88.82	88.82	92.21
7	25	75.00	100.00	75.00	75.00	75.00	100.00	100.00	100.00	100.00	100.00
8	25	89.98	54.78	82.05	30.30	77.62	89.28	90.91	93.71	93.71	97.90
9	25	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00	100.00
10	25	57.75	61.32	48.54	66.85	50.49	55.25	57.64	64.46	64.46	73.78
11	25	47.20	52.14	40.63	36.68	44.66	46.32	44.87	53.43	53.43	57.71
12	25	48.90	22.89	49.08	44.14	47.25	55.31	48.90	67.95	67.95	75.09
13	25	90.85	91.46	87.80	98.78	75.61	90.85	90.24	93.29	93.29	98.17
14	25	86.26	72.84	74.81	89.96	81.40	81.65	89.71	88.31	88.31	92.43
15	25	50.29	17.94	55.00	35.59	51.76	51.18	41.47	60.29	60.29	68.24
16	25	98.08	88.46	92.31	100.00	96.15	96.15	98.08	96.15	96.15	94.23
OA		59.43	46.28	54.97	53.46	54.91	58.94	61.06	65.90	72.22	70.13
AA		70.72	57.51	66.04	66.26	64.93	72.57	72.91	78.33	83.39	81.67
Kappa		0.54	0.39	0.49	0.49	0.49	0.54	0.56	0.66	0.72	0.66

TABLE 5. Classification accuracy of different methods for Purdue (25 training samples were randomly selected for each class).

class	Train	SVM	RF	CRC	ProCRC	GBDT	LightGBM	Meta-DES	DKCRCS	DMCRCS-SA	MS-SP-SSE
1	25	78.42	80.36	91.84	81.97	86.34	81.57	83.51	76.72	84.64	85.04
2	25	94.37	89.95	78.78	81.78	90.99	92.86	94.84	94.46	95.12	95.49
3	25	98.30	98.86	24.43	7.95	95.45	98.30	98.30	97.16	92.61	90.91
4	25	98.80	98.50	99.70	96.40	98.80	97.00	99.10	99.70	99.70	100.00
5	25	96.70	96.85	98.92	100.00	95.85	95.16	97.62	98.69	99.69	99.92
6	25	85.26	51.98	78.54	90.69	90.77	82.43	82.43	90.85	92.31	92.79
OA		89.55	81.47	86.16	86.80	91.68	88.84	90.41	90.97	93.36	93.66
AA		91.97	86.08	78.70	76.47	93.03	91.22	92.63	92.93	94.01	94.03
Kappa		0.87	0.77	0.83	0.83	0.90	0.86	0.88	0.89	0.92	0.92

B. EXPERIMENT SETUP

To ensure the comparability and fairness, the compared and proposed algorithms are under the same experimental conditions.

1. The number of SP N is set {1000, 1500, 2000, 2500, 3000}. The regularization parameter λ of multi-kernel CRC is set {1e-1, 1e-2, 1e-3, 1e-4, 1e-5}.

2. In order to assess the effectiveness of the proposed algorithm, the study employed a variety of classification algorithms for comparison. The baseline algorithms included traditional machine learning techniques like SVM [35], RF [35], CRC, and ProCRC [36], while more advanced ensemble methods, such as GBDT [37]and LightGBM [38], were also considered. Additionally, the paper utilized the META-DES [39], [40], [41]algorithm as an advanced comparative approach.

C. EXPERIMENT RESULTS

Table 4 shows the classification accuracy results for the Indian Pines dataset, including Overall Accuracy (OA),

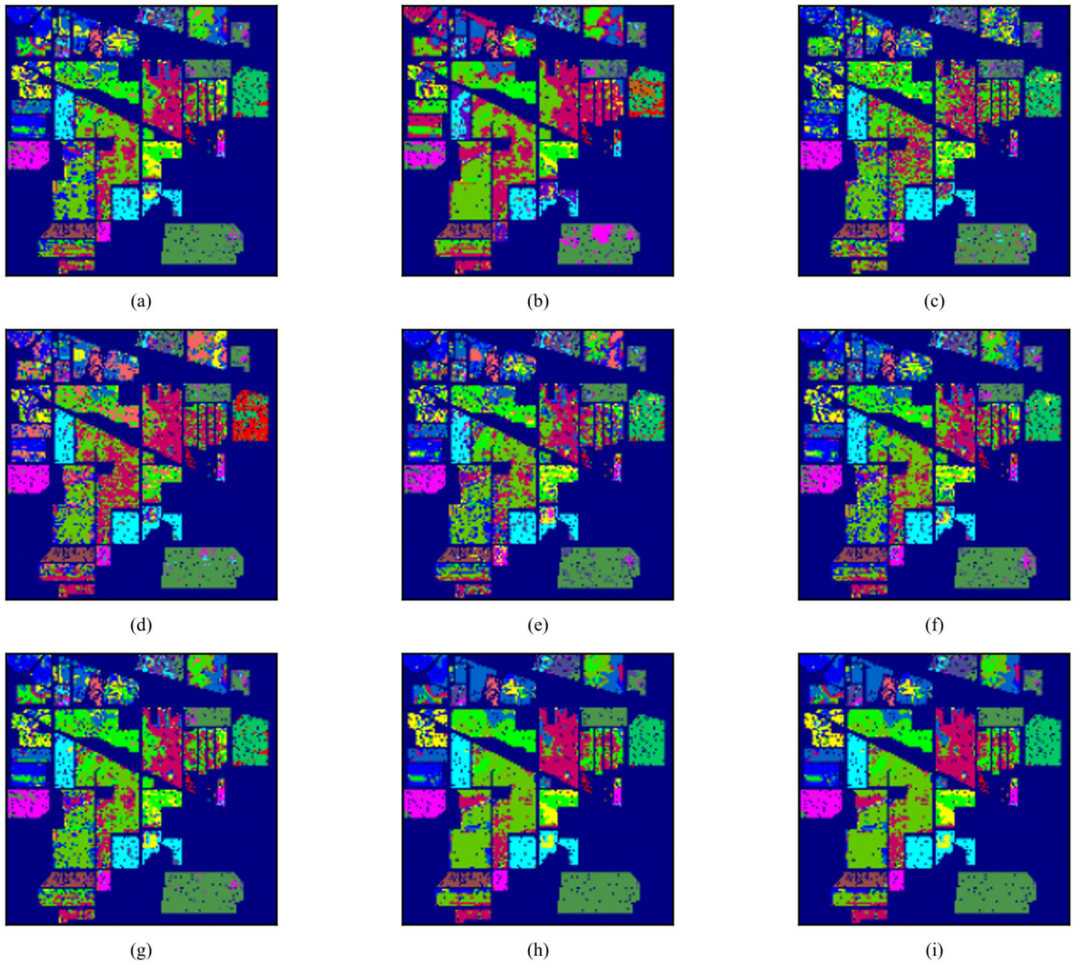

FIGURE 5. Classification maps obtained by different models of Indian Pines.

Average Accuracy (AA), and Kappa index. And The classification maps generated by the proposed methods and other comparative algorithms are shown in FIGURE 5. From the table, it can be observed that the Meta-DES algorithm achieved the highest OA of 61.06%. However, the three proposed algorithms, namely DKCRCS, DMCRCS-SA, and MS-SP-SSE, outperformed Meta-DES with OA values of 65.90%, 72.22%, and 70.13%, respectively. This indicates that the three proposed algorithms demonstrated significant improvements in classification accuracy compared to the comparative Meta-DES algorithm on the Indian Pines dataset. Regarding AA, the Meta-DES algorithm performed reasonably well with an AA of 72.91%. In contrast, the three proposed algorithms, DKCRCS, DMCRCS-SA, and MS-SP-SSE, achieved higher AAs of 78.33%, 83.39%, and 81.67%, respectively. This shows that the proposed algorithms surpassed Meta-DES in terms of average accuracy. Furthermore, looking at the Kappa index, the Meta-DES algorithm attained a Kappa value of 0.56, while the three proposed algorithms, DKCRCS, DMCRCS-SA, and MS-SP-SSE, achieved Kappa values of 0.66, 0.72, and 0.66, respectively. As Kappa considers the random classification factor when evaluating classifier performance, increasing the Kappa index indicates higher stability and accuracy of the classifier. The three proposed algorithms demonstrated significant improvements in Kappa index compared to Meta-DES.In conclusion, the analysis of the classification results for the Indian Pines dataset highlights the superiority of the three proposed algorithms, DKCRCS, DMCRCS-SA, and MS-SP-SSE, over the comparative Meta-DES algorithm in terms of OA, AA, and Kappa index. Thus, these algorithms exhibit significant advantages on the Indian Pines dataset and are expected to achieve excellent classification performance on other similar datasets.

Table 5 presents the classification accuracy of various algorithms on the Purdue dataset. The classification maps generated by the proposed methods and other comparative algorithms are shown in FIGURE 6. The evaluation metrics used are OA, AA, and Kappa index. Among the comparison algorithms, GBDT achieves the highest OA of 91.68%, followed by LightGBM with 88.84%, and SVM with 89.55%. However, the proposed algorithms, namely

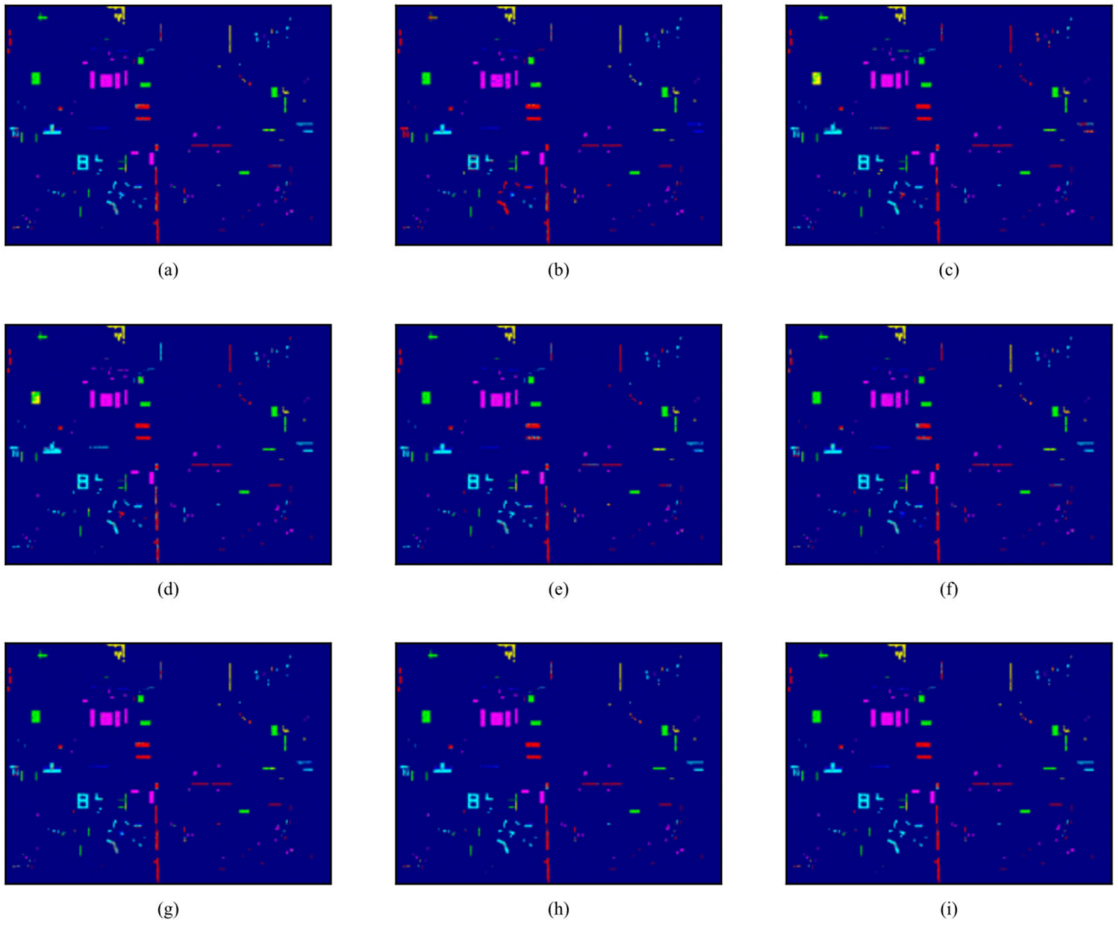

FIGURE 6. Classification maps obtained by different models of Purdue.

DKCRCS, DMCRCS-SA, and MS-SP-SSE, outperform the comparison algorithms. They achieve higher OA values of 90.97%, 93.36%, and 93.66%, respectively. The proposed algorithms also demonstrate superior AA compared to the comparison algorithms. The highest AA among the comparison algorithms is obtained by SVM with 91.97%, while the proposed algorithms achieve AA values of 92.93%, 94.01%, and 94.03%, respectively, showing a significant improvement. Furthermore, when considering the Kappa index, which measures the agreement beyond chance, the proposed algorithms again display better results compared to the comparison algorithms. The Kappa values for the proposed algorithms are 0.89, 0.92, and 0.92, indicating substantial agreement, whereas the comparison algorithms yield Kappa values ranging from 0.77 to 0.90. Overall, the analysis of the provided data reveals that the proposed algorithms, DKCRCS, DMCRCS-SA, and MS-SP-SSE, demonstrate notable advantages over the comparison algorithms SVM, RF, CRC, ProCRC, GBDT, LightGBM, and Meta-DES in terms of classification accuracy. They consistently outperform the comparison algorithms in terms of OA, AA, and Kappa, showcasing their effectiveness in classifying the Purdue dataset. These results indicate that the proposed methods have the potential to be valuable and competitive approaches for classification tasks, offering improved performance and accuracy compared to well-known and widely-used algorithms.

Table 6 presents the classification results for the KSC dataset using various classification algorithms, including SVM, RF, CRC, ProCRC, GBDT, LightGBM, Meta-DES, DKCRCS, DMCRCS-SA, and MS-SP-SSE. The classification maps generated by the proposed methods and other comparative algorithms are shown in FIGURE 7. The evaluation metrics used to assess the performance of these algorithms are OA, AA, and kappa.From Table 5, we can observe that the proposed algorithms (DKCRCS, DMCRCS-SA, and MS-SP-SSE) demonstrate competitive classification performance compared to the traditional algorithms (SVM, RF, CRC, ProCRC, GBDT, LightGBM, and Meta-DES). Specifically, the proposed algorithms achieve OA values of 83.26%, 91.44%, and 91.68%, while the best-performing traditional algorithm, LightGBM, obtains an OA of 86.23%.

TABLE 6. Classification accuracy of different methods for KSC (25 training samples were randomly selected for each class).

class	Train	SVM	RF	CRC	ProCRC	GBDT	LightGBM	Meta-DES	DKCRCS	DMCRCS-SA	MS-SP-SSE
1	25	88.64	31.14	87.10	65.64	83.17	88.64	89.90	87.52	98.60	98.04
2	25	81.54	40.00	84.10	100.00	79.49	80.00	82.56	79.49	86.67	85.64
3	25	89.57	81.52	92.42	85.31	78.20	84.36	91.00	91.00	93.36	94.31
4	25	60.39	0.00	34.78	2.90	57.00	49.28	53.62	43.96	50.24	40.58
5	25	56.78	9.32	57.63	0.00	64.41	70.34	65.25	55.08	69.49	67.80
6	25	53.51	35.14	69.73	44.32	52.97	61.08	50.81	47.57	49.73	52.97
7	25	95.65	69.57	98.55	37.68	75.36	85.51	95.65	94.20	100.00	100.00
8	25	72.92	16.93	66.93	65.89	69.27	79.95	77.86	64.84	88.54	91.67
9	25	85.11	58.09	69.15	0.85	76.38	87.66	86.81	86.81	95.53	95.96
10	25	81.79	61.90	84.59	98.04	86.55	91.60	82.35	91.04	95.80	98.04
11	25	93.53	89.49	98.11	95.96	93.26	95.15	92.99	93.53	98.92	98.38
12	25	83.22	55.99	68.85	71.46	79.96	85.40	82.79	79.30	93.90	97.17
OA		84.73	56.80	81.16	67.39	81.81	86.23	85.45	83.26	91.44	91.68
AA		80.17	49.93	77.74	58.95	76.57	81.36	80.88	77.96	86.21	86.15
Kappa		0.83	0.52	0.79	0.64	0.80	0.85	0.84	0.81	0.90	0.91

The AA values also reveal the superiority of the proposed algorithms. They achieve AA scores of 77.96%, 86.21%, and 86.15%, which outperform all the traditional algorithms, including LightGBM with an AA of 81.36%. Furthermore, the Kappa values further emphasize the advantage of the proposed algorithms. The proposed methods achieve Kappa values of 0.81, 0.90, and 0.91, respectively, surpassing all the traditional algorithms in Table 5, where the highest Kappa value among them is 0.85 for LightGBM. In summary, the analysis of Table 5 demonstrates that the proposed algorithms (DKCRCS, DMCRCS-SA, and MS-SP-SSE) show remarkable superiority over the traditional algorithms (SVM, RF, CRC, ProCRC, GBDT, LightGBM, and Meta-DES) in terms of overall accuracy, average accuracy, and kappa. These results suggest that the proposed algorithms offer more effective and accurate classification solutions for the Purdue dataset compared to the existing methods.

D. PARAMETER ANALYSIS

1) REGULARIZATION PARAMETER λ FOR DKCRCS

From FIGURE 8, it can be observed that the classification accuracy (OA) changes with different values of parameter λ in the Indian Pines, Purdue, and KSC datasets. FIGURE 8 (1) reveals that for the Indian Pines dataset, the OA generally increases as λ decreases from 1e-1 to 1e-4, but it slightly drops when λ reaches 1e-5. This trend suggests that a smaller λ tends to contribute to better accuracy in this dataset. And the optimal parameter value of λ is 1e-4. On the other hand, FIGURE 8 (2) shows that the Purdue dataset achieves higher OA values when λ is set to 1e-4 and 1e-5 compared to λ values of 1e-1 and 1e-2. This indicates that in the Purdue dataset, a larger λ value may be more favorable for obtaining higher classification accuracy. And the optimal parameter value of λ is 1e-5. Moreover, FIGURE 8 (3) illustrates that the KSC dataset exhibits a steady increase in OA with decreasing λ, implying that smaller λ values are more beneficial for classification accuracy in the KSC dataset. Overall, the impact of the parameter λ on classification accuracy varies across different datasets. And the optimal parameter value of λ is 1e-3. For the Indian Pines dataset, the OA generally increases with smaller λ values but slightly decreases with an extremely small λ (1e-5). For the Purdue dataset, higher OA values are achieved with larger λ values (1e-4 and 1e-5). In contrast, the KSC dataset experiences a consistent improvement in OA with smaller λ values. The results also highlight the robustness of the proposed algorithm to its parameters. Although the choice of λ can influence classification accuracy, the differences in OA for different λ values are relatively small, indicating that the algorithm is less sensitive to parameter tuning. This robustness is advantageous for real-world wetland classification problems, as the proposed algorithm can achieve high classification performance without requiring extensive effort in parameter optimization.

2) NUMBER OF SP N FOR DMCRCS-SA

From FIGURE 9, we can observe the impact of different values of parameter N on the classification accuracy (OA) for three datasets: Indian Pines, Purdue, and KSC. FIGURE 9 (1) present the effect of different values of parameter N

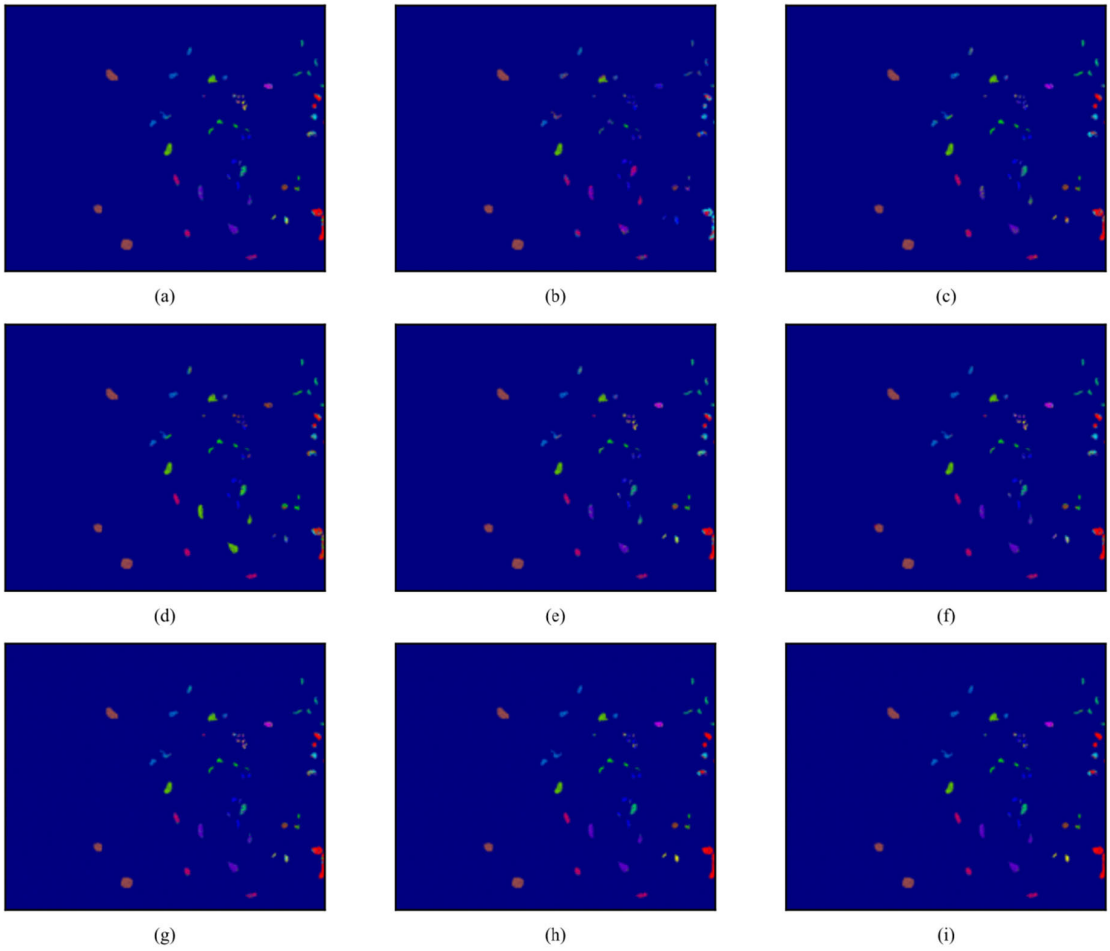

FIGURE 7. Classification maps obtained by different models of KSC.

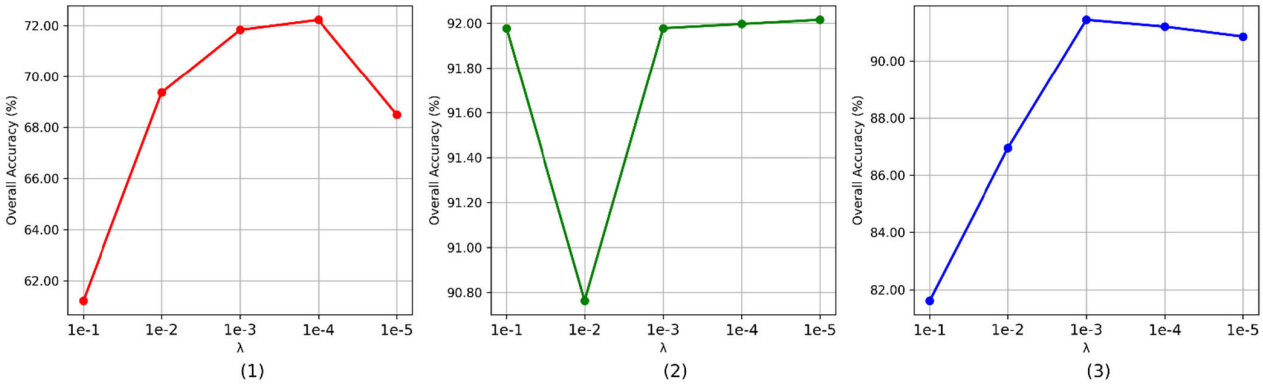

FIGURE 8. The effect of parameter λ on OA for DKCRCS, with 25 samples of each class in Indian Pines, Purdue and KSC.

on accuracy for the Indian Pines dataset. The x-axis represents five different values of N, and the y-axis shows the corresponding classification accuracy (OA). As N varies from 1000 to 3000, the OA values for Indian Pines dataset are 72.23%, 70.95%, 70.13%, 70.13%, and 70.13%, respectively. It can be observed that changing the value of N has a significant impact on the classification accuracy, as the OA values vary notably along the x-axis.

Similarly, FIGURE 9 (2) illustrate the effect of different values of parameter Non the classification accuracy for

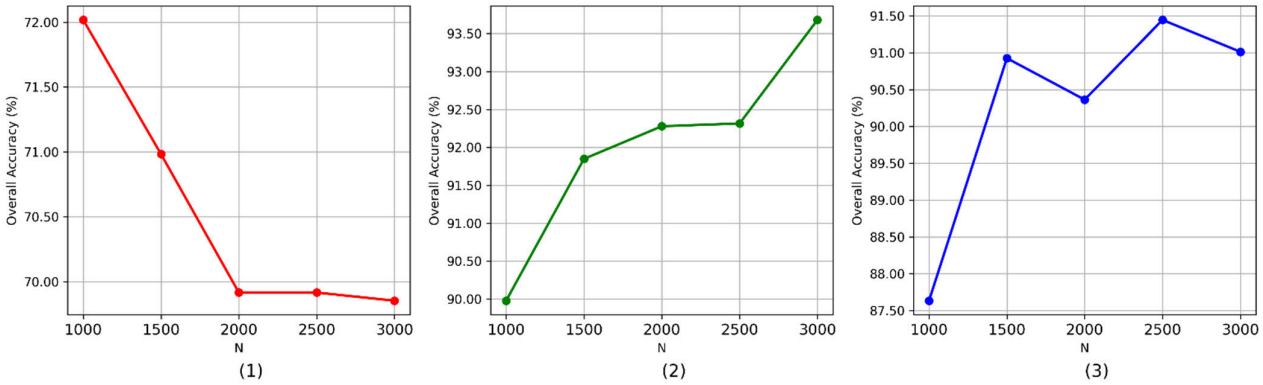

FIGURE 9. The effect of parameter *N* on OA for DKCRCS-SA, with 25 samples of each class in Indian Pines, Purdue and KSC.

the Purdue dataset. As N changes from 1000 to 3000, the OA values for Purdue dataset are 89.98%, 91.85%, 92.28%, 92.31%, and 93.68%, respectively. Here, we can observe a clear increasing trend in OA as the value of N decreases along the x-axis.

Furthermore, FIGURE 9 (3) demonstrate the impact of different values of parameter N on accuracy for the KSC dataset. As N varies from 1000 to 3000, the OA values for the KSC dataset are 87.64%, 90.93%, 90.36%, 91.45%, and 91.01%, respectively. Similar to the Purdue dataset, the KSC dataset also shows an increasing trend in OA as Ndecreases. Overall, the analysis of FIGURE 9 indicates that parameter N has a considerable influence on the classification accuracy for all three datasets. The accuracy improves as N decreases. However, it is worth noting that the effect of N on classification accuracy may vary across different datasets, as demonstrated by the different trends in OA observed for the Indian Pines dataset compared to the Purdue and KSC datasets.

V. CONCLUSION

In this paper, a novel hyperspectral image classification method is proposed, which effectively addresses the issue of low-confidence sample augmentation in existing semi-supervised methods by combining dynamic ensemble selection with SP consistency constraints. The use of SP block division transforms the hyperspectral image into multiple independent classification tasks. The dynamic ensemble selection strategy is applied to choose high-confidence samples from the unlabeled data, and pseudo-labels are introduced for sample expansion, thereby enhancing the quality of the training samples. Furthermore, employing a multiple-kernel collaborative representation classifier as the base classifier fully utilizes multiple feature representations of samples, strengthening the classifier's ability to capture sample similarities and correlations, and further improving the classification performance. The experimental results demonstrate the superior classification accuracy of the proposed method on multiple hyperspectral datasets compared to traditional methods, indicating its practicality and potential for application as an effective solution to hyperspectral image classification problems. However, it is also essential to consider the adaptability of parameter adjustments on different datasets to avoid risks of overfitting or underfitting. Overall, this method provides valuable exploration and insights for research in the field of hyperspectral image classification. While this study has achieved satisfactory results and made significant progress on small-scale datasets, we acknowledge the potential for broader impact and practical applications when applying this method in larger and more complex data environments. Future work will be dedicated to optimizing and scaling our model to accommodate large-scale datasets and addressing various challenges associated with this endeavor. These efforts will contribute to a deeper understanding and resolution of significant real-world issues, thereby providing sustainable and broader value for scientific research and practical applications.

REFERENCES

[1] D. Landgrebe, "Hyperspectral image data analysis," *IEEE Signal Process. Mag.*, vol. 19, no. 1, pp. 17–28, Jan. 2002, doi: 10.1109/79.974718.

[2] H. Su, Z. Wu, H. Zhang, and Q. Du, "Hyperspectral anomaly detection: A survey," *IEEE Geosci. Remote Sens. Mag.*, vol. 10, no. 1, pp. 64–90, Mar. 2022, doi: 10.1109/MGRS.2021.3105440.

[3] 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, Jul. 2019, doi: 10.3390/s19143071.

[4] S. Lu, Y. Shimizu, J. Ishii, S. Funakoshi, I. Washitani, and K. Omasa, "Estimation of abundance and distribution of two moist tall grasses in the Watarase wetland, Japan, using hyperspectral imagery," *ISPRS J. Photogramm. Remote Sens.*, vol. 64, no. 6, pp. 674–682, Nov. 2009, doi: 10.1016/j.isprsjprs.2009.06.003.

[5] Y. Chen, Y. Wang, Y. Gu, X. He, P. Ghamisi, and X. Jia, "Deep learning ensemble for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 6, pp. 1882–1897, Jun. 2019, doi: 10.1109/JSTARS.2019.2915259.

[6] J. Li, X. Huang, P. Gamba, J. M. Bioucas-Dias, L. Zhang, J. A. Benediktsson, and A. Plaza, "Multiple feature learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1592–1606, Mar. 2015, doi: 10.1109/TGRS.2014.2345739.

[7] M. S. Aydemir and G. Bilgin, "Semisupervised hyperspectral image classification using small sample sizes," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 621–625, May 2017, doi: 10.1109/LGRS.2017.2665679.
[8] W. Luo, C. Zhang, Y. Li, F. Yang, D. Zhang, and Z. Hong, "Deeply-supervised pseudo learning with small class-imbalanced samples for hyperspectral image classification," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 112, Aug. 2022, Art. no. 102949, doi: 10.1016/j.jag.2022.102949.
[9] M. Romaszewski, P. Głomb, and M. Cholewa, "Semi-supervised hyperspectral classification from a small number of training samples using a co-training approach," *ISPRS J. Photogramm. Remote Sens.*, vol. 121, pp. 60–76, Nov. 2016, doi: 10.1016/j.isprsjprs.2016.08.011.
[10] S. S. Sawant and M. Prabukumar, "A review on graph-based semi-supervised learning methods for hyperspectral image classification," *Egyptian J. Remote Sens. Space Sci.*, vol. 23, no. 2, pp. 243–248, Aug. 2020.
[11] J. Gu, L. Jiao, F. Liu, X. Zhang, X. Tang, and P. Chen, "Multi-feature weighted sparse graph for SAR image analysis," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 2, pp. 881–891, Feb. 2020.
[12] X. He, Y. Chen, and P. Ghamisi, "Dual graph convolutional network for hyperspectral image classification with limited training samples," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5502418, doi: 10.1109/TGRS.2021.3061088.
[13] S. Jozdani, D. Chen, D. Pouliot, and B. Alan Johnson, "A review and meta-analysis of generative adversarial networks and their applications in remote sensing," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 108, Apr. 2022, Art. no. 102734, doi: 10.1016/j.jag.2022.102734.
[14] A. Jamali, M. Mahdianpari, F. Mohammadimanesh, and S. Homayouni, "A deep learning framework based on generative adversarial networks and vision transformer for complex wetland classification using limited training samples," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 115, Dec. 2022, Art. no. 103095, doi: 10.1016/j.jag.2022.103095.
[15] J. Kim, T. Kim, and J.-G. Ryu, "Multi-source deep data fusion and super-resolution for downscaling sea surface temperature guided by generative adversarial network-based spatiotemporal dependency learning," *Int. J. Appl. Earth Observ. Geoinf.*, vol. 119, May 2023, Art. no. 103312, doi: 10.1016/j.jag.2023.103312.
[16] X. Wang, K. Tan, Q. Du, Y. Chen, and P. Du, "CVA2E: A conditional variational autoencoder with an adversarial training process for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 8, pp. 5676–5692, Aug. 2020, doi: 10.1109/TGRS.2020.2968304.
[17] Y. Ding, Z. Zhang, X. Zhao, Y. Cai, S. Li, B. Deng, and W. Cai, "Self-supervised locality preserving low-pass graph convolutional embedding for large-scale hyperspectral image clustering," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5536016, doi: 10.1109/TGRS.2022.3198842.
[18] C. Yu, J. Huang, M. Song, Y. Wang, and C.-I. Chang, "Edge-inferring graph neural network with dynamic task-guided self-diagnosis for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5535613, doi: 10.1109/TGRS.2022.3196311.
[19] C.-D. Wang, J.-H. Lai, and P. S. Yu, "Multi-view clustering based on belief propagation," *IEEE Trans. Knowl. Data Eng.*, vol. 28, no. 4, pp. 1007–1021, Apr. 2016, doi: 10.1109/TKDE.2015.2503743.
[20] B. Palsson, M. O. Ulfarsson, and J. R. Sveinsson, "Synthetic hyperspectral images with controllable spectral variability and ground truth," *IEEE Geosci. Remote Sens. Lett.*, vol. 19, pp. 1–5, 2022, doi: 10.1109/LGRS.2022.3150245.
[21] Y. Su, J. Li, A. Plaza, A. Marinoni, P. Gamba, and S. Chakravortty, "DAEN: Deep autoencoder networks for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 7, pp. 4309–4321, Jul. 2019, doi: 10.1109/TGRS.2018.2890633.
[22] L. Jiao, W. Sun, G. Yang, G. Ren, and Y. Liu, "A hierarchical classification framework of satellite multispectral/hyperspectral images for mapping coastal wetlands," *Remote Sens.*, vol. 11, no. 19, p. 2238, Sep. 2019, doi: 10.3390/rs11192238.
[23] K. Zhang, Y. Su, X. Guo, L. Qi, and Z. Zhao, "MU-GAN: Facial attribute editing based on multi-attention mechanism," 2020, *arXiv:2009.04177*.
[24] H. Xu, H. Zhang, W. He, and L. Zhang, "Superpixel-based spatial–spectral dimension reduction for hyperspectral imagery classification," *Neurocomputing*, vol. 360, pp. 138–150, Sep. 2019, doi: 10.1016/j.neucom.2019.06.023.
[25] J. Li, H. Zhang, and L. Zhang, "Efficient superpixel-level multitask joint sparse representation for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 10, pp. 5338–5351, Oct. 2015, doi: 10.1109/TGRS.2015.2421638.
[26] S. Jia, X. Deng, J. Zhu, M. Xu, J. Zhou, and X. Jia, "Collaborative representation-based multiscale superpixel fusion for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 10, pp. 7770–7784, Oct. 2019, doi: 10.1109/TGRS.2019.2916329.
[27] L. Zhang, M. Yang, and X. Feng, "Sparse representation or collaborative representation: Which helps face recognition?" in *Proc. Int. Conf. Comput. Vis.*, Nov. 2011, pp. 471–478, doi: 10.1109/ICCV.2011.6126277.
[28] W. Li, Q. Du, F. Zhang, and W. Hu, "Collaborative-representation-based nearest neighbor classifier for hyperspectral imagery," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 389–393, Feb. 2015, doi: 10.1109/LGRS.2014.2343956.
[29] Z.-Q. Li, J. Sun, X.-J. Wu, and H.-F. Yin, "Multiplication fusion of sparse and collaborative-competitive representation for image classification," *Int. J. Mach. Learn. Cybern.*, vol. 11, no. 10, pp. 2357–2369, Oct. 2020, doi: 10.1007/s13042-020-01123-5.
[30] C. Zhao and W. Li, "Collaborative representation-based binary hypothesis model with multi-features learning for target detection in hyperspectral imagery," *J. Indian Soc. Remote Sens.*, vol. 46, no. 5, pp. 847–862, May 2018, doi: 10.1007/s12524-018-0752-8.
[31] M. Yang, L. Zhang, D. Zhang, and S. Wang, "Relaxed collaborative representation for pattern classification," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit.*, Jun. 2012, pp. 2224–2231, doi: 10.1109/CVPR.2012.6247931.
[32] A. H. R. Ko, R. Sabourin, and A. S. Britto Jr., "From dynamic classifier selection to dynamic ensemble selection," *Pattern Recognit.*, vol. 41, no. 5, pp. 1718–1731, May 2008, doi: 10.1016/j.patcog.2007.10.015.
[33] R. M. O. Cruz, H. H. Zakane, R. Sabourin, and G. D. C. Cavalcanti, "Dynamic ensemble selection VS K-NN: Why and when dynamic selection obtains higher classification performance?" in *Proc. 7th Int. Conf. Image Process. Theory, Tools Appl. (IPTA)*, Nov. 2017, pp. 1–6, doi: 10.1109/IPTA.2017.8310100.
[34] S. García, Z.-L. Zhang, A. Altalhi, S. Alshomrani, and F. Herrera, "Dynamic ensemble selection for multi-class imbalanced datasets," *Inf. Sci.*, vols. 445–446, pp. 22–37, Jun. 2018, doi: 10.1016/j.ins.2018.03.002.
[35] H. Lu, H. Su, J. Hu, and Q. Du, "Dynamic ensemble learning with multi-view kernel collaborative subspace clustering for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 2681–2695, 2022, doi: 10.1109/JSTARS.2022.3158761.
[36] J. Liu, Z. Wu, J. Li, A. Plaza, and Y. Yuan, "Probabilistic-kernel collaborative representation for spatial–spectral hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 54, no. 4, pp. 2371–2384, Apr. 2016, doi: 10.1109/TGRS.2015.2500680.
[37] Z. Zhang and C. Jung, "GBDT-MO: Gradient-boosted decision trees for multiple outputs," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 32, no. 7, pp. 3156–3167, Jul. 2021, doi: 10.1109/TNNLS.2020.3009776.
[38] X. Sun, L. Guo, W. Zhang, Z. Wang, and Q. Yu, "Small aerial target detection for airborne infrared detection systems using LightGBM and trajectory constraints," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 9959–9973, 2021, doi: 10.1109/JSTARS.2021.3115637.
[39] R. M. O. Cruz, R. Sabourin, G. D. C. Cavalcanti, and T. I. Ren, "META-DES: A dynamic ensemble selection framework using meta-learning," *Pattern Recognit.*, vol. 48, no. 5, pp. 1925–1935, May 2015, doi: 10.1016/j.patcog.2014.12.003.
[40] R. M. O. Cruz, R. Sabourin, and G. D. C. Cavalcanti, "META-DES.H: A dynamic ensemble selection technique using meta-learning and a dynamic weighting approach," in *Proc. Int. Joint Conf. Neural Netw. (IJCNN)*, Jul. 2015, pp. 1–8, doi: 10.1109/IJCNN.2015.7280594.
[41] R. M. O. Cruz, R. Sabourin, and G. D. C. Cavalcanti, "On meta-learning for dynamic ensemble selection," in *Proc. 22nd Int. Conf. Pattern Recognit.*, Aug. 2014, pp. 1230–1235, doi: 10.1109/ICPR.2014.221.

XIANG GE was born in Huainan, Anhui, in August 1994. He received the degree from the Anhui University of Science and Technology. He is currently with the Huainan City Construction and Development Planning and Design Institute, mainly engaged in oblique photography and remote sensing technology. His research interests include geographic information and remote sensing applications.

XUEXIANG YU was born in Taihu, Anhui. He received the Ph.D. degree in engineering from the School of Geodesy and Surveying Engineering, Wuhan University, in 2002, and the degree from the Postdoctoral Mobile Station of Control Science and Engineering, Southeast University, in 2007. He was a Postdoctoral Researcher. He is currently the Dean and the Deputy Secretary of the Party Committee of the School of Spatial Information and Surveying Engineering, Anhui University of Science and Technology. He is also a Professor and a Ph.D. Supervisor. He is mainly engaged in the teaching and research of satellite positioning technology and application, deformation monitoring automation, software engineering practice, and other aspects. He presided over more than ten vertical scientific research projects, such as the National Natural Fund Project and the Anhui Province Science and Technology Major Project. He has trained nine Ph.D. students and more than 90 master's students. He has 12 sets of software that won the computer software copyright. He holds five invention patents. He published four monographs, such as ''Automatic Monitoring System of Coal Mining Subsidence'' and more than 100 academic articles, such as ''Research on Single Epoch Solution Method of GPS Deformation Monitoring Information.'' He is also a member of the International Mining Survey Association, a member of the International Exchange and Cooperation Working Committee of the China Geographic Information Industry Association, the Director of the Coal Industry Engineering Research Center for Collaborative Monitoring of Mining Environment and Disaster, a member of the Coal Mining Damage Technology Appraisal Committee of the China Coal Society, a member of the Education Working Committee of the China Surveying and Geoinformation Society, a member of the Mining Survey Professional Committee, and a China Satellite Member of the Education and Development Professional Committee of the Navigation and Positioning Association. He won the Third Prize of the Anhui Province Science and Technology Award, the Second Prize of the China Coal Industry Science and Technology Award, the Second Prize of the Satellite Navigation and Positioning Science and Technology Award, and the Third Prize of the Nanjing Science and Technology Progress Award. He is the Chief Editor of three planning textbooks, such as Principles and Applications of GNSS Navigation and Positioning. He is the Deputy Chief Editor of ten textbooks ''Research and Practice on Cultivation of Innovative Talents in Coal Surveying and Mapping Major Under the Background of New Engineering'' won one first prize in teaching achievements of the China Coal Industry Association and ''Construction and Practice of MOOCs on Principles and Applications of Satellite Navigation and Positioning'' won one second prize of teaching achievements of the China Satellite Navigation and Positioning Association.

XU YANG was born in Hanzhong, Shaanxi, in February 1989. He received the Ph.D. degree in geodesy and surveying engineering from the China University of Mining and Technology, in 2019. From 2001 to 2023, he was a Postdoctoral Student with the Nanjing University of Information Science and Technology. He presided over one project of the Natural Science Foundation of Anhui Province, one key project of the Natural Science of Education Department of Anhui Province, three provincial open fund projects, one provincial teaching and research project, one authorized invention patent, and five applied for invention patents. He obtained one software copyright. He participated in the compilation of one planning textbook and published more than 25 academic articles. He is engaged in the teaching and research of satellite navigation and positioning theory and application and warning of typical geological disasters, intelligent mapping, and other aspects.