# Novel Sample Augmentation Approach for Improving Classification Performance With High-Resolution Remote Sensing Imagery

Zhiyong Lv, *Senior Member, IEEE*, Pengfei Zhang, Pei Wang, Weiwei Sun, *Senior Member, IEEE*, Minghua Zhao, and Rui Zhu

***Abstract*—Achieving satisfactory land cover classification performance with high-resolution remote sensing images (HRSIs) usually requires sufficient samples for a supervised classifier. However, labeling sufficient samples is labor-intensive and time-consuming. In this article, a novel sample augmentation approach (NSAA) is proposed to synthesize new samples and improve classification accuracies for HRSI when initial known samples are very limited. First, a very small sample set of each class is prepared manually for the algorithm's initialization. Second, a sample generator based on normal cloud model is proposed, and an adaptive region growing algorithm is suggested to explore some potential samples around a known sample for parameter estimation of the sample generator. Third, to further refine the generated samples around an initial known sample, a near-to-far space constraint strategy (NFSC) is proposed based on the *K*-means clustering algorithm to improve the quality of the generated samples. The proposed sample augmentation approach is incorporated with a classifier iteratively, and a sample balancing strategy is suggested in the iterative progress. Experimental results based on six real HRSIs and compared with eight state-of-the-art methods demonstrate the feasibility and superiorities of the proposed sample augmentation approach. Moreover, the reliability and robustness of the generated samples are verified by popular deep-learning networks and typical traditional classifiers. The improvement achieved by our proposed approach is about 0.12%–0.95% in terms of the overall accuracy (OA).**

***Index Terms*—Land cover classification, remote sensing images, sample generation.**

## I. Introduction

LAND cover classification with high-resolution remote sensing images (HRSIs) has been an important topic for remote sensing applications [1]. In recent years, various HRSI classification approaches have been promoted and widely used in practical applications, such as land cover mapping [2], crop-type classification [3], and water body classification [4]. In a classification task, a label is assigned to each pixel in an image scene. The classification approaches can be divided into methods based on supervision (with training samples) and nonsupervision (without training samples). Many studies have indicated that training samples play an important role in achieving satisfactory classification performance for supervised classifiers [5], [6]. However, supervised classification with HRSIs usually encounters the following challenges.

1) *High Spectral Intraclass Variance Deduces the Separability of Land Cover Classes:* HRSI has some advantages in capturing the ground details and presenting an excellent visual observation [7]. However, higher spatial resolution does not mean higher classification accuracies: a) because high spatial resolution usually captures the small ground objects, such as cars, ships, and even water tanks on roofs, which may be misclassified and reduced the classification accuracies [8] and b) many studies have indicated that a higher spatial resolution strengthens the correlation among pixels and enlarges the intraclass variance, but the variability of different entities within an intraclass is nonlinear [9], [10]. Thus, the uncertain variability of intraclass samples causes challenges in separating different targets [9].
2) *Labeling Many Samples of HRSI for Training a Supervised Classifier Is Time-Consuming and Labor-Intensive:* Samples are required for training supervised classifiers or neural networks, and the test performance of a trained classifier or network is tightly related to the quality of training samples [11], [12]. However, making a high-quality sample set is a challenge for the following reasons. First, visual interpretation is a widely used approach to make samples of HRSI, but collecting high-quality samples needs the professional background knowledge of practitioners, and labeling sufficient samples is time-consuming. Second, the quality of labeling samples via fieldwork may be affected by the precision of the global position system (GPS) [12], [13]. Although GPS is one of the most reliable equipments for location, precisely locating a position from HRSI to a GPS is difficult because imaging and equipment errors are inevitable in the making of orthographic images with

Received 19 August 2025; accepted 12 September 2025. Date of publication 16 September 2025; date of current version 25 September 2025. This work was supported in part by the National Natural Science Foundation of China under Grant 42271385, Grant 41971296, Grant 42122009, and Grant 62272383; and in part by the Interdisciplinary Project of Xi'an University of Technology under Grant 104-206022504. *(Corresponding author: Pei Wang.)*

Zhiyong Lv is with Shaanxi Key Laboratory for Network Computing and Security Technology, School of Computer Science and Engineering, Xi'an University of Technology, Xi'an, Shannxi 710048, China (e-mail: Lvzhiyong_fly@hotmail.com).

Pengfei Zhang and Minghua Zhao are with the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an 710048, China (e-mail: Zhangpengfei_life@hotmail.com; zhaominghua@xaut.edu.cn).

Pei Wang is with Shaanxi Bureau of Surveying and Mapping Geographic Information, Xi'an 710054, China (e-mail: wangpei_firps@163.com).

Weiwei Sun is with the Department of Geography and Spatial Information Techniques, Ningbo University, Ningbo, Zhejiang 315211, China (e-mail: nbsww@outlook.com).

Rui Zhu is with the Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore 138634 (e-mail: zhur@ihpc.a-star.edu.sg).

Digital Object Identifier 10.1109/TGRS.2025.3610323

HRSI. Therefore, how to label samples with high quality remains a challenge for supervised classification with HRSI.

3) *Classification With Remote Sensing Images Is a Typical Imbalance Dataset Classification Problem [14]:* A remote sensing image describes the ground targets on the Earth's surface in a geographical area, and the quantity of the candidate classified targets is usually unknown and unbalanced. Balancing the quantity of samples for different classes helps improve the overall accuracies and the user's accuracy for each class [12], [15]. Therefore, how to consider the balance of samples correspondingly when collecting samples for each class becomes attractive.

The above introduction indicates that achieving satisfactory classification performance with a small number of samples is important for practical application with HRSI. In recent years, various classification methods with small samples have been proposed in remote sensing images. The details of related work were reviewed in the following.

1) *Basic and Intuitive Way to Improve Classification Performance Is Sample Augmentation [16]:* The idea of sample augmentation is to generate synthetic samples directly in order to improve classification performance. One of the typical sample augmentation algorithms is the synthetic minority oversampling technique (SMOTE) [17]. The idea of SMOTE is to generate new synthetic samples close to the minority class in feature space. Due to the simplicity and wide adaptability of SMOTE, it has been extended and improved in many versions, such as borderline-SMOTE and $K$-means-SMOTE [18]. In addition, potential samples around each labeled sample can be explored to augment a known sample set, such as multilabel sample augmentation [19], iterative sample augmentation [11], and suitable neighboring sample exploration [9], [20]. In addition, sample optimization reduces the interclass similarity and improves the utilization of training samples [21]. With the rapid development of computer vision (CV), many sample augmentation methods have provided valuable insights for our research. Random cropping, flipping, and rotation have been widely used to enhance model generalization [22]. Recently, Supermix [23] and Token-mixup [24] have indicated significant improvements for classification by creating new training samples via mixing or cutting images. These methods have demonstrated their effectiveness in handling limited training data and may have potential benefit for remote sensing image classification. In addition, adversarial sample generation with generative adversarial network, which learn the minority class distribution in the training stage and generate representative samples to improve HRSI classification accuracy [25], and sample augmentation based on convolutional neural networks (CNNs) [26]. In summary, many studies have indicated that sample augmentation, including sample synthetics and sample exploration, is an effective way to improve classification performance, and it also has an advantage in avoiding overfitting problems in deep learning classification [27].
2) *Another Way to Enhance the Classification Performance Is via Improving the Quality of Samples Through Noise-Label Detection and Correction:* Sample quality refers to the representative ability of a sample to deliver the diversity and representativeness of a class [28]. Different entities within an intraclass performed with different spectral reflectances. Thus, collecting different points to make samples for a class leads to different sample qualities, which also depends on the practitioner's experience. Some studies have indicated that supervised classification accuracies mainly depend on the quality of the sample and various approaches have been promoted for improving classification accuracies via correcting the sample's quality. For example, Tu et al. [12] and Qian et al. [29] proposed a series of noise-label detection approaches, including a super-pixel-to-pixel weighting distance-based approach, density peak-based approach [12], and dual-channel residual network for noisy label detection. Kang et al. [13] proposed an approach to detect and correct mislabeled samples. Noise-robust neural networks have also been proposed directly to conquer the negative effects of noise on classification performance. Such methods include wavelet integrated CNNs [30] and robust spatial–spectral graphs for noisy labels [31]. In summary, various studies demonstrated that the detection and correction of noise labels can reduce the negative effect of noisy samples and improve classification performance. More information about the quality of samples and consequent classification performance for remote sensing images can be reviewed in [32].
3) *Recently, Some Deep Learning Neural Networks Have Been Proposed to Obtain Satisfactory Performance With Limited Samples:* Due to the advantages of deep learning neural networks in exploring deep and abstract features for enhancing classification performance [33], some deep learning neural networks have been proposed for classification with very limited samples. One of the well-known deep learning neural networks for classification with limited samples is the few-shot learning network, which has been implemented as the dual-metric few-shot learning network [34], spectral–spatial Siamese network [35], and cross-domain few-shot learning [36]. In addition, active learning [37], contrastive learning [38], [39], and multitask deep learning [40] have been applied successfully for remote sensing image classification with small samples. In contrast to the above sample augmentation and sample's quality enhancement for achieving classification with small samples, deep learning networks improve the classification performance with limited samples via optimizing the neural network's structure and training strategy.
4) *In Addition to the Sample Augmentation-Related Work Above, Sample Balancing Also Plays an Important Role in Land Cover Classification With Remote Sensing Images [41]:* Intuitively, the quantity of candidate

classified pixels within an image scene is unpredicted and unknown, and remote sensing images are well-known class imbalance datasets [42]. Various methods have also been proposed for dealing with the classification problem of class imbalance in datasets, such as SMOTE-based deep learning methods [43], spectral–spatial-dependent global learning frameworks [44], and CNN-based imbalance classification methods [42]. These studies indicated that considering the class imbalance phenomenon in classification helps improve classification performance [9], [11]. Therefore, the different numbers of samples must be used for different classes while classifying an imbalanced dataset.

In this article, we concentrate on proposing a sample augmentation approach to synthesize new samples for improving classification performance when initial known samples are very limited. The class imbalance problem is also taken into account in the process of synthesizing new samples. The main contributions of this work can be summarized as follows.

1) A sample augmentation approach is proposed based on a new sample generator with normal cloud model. To the best of our knowledge, the normal cloud model is first used to generate samples and enhance classification performance in HRSI applications. Moreover, a parameter estimation approach is proposed for estimating and adjusting the parameter of the cloud model in each iteration. Therefore, the proposed sample augmentation approach is no-parameter's requirement, but it can generate unlimited virtual new samples for improving classification performance.
2) A sample balancing strategy is proposed to balance the sample's quantity for each class. Land cover classification with remote sensing images is a typical unbalancing classification problem. Our previous studies have shown that adjusting the sample's quantity is effective in conquering the imbalanced class in classification with remote sensing images [9], [11]. Therefore, a sample balancing strategy is proposed to adjust automatically the sample's quantity of each class and seek to balance the user's accuracy of each class.
3) An iterative image classification approach is proposed based on the new sample augmentation approach and sample balancing strategy. From the perspective of methodology, the proposed approach is intuitive and reasonable. Experimental results indicated that the proposed approach has some advantages in improving classification accuracies compared with that of the state-of-the-art methods, and the generated samples are robust and effective for different classifiers.
4) The Novel Sample Augmentation Approach (NSAA) is effective in improving the classification performance of some typical traditional classifiers and popular deep-learning neural networks when the training samples are very limited. In the experiments, two traditional classifiers and four deep learning networks have verified the robustness of the proposed NSSA. Compared to classification without using generated samples, adding generated samples through NSSA can improve classification accuracy and performance.

The rest of this article is organized as follows. Section II elaborates on the details of the proposed method. The extensive experiments and discussion are presented in Section III. Finally, the conclusion is drawn in Section IV.

## II. Proposed Sample Augmentation Approach

### A. Overview

The proposed sample augmentation approach comprises three parts: parameter estimation, sample generator, and refinement of generated samples. An overview of the classification framework based on the proposed sample augmentation is shown in Fig. 1. First, the classical and widely used dimensionality reduction method named principal component analysis (PCA) was adopted for band selection from HRSI. In our proposed approach, the PCA with more than 99% of the information of the raw data was used instead of the original spectral images. Second, an initial sample set was selected manually from the input image, and a classification map was obtained via a supervised classifier with the initial sample set. Then, the proposed sample augmentation is used to synthesize new samples based on the initial sample set, and the newly generated samples are appended to the initial samples. Another classification map is obtained via the supervised classifier with the augmented sample set. Finally, the proposed steps are fused into an iterative algorithm, and the similarity is calculated between the two classification maps with different training sets in each iteration. When the similarity is less than a predefined threshold, the final classification map is obtained.

As shown above, different from many existing semi-supervised approaches, the proposed framework is an iterative classification progress. In each iteration, training samples are updated iteratively with the proposed sample augmentation, and a classification map is obtained correspondingly with augmented samples. The termination of the iteration is determined as follows:

$$\left| \frac{S_{c_k}^{t-1}}{N_{c_k}^{t-1}} - \frac{S_{c_k}^{t}}{N_{c_k}^{t}} \right| \leq \varepsilon \tag{1}$$

where $S_{c_k}^{t}$ and $N_{c_k}^{t}$ denote the quantity of samples and the classified pixel and number for the $c$th class in the $t$th iteration, respectively. The initialization value of $S_{c_k}^{t}$ is larger than 3 for each class. Equation (1) implies that the proposed method terminates the iteration until the pairwise classification maps from two adjacent iterations are similar enough for each class. The proposed approach enhances the classification performance via sample augmentation with the proposed sample generator and adjusts the sample's quality of each class to deal with the imbalance problem in the classification process. The proposed sample augmentation plays an important role in improving classification accuracies. Here, we take the labeled sample set $S^t$ of the $t$th iteration as an example, and three parts of the proposed sample augmentation approach are elaborated in the following.

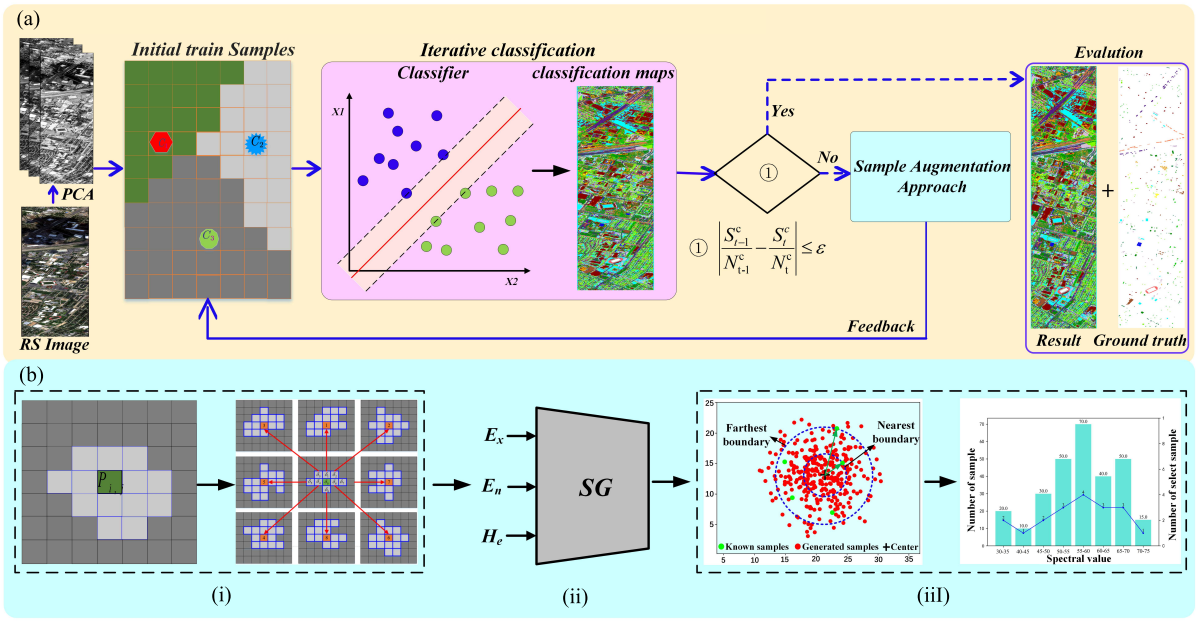



Fig. 1. Flowchart of the classification framework based on the proposed sample augmentation approach. (a) Proposed NSAA for classification with HRSI. (b) Sample augmentation approach (i) generator's parameters spatial constraint and calculation, (ii) sample generator output, and (iii) generated sample refinement and extraction.

### B. Generator's Parameter Estimation

In our proposed sample augmentation approach, three parameters (Ex, En, and He) for the classical normal cloud model are required. Therefore, we first proposed a parameter estimation strategy instead of giving them empirically. For clarity, a sample in the $c_k$ class sample subset $S_{c_k}$ is symbolized as $P_{i,j}^{c_k}$, where $i$, $j$ denotes the position of $P_{i,j}^{c_k}$. The details of the generator's parameter estimation strategy are presented as follows.

First, to estimate the parameter of the proposed generator, some pixels with high spectral similarity were explored $P_{i,j}^{c_k}$. In our previous study [15], the adaptive region is composed of a group of pixels characterized by spectral similarity and spatial continuity. Therefore, exploring potential samples around an initial sample in an adaptive region may be more precise than using a regular window or mathematical model. In the present study, the algorithm was adopted to explore the pixels with high spectra similarity around $P_{i,j}^{c_k}$. The details of the pixel selection approach are briefly reviewed in the following: 1) an adaptive region named $R_{i,j}$ was generated around $P_{i,j}^{c_k}$ with the $T_1$ and $T_2$ parameters, where $T_1$ restricted the spectral similarity between $P_{i,j}^{c_k}$ and its neighbors and $T_2$ constrainted the total number of pixels within the adaptive region; 2) calculate the standard deviation of $R_{i,j}$ in the spectral dimension denoted as $Std_{i,j}$, and take each pixel within $R_{i,j}$ as the central pixel ($P_x$) and an adaptive region around $P_x$ will be extended, the corresponding standard deviation of each adaptive region for the pixels within $R_{i,j}$ can be calculated and denoted as $Std_x$; 3) if $Std_x \leq Std_{i,j}$ is true, then $P_x$ will be labeled as the pixels that have the high spectral similarity with $P_{i,j}^{c_k}$; and 4) after all the pixels are within $R_{i,j}$, the pixels that have a high spectral similarity with $P_{i,j}^{c_k}$ can be explored and denoted as $PS_{i,j}^{c_k} = \{x_1, x_1, x_1, \ldots, x_s\}$.

Second, the parameters for the sample generator around $P_{i,j}^{c_k}$ can be estimated with $PS_{i,j}^{c_k} = \{x_1, x_1, x_1, \ldots, x_s\}$, as shown in the following:

$$\mathrm{Ex} = \frac{1}{s}\sum_{i=1}^{i=s} x_i \tag{2}$$

$$\mathrm{En} = \sqrt{\frac{\pi}{2}} \times \frac{1}{s}\sum_{i=1}^{i=s} |x_i - \mathrm{Ex}| \tag{3}$$

$$\mathrm{He} = \sqrt{\left(\frac{1}{s-1} \times \sum_{i=1}^{i=s} |x_i - \mathrm{Ex}|\right)^2 - \mathrm{En}^2} \tag{4}$$

where $s$ is the number of sample sets $PS_{i,j}^{c_k}$, Ex is an expectation of spectral distribution the sample set, which represents the qualitative concept of the sample set, En measured the degree of uncertainty of the sample set, and He is the uncertainty measurement of En, which reflects the degree of aggregation of the sample set. The proposed parameters estimation approach concentrates on exploring the potential samples around a known sample. To achieve the objective, two aspects, namely, the adaptive region spatial constraint and the spectral similarity constraint, are coupled to explore highly homogeneous pixels around the known sample. Therefore, the explored pixels and the known sample have high homogeneity in terms of spectra reflectance, and they have a high probability of belonging to the same class compared with the known sample. Thus, the explored pixels around a known sample can be used to estimate the parameter of the sample generator for the known sample.

### C. Proposed Sample Generator

Compared with medium–low remote sensing images, HRSI usually demonstrates a higher spectral variance for

**Algorithm 1** Algorithm of the Proposed Sample Generator

**Input:** $PS = \{PS^{c_1}, PS^{c_2}, PS^{c_3}, \ldots, PS^{c_n}\}$; where $c_1, c_2, c_3, \ldots, c_n$ denote the class number, and $n$ is the total classes within an image scene. $PS^{c_k}$ is a sample subset augmented by adaptive region for the $S_{c_k}$, $k \in \{1, 2, 3, \ldots, n\}$.

**Output:** $S' = \{S'_{c_1}, S'_{c_2}, S'_{c_3}, \ldots, S'_{c_n}\}$; where $S'$ is the sample set with the newly generated samples for all the classes.

**For k=1; k++; k≤n do**

**Step-1:** Select one sample subset from $S$, and denote it as $S_{c_k}$.

**Step-2:** Calculate $\mathrm{Ex}(S_{c_k})$, $\mathrm{En}(S_{c_k})$, $\mathrm{He}(S_{c_k})$, which are the *Expectation*, *Entropy*, and *Hyper − entropy* for the sample subset $S_{c_k}$.

**Step-3:** A new sample for the class-k can be generated based on the normal cloud model: $SG(x'^{i}_{c_k}, \mu'^{i}_{c_k}, T)$, where $x'^{i}_{c_k}$ is the new value of a generated sample, and $\mu'^{i}_{c_k}$ is the certainty degree of $x'^{i}_{c_k}$ belongs to the class-k, $T$ is the total quantity of the expected new samples.

**Step-4:** Repeat Step-3 until the number of generated samples reaches $T$.

**Step-5:** Put the generated new samples to a container: $S'_{c_k}$.

**End for**

intraclass samples [45]. Therefore, differentiated samples must be adopted to provide a sufficient representation of a class. In this article, considering the randomness and fuzziness of the pixels within an intraclass in terms of spectral reflectance, a sample generator was proposed based on the normal cloud model, which is a useful tool for describing the randomness and fuzziness of a phenomenon. The normal cloud model has an advantage in implementing the uncertain transformation between a qualitative concept and its quantitative instantiations [46]. In this article, the normal cloud model was first adopted to synthesize virtual samples in our proposed sample generator. The details of the proposed sample generator are given in Algorithm 1.

In the proposed Algorithm 1, $x'^{i}_{c_k}$ is a new sample. When we observe $x'^{i}_{c_k}$ from the viewpoint of the normal cloud model [46], it is a norm random number, and it is obtained with $x'^{i}_{c_k} \in \mathrm{NORM}(\mathrm{Ex}, (\mathrm{En}'_i)^2)$. $\mathrm{En}'_i$ is a normally distributed random variable with expectation En and variance $\mathrm{He}^2$, $\mathrm{En}'^{2}_{i} \in \mathrm{NORM}(\mathrm{En}, (\mathrm{He})^2)$. Here, NORM$(\cdot)$ denotes the Gaussian normal distribution, and it can be written in a universal form: $\mathrm{NORM}(\mu, \sigma) = (1/(2\pi\sigma)^{1/2})e^{-((x-\mu)^2/2\sigma^2)}$, where $\mu$ is the expectation Ex and $\sigma$ is a standard deviation. Then, the random $x$ is subject to a normal distribution with the constraints of $\mu$ and $\sigma$. The normal cloud model plays an important role in the proposed sample generator. Due to the parameters of the normal cloud model being estimated from the initial samples, the characteristics of the samples of a class can be conveyed to constrain the new sample's generation of the class. The randomness and fuzziness of the pixels composed of a class can be modeled by the normal cloud model in the proposed sample generator.

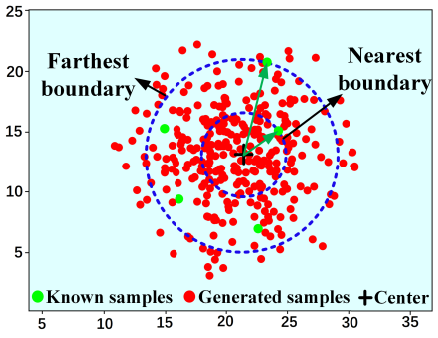



Fig. 2. Schematic of the proposed NFSC to select samples with high reliability.

### D. Refinement of Generated Samples

When a sample generator for creating new samples around a class $S_{c_k}$, some noise samples may exist. Therefore, a simple yet helpful strategy sample refining approach named the near-to-far space constraint strategy (NFSC) is promoted to select the generated samples with high reliability. For clarity, the original initial sample set and the corresponding generated sample set are denoted as $S_{c_k} = \{c_1, c_2, c_3, \ldots, c_k\}$ and $S'_{c_k} = \{c'_1, c'_2, c'_3, \ldots, c'_m\}$, respectively. The details of the proposed NFSC are given as follows.

First, the $K$-mean clustering algorithm is applied to the generated samples $S'_{c_k} = \{c'_1, c'_2, c'_3, \ldots, c'_m\}$, and $S_{c_k} = \{c_1, c_2, c_3, \ldots, c_k\}$ to measure the sample's distribution in terms of spectral feature, and the clustering center is obtained and denoted as $S_{cen}$. Then, as shown in Fig. 2, the first boundary is outlined by the nearest original sample to the clustering center $S_{cen}$, and the second boundary is outlined by the farthest original sample to the clustering center $S_{cen}$. Here, the distance between an originally known sample and the clustering center $S_{cen}$ is based on Euclidean distance and spectral features. Based on the proposed NFSC, some samples within the near-to-far space are anchored for further refining. Other samples are assumed as noisy labels and excluded for further refining.

Second, to cover the spectral variation within an intraclass, the samples within the near-to-far space will be divided into different levels with histograms. After samples within the near-to-far space were counted into a histogram, some new samples were selected randomly from each bin, $Sb_i = \lfloor (Nb_i/NH) \times \mathrm{count} \rfloor$, where $Sb_i$ is the selected new sample's quantity from the $i$th bin. $Nb_i$, $NH$, and count are the sample's quantity of the $i$th bin, the total samples referred to in the histogram, and the aiming quantity of the selection sample from the $i$th bin, respectively. The refining strategy shows that new samples can be selected from each bin, and so the refined samples cover the spectral variance of an intraclass. The refined sample is appended to the last iteration sample set $S^t$ and constructed $S^{t+1}$. Just like iteration termination (1), the proposed method will be terminated until all classes satisfy the iteration termination condition.

## III. Experiment

Three experiments are designed based on six real HRSIs in this section. The first experiment aims to verify the feasibility


Fig. 3. Three-band false-color image of each dataset. (a) Data-1. (b) Data-2. (c) Data-3. (d) Data-4. (e) Data-5. (f) Data-6.

TABLE I
DETAILED DESCRIPTION FOR EACH DATASET

| Dataset | Location | Platform and Sensor | Size (pixel) | Resolution (m/pixel) |
|---|---|---|---|---|
| Data-1 | Califoria, USA | Aerial places/RMKA30 | $560 \times 360 \times 3$ | 0.32 |
| Data-2 | Pavia Center, Italy | Aerial planes / ROSIS sensor | $715 \times 1096 \times 120$ | 1.30 |
| Data-3 | Jiangxi, China | UAV/Canon 5D-Mark-II | $1400 \times 1000 \times 3$ | 0.10 |
| Data-4 | Houston University, USA | Airborne/Spectrographic Imager | $1905 \times 349 \times 144$ | 2.50 |
| Data-5 | Hubei, China | Gaofen-2 Satellite Sensor | $1550 \times 600 \times 3$ | 0.80 |
| Data-6 | Trento, Italy | AISA Eagle sensor | $600 \times 166 \times 63$ | 1.00 |

and advantages of the proposed approach by comparing it with some state-of-the-art methods. The second experiment is designed to investigate the robustness of the proposed approach by applying it to different classifiers. The third experiment concentrated on discussing the relationship between the parameter setting and classification accuracies. The details of each experiment are presented as follows.

### A. Datasets Description

Six real HRRS images were adopted in experiments. A detailed description of the dataset is presented in Table I and Fig. 3. When the dataset is a hyperspectral remote sensing image, a classical approach named PCA is applied for dimensionality reduction. The experimental datasets covered different scenes, including city, urban, and downtown areas. Moreover, many targets have similar spectral reflectance, such as grass and trees and roads and shadows. Therefore, classifying these images into semantic maps with very limited samples is a challenging task.

### B. Experimental Settings and Evaluation Metrics

In the experiments, ten methods, comprising two nondeep learning methods (Richards's method [20] and Tu's method [12]) and eight deep learning-based methods (Li's method [37], MDL-Net [40], S3-Net [35], Gia-CFSL [36], DM-MRN [34], PRCL-FSL [38], ADGAN [47], and 2D-HyperGAMO [48]), are used for comparisons. The details of each method are given as follows.

*1) Nondeep Learning Methods:*

1) *Richards's Method [20]:* This method achieves sample augmentation based on spatial neighborhood information. The training set is expanded by city block distance and spectral angle similarity measures.
2) *Tu's Method [12]:* This method aims at improving the sample's quality by detecting and removing noise labels. In this article, four representative distance metrics are evaluated with a density-peak clustering algorithm to detect noisy labels.

*2) Deep Learning Methods:*

1) *Li's Method [37]:* In this article, a probability model based on a discriminative random field is proposed to learn sample distribution. Then, a loopy belief propagation is adopted for sample augmentation.
2) *MDL-Net [40]:* This network is developed to identify unknown classes. By using extreme value theory to reconstruct the loss, it is possible to maximize the distance between classes in order to achieve classification, which is especially effective for small samples.
3) *S3-Net [35]:* This network has two branches to augment samples by feeding sample pairs into each branch and thus enhancing the model separability, where negative samples are randomly selected to avoid redundancy.
4) *Gia-CFSL [36]:* The network utilizes graph information to represent intra and interclass relationships and enhances the separability through dual modules. In experiments, an excellent classification performance has been observed for few samples.
5) *DM-MRN [34]:* This network concentrates on the underutilization of scarce labeled samples. A sample recombination strategy is designed to increase the effectiveness and robustness of the classification model.
6) *RPCL-FSL [38]:* This network solves prototype instability and domain shift between training and testing datasets. A fusion training strategy is designed to reduce the feature differences between training samples and testing samples to improve classification performance.
7) *ADGAN [47]:* This network includes an output in the discriminator, which has been proven effective in handling minority class samples. In addition, ADGAN generates masks with adaptive shapes, which can improve classification performance.
8) *3D-HyperGAMO [48]:* This model uses the generative adversarial minority oversampling technique, which automatically generates high-quality samples for minority classes at the training stage using the existing samples of that class.

The parameter settings of the classifiers are as follows: SVM [49]: trained with a radial basis function kernel, C set to 10, and along with fivefold cross-validation. RF [50]: configured with ten decision trees and a maximum depth of 10. KNN [51]: configured with $k = 5$, using Euclidean distance. FCN [52]: trained for 200 epochs with a learning rate of 0.001. CNN [53]: trained for 200 epochs with a patch size of 5, a learning rate of

TABLE II
BRIEF SUMMARY OF THE EVALUATION INDICATORS

| | |
|---|---|
| Overall Accuracy (OA) [56] | $OA = \frac{TP+TN}{TP+TN+FP+FN}$ |
| Average Accuracy (AA) [57] | AA = $(\frac{TP}{TP+FP} + \frac{TN}{TN+FN})/T$ |
| Kappa Coefficient (Ka) [58] | $P_0 = \frac{TP+TN}{TP+TN+FP+FN}$, $P_e = \frac{(TP+FP)\times(TP+FN)+(FN+TN)\times(FP+TN)}{(TP+TN+FP+FN)^2}$, $Ka = \frac{P_0-P_e}{1-P_e}$ |
| F1-score (F1-score) [58] | $F1-score = \frac{2TP}{2TP+FN+FP}$ |
| Standard Deviation User's Accuracy (SDUA) [11] | $SDUA = \sqrt{\frac{1}{T}\sum_{i=1}^{i=T}\left(UA_i - \overline{UA}\right)^2}$ |

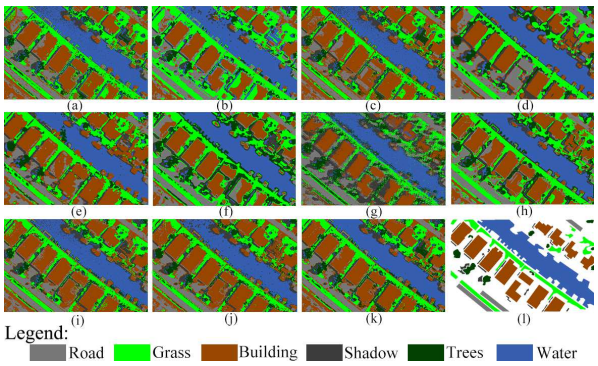



Fig. 4. Classification maps for the Data-1. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], (k) NSAA, and (l) ground truth.

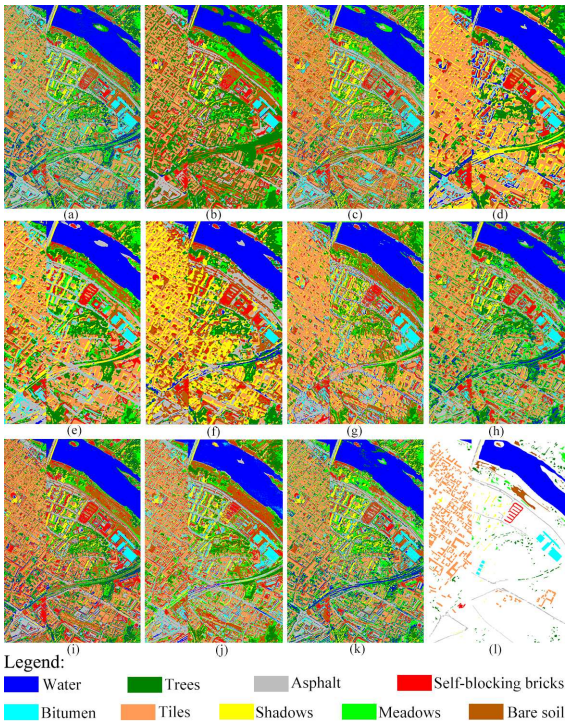



Fig. 5. Classification maps for the Data-2. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], (k) NSAA, and (l) ground truth.

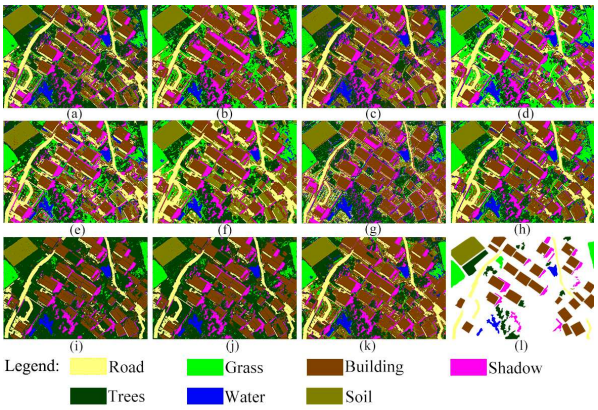



Fig. 6. Classification maps for Data-3. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], (k) NSAA, and (l) ground truth.

0.001. Res-Net [54]: using ResNet-50 architecture, trained for 200 epochs with a learning rate of 0.001. Transformer [55]: using Vit architecture, trained for 200 epochs with a learning rate of 0.001, embedding dimension is 768, and 12 attention heads.

To guarantee comparative fairness, five pixels with labels for each class were selected randomly from the ground truth as the initial sample set. In addition, for the deep learning methods, the learning rate is fixed at $e^{-5}$. The parameters of our proposed approach are as follows: $T_1 = 5$ and $T_2 = 100$ are set for Data-2, Data-4, and Data-6. $T_1 = 5$ and $T_2 = 400$ are set for Data-1, Data-3, and Data-5. In addition, all the experiments were implemented with Python as the backend, which is powered by a workstation with an Intel[1] Core[2] i7-7700 CPU, and NVIDIA GeForce RTX 1080 Ti GPU.

Four widely used indicators, including overall accuracy (OA) [56], average accuracy (AA) [57], kappa coefficient (Ka) [58], and $F1$-score [58], are adopted for quantitative comparisons. In addition, the standard deviation of the user accuracy (SDUA) [11] is adopted to evaluate the accuracy balanced ability for each class. The detailed descriptions of the evaluation metrics are summarized in Table II.

### C. Experimental Results

The first experiments aimed at verifying the performance of the proposed approach by comparing with the cognate methods: as shown in Tables III–VIII, the results based on six real HRSIs and comparing with eight state-of-the-art methods indicated that our proposed approach could achieve at least three best accuracies within the five widely used evaluation metrics. Moreover, compared with the best performance of the existing cognate methods, our approach achieved an improvement of about 0.14% to 1.26% in terms of OA for the six datasets. Regarding the SDUA metric, which measures the balancing ability of each approach for the user's accuracy, our proposed approach achieved the best SDUA when applying it to Data-2, Data-3, Data-4, and Data-6. Particularly, the proposed approach obtained the best Ka in comparison with

[1]Registered trademark.
[2]Trademarked.

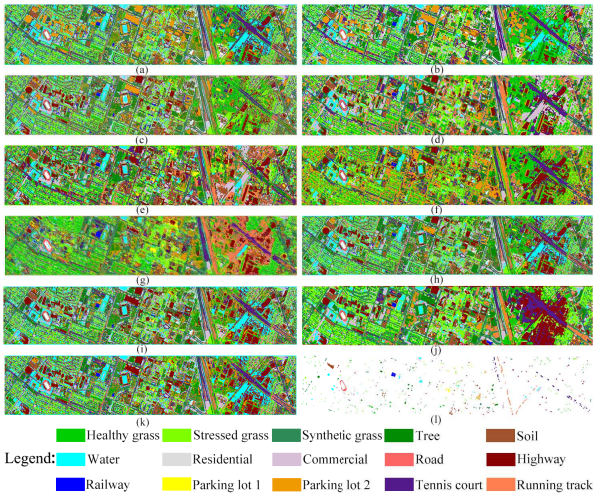



Fig. 7. Classification maps for Data-4. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], (k) NSAA, and (l) ground truth.

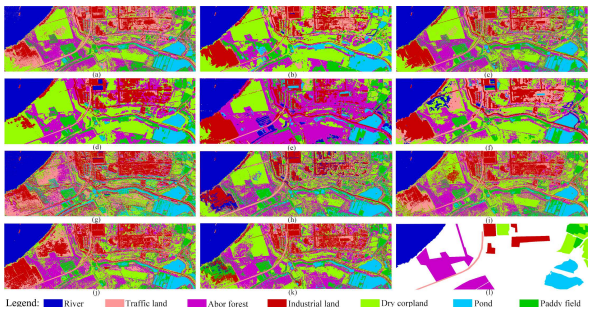



Fig. 8. Classification maps for Data-5. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], (k) NSAA, and (l) ground truth.

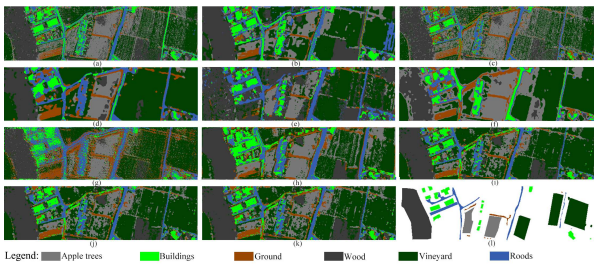



Fig. 9. Classification maps for Data-6. (a) Richard's method [20], (b) Li's method [37], (c) Tu's method [12], (d) MDL-Net [40], (e) S3-Net [35], (f) Gia-CFSL [36], (g) DM-MRN [34], (h) RPCL-FSL [38], (i) ADGAN [47], (j) 3D-HyperGAMO [48], and (k) our proposed approach, and (l) ground truth.

all methods for each dataset. The quantitative comparison indicated that the proposed approach could generate effective samples to improve classification performance with very limited initial samples. In addition, the time-consuming of the proposed NSSA and comparison methods is evaluated for each dataset. The comparison indicated that the proposed NSSA has better time complexity than Li's method [37], Gia-CFSL [36], DM-MRN [34], PRCL-FSL [38], ADGAN [47], and 2D-HyperGAMO [48]. Compared with other methods, the NSSA needed more time to achieve the classification

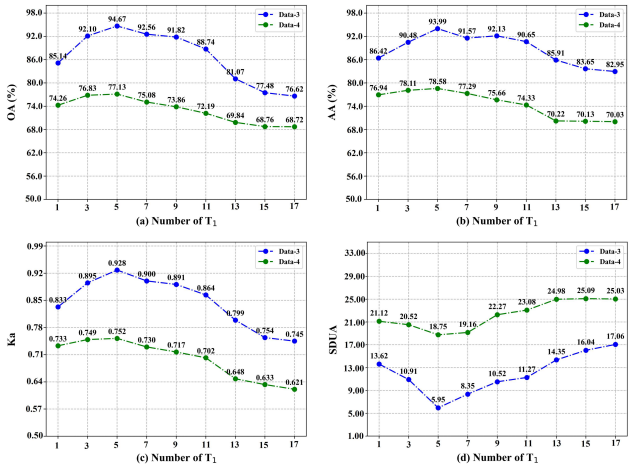



Fig. 10. Relationship between $T_1$ and classification accuracy when $T_2$ is fixed at 400 and 100 for Data-3 and Data-4. (a) Relationship between OA and $T_1$ of the proposed approach. (b) Relationship between AA and $T_1$ of the proposed approach. (c) Relationship between Ka and $T_1$ of the proposed approach. (d) Relationship between SDUA and $T_1$ of the proposed approach.

TABLE III

CLASSIFICATION RESULTS FOR DATA-1 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 85.39 | 72.40 | 0.802 | 75.71 | 22.91 | **14.3** |
| Li' method [37] | 86.06 | 75.83 | 0.810 | 79.13 | 22.59 | 133.0 |
| Tu' method [12] | 83.56 | 70.81 | 0.780 | 74.20 | 24.94 | 29.6 |
| MDL-Net [40] | 88.28 | 77.55 | 0.841 | 80.42 | 22.49 | 66.2 |
| S3-Net [35] | 88.12 | 76.23 | 0.837 | 79.58 | **20.49** | 47.4 |
| Gia-CFSL [36] | 86.40 | 73.86 | 0.817 | 77.98 | 21.85 | 103.3 |
| DM-MRN [34] | 74.26 | 62.55 | 0.662 | 62.59 | 31.29 | 197.4 |
| RPCL-FSL [38] | 89.80 | 78.90 | 0.861 | 82.55 | 22.14 | 97.1 |
| ADGAN [47] | 88.29 | 77.03 | 0.841 | 80.26 | 23.83 | 152.6 |
| 3D-HyperGAMO [48] | 90.64 | 80.97 | 0.873 | 83.87 | 23.33 | 125.4 |
| Proposed NSAA (#13) | **91.06** | **79.60** | **0.878** | **83.10** | 21.74 | 93.7 |

TABLE IV

CLASSIFICATION RESULTS FOR DATA-2 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 87.95 | 73.58 | 0.832 | 75.99 | 19.97 | **21.4** |
| Li' method [37] | 88.03 | **76.50** | 0.836 | 78.15 | 24.79 | 193.2 |
| Tu' method [12] | 82.89 | 68.42 | 0.767 | 70.62 | 25.87 | 43.7 |
| MDL-Net [40] | **90.73** | 75.81 | 0.857 | 73.52 | 18.71 | 104.7 |
| S3-Net [35] | 89.54 | 75.28 | 0.855 | **79.13** | 20.86 | 74.3 |
| Gia-CFSL [36] | 84.93 | 69.28 | 0.792 | 69.00 | 27.06 | 193.6 |
| DM-MRN [34] | 85.75 | 65.34 | 0.803 | 66.37 | 23.55 | 223.9 |
| RPCL-FSL [38] | 87.22 | 73.22 | 0.822 | 75.55 | 21.99 | 179.4 |
| ADGAN [47] | 86.74 | 73.73 | 0.816 | 74.91 | 24.83 | 207.5 |
| 3D-HyperGAMO [48] | 88.87 | 72.56 | 0.846 | 74.86 | 24.33 | 186.3 |
| Proposed NSAA (#13) | 89.78 | 75.08 | **0.857** | 76.57 | **18.10** | 93.5 |

map for the same dataset, because the NSSA was an iterative algorithm that needed more time to augment and refine each class of samples. The visual performance comparisons further confirmed the conclusion based on the quantitative comparison. As shown in Figs. 4–10, compared with the Richard method [20], the proposed NSAA demonstrates a robustness in processing HRSIs, especially under the condition of limited samples. While the Tu method [12] was excellent in noise label detection, it still has improvement space in sample augmentation. Few shot methods, such as MDL-Net [40], S3-Net [35], and RPCL-FSL [38], have shown excellent performance on some datasets, but they are not as effective as NSAA in addressing class imbalance.

TABLE V

CLASSIFICATION RESULTS FOR DATA-3 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 85.48 | 86.29 | 0.814 | 88.32 | 14.08 | **39.7** |
| Li' method [37] | 94.53 | **95.02** | 0.919 | 93.26 | 8.27 | 223.3 |
| Tu' method [12] | 87.52 | 85.96 | 0.835 | 87.01 | 12.87 | 67.4 |
| MDL-Net [40] | 88.04 | 84.19 | 0.844 | 87.28 | 12.78 | 145.1 |
| S3-Net [35] | 87.96 | 85.14 | 0.843 | 88.01 | 12.4 | 103.9 |
| Gia-CFSL [36] | 87.25 | 88.65 | 0.835 | 89.88 | 14.98 | 186.5 |
| DM-MRN [34] | 88.31 | 85.92 | 0.843 | 86.38 | 11.13 | 261.3 |
| RPCL-FSL [38] | 93.67 | 90.72 | 0.915 | 92.57 | 8.34 | 221.8 |
| ADGAN [47] | 93.18 | 94.63 | 0.909 | 94.54 | 11.94 | 239.7 |
| 3D-HyperGAMO [48] | 93.87 | 94.40 | 0.918 | 94.48 | 10.60 | 219.4 |
| Proposed NSAA (#12) | **94.67** | 93.99 | **0.928** | **94.59** | **5.95** | 147.3 |

TABLE VI

CLASSIFICATION RESULTS FOR DATA-4 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 73.57 | 76.24 | 0.715 | 74.84 | 19.20 | **23.6** |
| Li' method [37] | 75.70 | 78.04 | 0.738 | 75.47 | 20.78 | 187.3 |
| Tu' method [12] | 62.44 | 65.53 | 0.594 | 63.95 | 21.92 | 43.8 |
| MDL-Net [40] | 74.05 | 77.10 | 0.719 | 76.93 | 19.57 | 78.4 |
| S3-Net [35] | 74.55 | 77.26 | 0.725 | 77.44 | 19.63 | 66.5 |
| Gia-CFSL [36] | 70.02 | 75.16 | 0.677 | 72.74 | 19.46 | 151.9 |
| DM-MRN [34] | 67.13 | 67.19 | 0.647 | 66.93 | 21.74 | 217.7 |
| RPCL-FSL [38] | 76.45 | 77.76 | 0.745 | 77.11 | 19.19 | 124.3 |
| ADGAN [47] | 76.45 | 77.76 | 0.745 | 77.11 | 19.19 | 193.6 |
| 3D-HyperGAMO [48] | 76.83 | 78.11 | 0.749 | 77.38 | 19.03 | 173.6 |
| Proposed NSAA (#12) | **77.13** | **78.58** | **0.753** | **78.03** | **18.75** | 102.8 |

TABLE VII

CLASSIFICATION RESULTS FOR DATA-5 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 77.36 | 70.31 | 0.722 | 69.74 | 26.75 | **20.3** |
| Li' method [37] | 73.98 | 63.04 | 0.682 | 62.25 | 30.19 | 196.7 |
| Tu' method [12] | 75.32 | 66.90 | 0.695 | 65.70 | 27.95 | 73.2 |
| MDL-Net [40] | 78.64 | 72.47 | 0.736 | 69.21 | 21.78 | 88.7 |
| S3-Net [35] | 77.87 | 70.84 | 0.722 | 67.92 | 25.33 | 66.2 |
| Gia-CFSL [36] | 72.77 | 68.23 | 0.666 | 65.94 | 25.26 | 191.9 |
| DM-MRN [34] | 66.18 | 57.65 | 0.594 | 56.93 | 27.02 | 222.7 |
| RPCL-FSL [38] | 75.21 | 69.50 | 0.695 | 69.74 | **23.05** | 188.0 |
| ADGAN [47] | 78.23 | 69.52 | 0.731 | 70.43 | 25.23 | 202.6 |
| 3D-HyperGAMO [48] | 78.98 | 70.07 | 0.733 | 70.65 | 25.02 | 183.8 |
| Proposed NSAA (#15) | **79.10** | **70.72** | **0.741** | **70.91** | 24.78 | 119.7 |

TABLE VIII

CLASSIFICATION RESULTS FOR DATA-6 (# DENOTES THE TOTAL ITERATIONS OF THE PROPOSED APPROACH)

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA | Time (min) |
|---|---|---|---|---|---|---|
| Richard' method [20] | 78.47 | 72.23 | 0.719 | 73.12 | **14.49** | **15.4** |
| Li' method [37] | 82.94 | 74.10 | 0.769 | 71.76 | 19.80 | 121.8 |
| Tu' method [12] | 73.79 | 67.29 | 0.662 | 65.58 | 24.27 | 37.9 |
| MDL-Net [40] | 83.97 | 74.03 | 0.784 | 72.69 | 22.58 | 71.9 |
| S3-Net [35] | 81.99 | 75.63 | 0.762 | **77.00** | 18.01 | 56.3 |
| Gia-CFSL [36] | 80.13 | 66.66 | 0.734 | 66.34 | 22.32 | 137.0 |
| DM-MRN [34] | 84.52 | 72.19 | 0.786 | 72.36 | 25.55 | 163.5 |
| RPCL-FSL [38] | 82.91 | 75.56 | 0.776 | 73.89 | 18.84 | 111.3 |
| ADGAN [47] | 83.41 | 75.97 | 0.774 | 72.16 | 17.61 | 147.3 |
| 3D-HyperGAMO [48] | 84.00 | 75.74 | 0.782 | 73.16 | 17.75 | 139.8 |
| Proposed NSAA (#16) | **84.52** | **75.78** | **0.790** | 74.11 | 17.64 | 88.3 |

The second experiment aims at verifying the robustness and adaptive ability of the proposed approach by applying it to different classifiers, including classic classifiers (random forest (RF) [50], K-nearest neighbors (KNN) [51], and deep learning classifiers [fully connected network (FCN) [52], CNN [53], Res-Net [54], and transformer [55]]. The parameter details of the classifiers can be found in Section IV. Based on the above clarifiers, classification accuracies between a classifier using the initial samples and using our proposed approach are summarized in Table IX. The comparisons based on the six datasets and four widely used classifiers effectively demonstrated the advances and robustness of these classifiers.

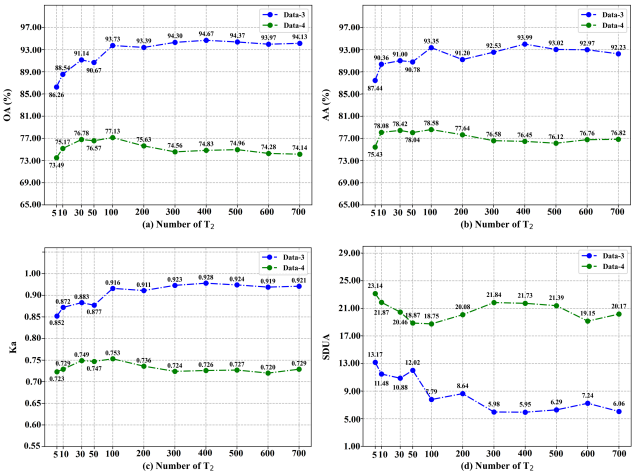



Fig. 11. Relationship between $T_2$ and classification accuracy when $T_1$ is fixed at 5 for Data-3 and Data-4. (a) Relationship between OA and $T_1$ of the proposed approach. (b) Relationship between AA and $T_1$ of the proposed approach. (c) Relationship between Ka and $T_1$ of the proposed approach. (d) Relationship between SDUA and $T_1$ of the proposed approach.

Although experiments with some real datasets verified the feasibility and advantages of the NSAA, it may perform less effectively for classes with extremely high intraclass variance, because the sample generator is based on a normal cloud model, and the modeling of statistical features may capture the complex variation insufficiently. In the future, we will explore more complex generative models to solve this limitation.

### D. Discussion

For the ubiquity of the proposed NSSA in practical applications, here, we discussed the impact of parameter tuning on the classification performance of the NSAA, the relationship between the iteration and classification accuracy of the NSAA, as well as the classification ability of NSAA in few-sample scenarios.

First, we investigated the relationship between classification accuracies and the parameters ($T_1$ and $T_2$) of the refereed adaptive region extension algorithm in [15]. Fig. 10 indicates that different datasets have different fluctuations with the increment of $T_1$ when $T_2$ is fixed. Because $T_1$ denotes the spectral similarity between a pixel and its neighboring pixels, when $T_1$ is too small, it is insufficient to extend an adaptive region and explore enough contextual information. With the increment of $T_1$, the size of the adaptive region was increased correspondingly, and more contextual information will be utilized. In addition, $T_2$ is a predefined number of the total pixels within an extended region. Therefore, the shape and size of an adaptive region around a pixel depend on the setting of $T_1$ and $T_2$. Based on the setting of $T_1$ and $T_2$, different evaluation metrics performed different curves. For example, OA is increased from 85.14% to 94.67% with $T_1 = 1$ to $T_1 = 5$ for Data-3, and OA is maintained at about 76.00% when $T_1$ is increased from 1 to 5 for Data-4. On the contrary, an observation of the relationship between $T_2$ and classification

TABLE IX

CLASSIFICATION PERFORMANCE COMPARISONS BETWEEN USING THE INITIAL SAMPLES AND USING THE AUGMENTED SAMPLES OBTAINED BY THE PROPOSED NSAA FOR SOME CLASSIFIERS. ($Ka \in [-1, +1]$, OTHERS MEASURED IN PERCENTAGE)

| Dataset | Classifier | OA (%) | | AA (%) | | Ka | | F1-score (%) | | SDUA | |
|---|---|---|---|---|---|---|---|---|---|---|---|
| | | Initial | NSAA | Initial | NSAA | Initial | NSAA | Initial | NSAA | Initial | NSAA |
| Data-1 | RF [50] | 83.33 | **87.03** | 70.59 | **75.20** | 0.778 | **0.825** | 74.17 | **78.55** | 24.87 | **21.79** |
| | KNN [51] | 81.29 | **88.35** | 70.00 | **76.06** | 0.752 | **0.842** | 72.59 | **79.58** | 25.60 | **20.26** |
| | FCN [52] | 83.35 | **88.36** | 72.20 | **76.50** | 0.779 | **0.841** | 74.39 | **79.19** | 30.34 | **22.75** |
| | CNN [53] | 78.16 | **90.46** | 68.14 | **82.29** | 0.706 | **0.869** | 71.36 | **85.80** | 22.86 | **18.43** |
| | ResNet [54] | 84.18 | **91.36** | 70.68 | **81.30** | 0.732 | **0.873** | 73.27 | **86.94** | 21.34 | **17.32** |
| | Transformer [55] | 85.42 | **93.73** | 72.17 | **83.82** | 0.769 | **0.882** | 74.56 | **87.51** | 21.23 | **15.61** |
| Data-2 | RF [50] | 79.08 | **90.07** | 60.65 | **76.93** | 0.713 | **0.863** | 62.23 | **80.47** | 27.02 | **19.36** |
| | KNN [51] | 79.91 | **90.74** | 65.93 | **78.13** | 0.722 | **0.872** | 60.17 | **81.43** | 25.05 | **18.28** |
| | FCN [52] | 81.99 | **91.73** | 60.67 | **79.74** | 0.754 | **0.885** | 62.68 | **82.22** | 27.97 | **17.05** |
| | CNN [53] | 82.13 | **93.33** | 63.49 | **84.19** | 0.748 | **0.907** | 49.62 | **85.71** | 25.24 | **15.17** |
| | ResNet [54] | 82.34 | **93.67** | 64.32 | **84.33** | 0.749 | **0.894** | 61.36 | **85.80** | 26.32 | **15.03** |
| | Transformer [55] | 82.79 | **94.53** | 65.17 | **85.71** | 0.752 | **0.910** | 62.03 | **86.67** | 25.87 | **14.36** |
| Data-3 | RF [50] | 85.85 | **94.21** | 83.62 | **91.48** | 0.816 | **0.922** | 85.62 | **92.94** | 13.96 | **5.04** |
| | KNN [51] | 80.67 | **94.89** | 81.27 | **91.78** | 0.754 | **0.931** | 82.34 | **93.28** | 18.34 | **5.53** |
| | FCN [52] | 84.29 | **96.80** | 82.55 | **95.46** | 0.798 | **0.957** | 85.35 | **95.93** | 12.69 | **3.36** |
| | CNN [53] | 81.99 | **94.73** | 77.63 | **91.37** | 0.762 | **0.930** | 79.00 | **93.47** | 12.01 | **7.45** |
| | ResNet [54] | 82,03 | **95.24** | 81.67 | **91.35** | 0.773 | **0.953** | 84.27 | **94.60** | 12.24 | **4.36** |
| | Transformer [55] | 82.96 | **94.53** | 82.93 | **92.09** | 0.782 | **0.961** | 84.78 | **96.13** | 12.10 | **2.97** |
| Data-4 | RF [50] | 58.42 | **78.12** | 60.72 | **78.65** | 0.552 | **0.764** | 58.97 | **78.89** | 19.76 | **16.13** |
| | KNN [51] | 55.43 | **77.35** | 57.79 | **77.73** | 0.520 | **0.755** | 54.50 | **77.70** | 19.52 | **17.38** |
| | FCN [52] | 66.00 | **79.40** | 66.49 | **81.46** | 0.634 | **0.778** | 66.50 | **80.58** | 20.83 | **16.63** |
| | CNN [53] | 53.21 | **80.71** | 54.71 | **80.90** | 0.495 | **0.792** | 51.72 | **81.30** | 15.11 | **10.78** |
| | ResNet [54] | 63.27 | **78.36** | 64.32 | **80.96** | 0.601 | **0.785** | 59.21 | **75.04** | 24.98 | **19.36** |
| | Transformer [55] | 66.21 | **81.27** | 65.14 | **82.13** | 0.627 | **0.803** | 60.23 | **76.39** | 24.67 | **18.43** |
| Data-5 | RF [50] | 52.53 | **77.28** | 49.18 | **69.34** | 0.433 | **0.721** | 45.24 | **69.75** | 31.88 | **25.21** |
| | KNN [51] | 57.35 | **77.64** | 52.71 | **69.77** | 0.470 | **0.724** | 42.09 | **70.03** | 27.48 | **24.92** |
| | FCN [52] | 69.49 | **75.31** | 62.61 | **68.68** | 0.621 | **0.699** | 59.93 | **68.22** | 27.42 | **24.65** |
| | CNN [53] | 48.20 | **80.39** | 41.39 | **80.39** | 0.380 | **0.759** | 36.13 | **75.91** | 23.45 | **19.72** |
| | ResNet [54] | 67.36 | **78.64** | 58.34 | **77.21** | 0.584 | **0.743** | 57.23 | **72.36** | 24.36 | **21.72** |
| | Transformer [55] | 70.16 | **80.93** | 61.70 | **80.45** | 0.609 | **0.763** | 59.61 | **75.83** | 23.91 | **18.69** |
| Data-6 | RF [50] | 70.17 | **80.90** | 64.00 | **78.26** | 0.604 | **0.749** | 62.35 | **79.91** | 18.69 | **9.65** |
| | KNN [51] | 62.66 | **81.11** | 65.02 | **78.64** | 0.535 | **0.751** | 58.21 | **80.07** | 24.58 | **9.45** |
| | FCN [52] | 71.78 | **84.51** | 62.42 | **73.99** | 0.633 | **0.797** | 62.92 | **75.37** | 21.99 | **19.33** |
| | CNN [53] | 68.99 | **89.78** | 62.41 | **84.57** | 0.610 | **0.865** | 61.25 | **84.76** | 30.81 | **18.86** |
| | ResNet [54] | 69.32 | **88.67** | 64.53 | **82.96** | 0.624 | **0.857** | 62.36 | **82.93** | 21.67 | **18.54** |
| | Transformer [55] | 70.21 | **91.53** | 65.21 | **85.31** | 0.635 | **0.864** | 63.17 | **85.30** | 21.03 | **17.64** |

accuracy when $T_1$ is fixed, as shown in Fig. 11, different datasets have different fluctuations with the increment of $T_1$. As $T_2$ increases, all indicators have increased. Due to spatial uncertainty within the image scene and the assumption of spatial homogeneity in geographic areas, the limitation of $T_1$, as $T_2$ continues to increase, different indicators appear to be horizontally floating.

Second, the total iteration of the proposed NSSA was discussed. The NSSA is an iterative framework, and the iteration is terminated until all the classes satisfy a predefined condition. As shown in Fig. 12, the accuracies in terms of OA, AA, and Ka were first increased with the increment of the iteration times from 1 to 5, and then, the accuracies changed slightly but tended to be at a horizontal level after the increment of iterations was larger than 6. By contrast, SDUA was decreased with the iteration from 1 to 5, and then, it also tends to be a horizontal level when the iteration is larger than 6. The observation demonstrated that 1) the NSSA terminated at different iterations for different datasets; 2) the augmented samples using the NSSA are effective for improving classification accuracies with HRSIs; and 3) the predefined iteration termination conditions are effective in dynamically adjusting the termination of the iteration of the NSSA and balancing the users' accuracy for different datasets.

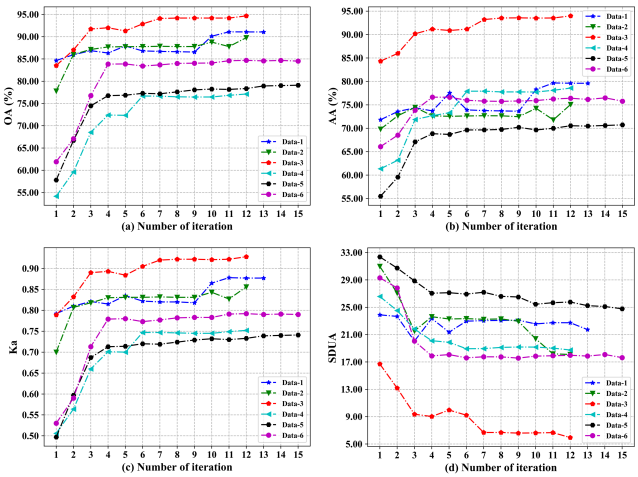



Fig. 12. Relationship between classification accuracies and iterations of the proposed approach coupled with SVM. (a) Relationship between OA and $T_1$ of the proposed approach. (b) Relationship between AA and $T_1$ of the proposed approach. (c) Relationship between Ka and $T_1$ of the proposed approach. (d) Relationship between SDUA and $T_1$ of the proposed approach.

Third, investigating the relationship between the initial samples quantity and the final classification accuracies based on NSSA is valuable. As shown in Fig. 13, the accuracies in terms of OA, AA, and Ka were increased gradually with the

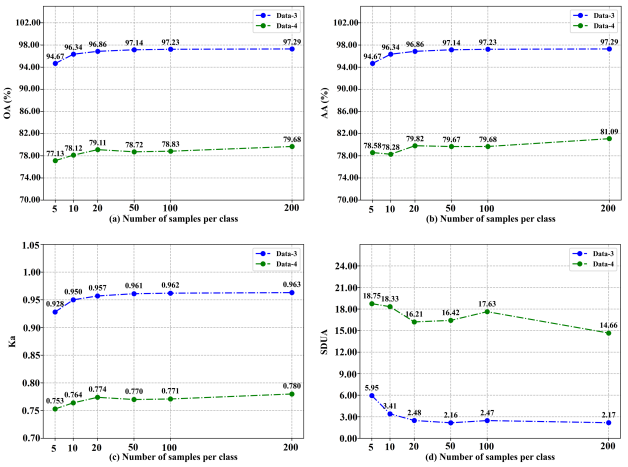



Fig. 13. Relationship between classification accuracies and number of samples of the proposed approach coupled with SVM for Data-3 and Data-4. (a) Relationship between OA and $T_1$ of the proposed approach. (b) Relationship between AA and $T_1$ of the proposed approach. (c) Relationship between Ka and $T_1$ of the proposed approach. (d) Relationship between SDUA and $T_1$ of the proposed approach.

TABLE X

COMPARE THE CLASSIFICATION PERFORMANCE OF DATA-1 USING DIFFERENT BALANCING STRATEGIES

| Method | OA (%) | AA (%) | Ka | F1-score (%) | SDUA |
|---|---|---|---|---|---|
| Base | 88.27 | 76.32 | 0.844 | 81.61 | 23.74 |
| Under-sampling [59] | 87.69 | 76.14 | 0.847 | 81.13 | 23.43 |
| SMOTE [17] | 89.76 | 78.83 | 0.860 | 82.17 | 22.59 |
| K-Means-SMOTE [18] | 90.53 | 79.41 | 0.872 | 82.94 | 21.97 |
| Proposed NSAA (#13) | **91.06** | **79.60** | **0.878** | **83.10** | **21.74** |

increment of initial samples from 5 to 200 pixels/class, but the improvement degree is not significant. Regardless of how many initial samples were selected for our proposed NSSA, the initial sample will be augmented iteratively to a balanced status in the classification progress. In addition, when we observe, the SDUA reduces slightly with the increment of the quantity of initial samples from 5 to 200 pixels/class. This is because more initial samples will provide more reliable parameter estimation for the proposed sample generator.

Fourth, we add a class balanced evaluation experiment based on Data-1, which will be trained using the same number of samples generated by our method for each class, but these samples will be balanced through methods, such as undersampling [59], SMOTE [17], and $K$-means-SMOTE [18]. Table X indicates that the NSAA outperforms other balancing strategies and achieves the highest classification performance. In addition, NSSA not only balances the number of samples each class but also enhances sample quality.

## IV. CONCLUSION

In this article, we have proposed a sample augmentation approach for improving classification performance with HRSIs when initial known samples are very small. In the proposed NSSA, a novel sample generator based on the normal cloud model is promoted to generate a large number of samples based on parameter estimation with initial known samples, and then, a novel near-to-far space constraint approach is designed to refine the generated samples and obtain the final reliable sample set. To verify the feasibility and advantages of the NSSA, a sample generator was combined with a sample balancing strategy, and the generated samples were used as training samples for a supervised classifier to improve classification performance when initial known samples are very limited.

The NSSA has some intuitive advances in generating the sample's quantity, and sample quality, such as it only needs 10 pixels/class for Data-1 to Data-6, but obtained OA = 91.06%, 89.78%, 94.67%, 77.13%, 79.10%, and 84.52%, respectively, which are better than that of some state-of-arts methods. In addition, the NSSA has an advantage in balancing the user's accuracies, such as the SDUA of the proposed method is significantly reduced by 5.79–9.82 compared with the initial classification map. Although the comparisons with some real datasets verified the feasibility and advantages of the proposed NSAA, the proposed NSAA was promoted based on a normal distribution assumption. Although the distribution of the pixels within a local area usually follows a normal distribution, the speckle noise of remote sensing images may affect the performance of the proposed NSAA. Moreover, the NSAA referred to two parameters, and the optimal setting of the parameters is time-consuming for a specific dataset. Therefore, in our future research, we will concentrate on promoting an estimation function for describing the distribution of the pixels within a local area and then proposing a robust sample augmentation approach to improve classification performance.

## REFERENCES

[1] G. Cheng, X. Xie, J. Han, L. Guo, and G.-S. Xia, "Remote sensing image scene classification meets deep learning: Challenges, methods, benchmarks, and opportunities," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 3735–3756, 2020.

[2] L. Gao, H.-M. Hu, X. Xue, and H. Hu, "From appearance to inherence: A hyperspectral image dataset and benchmark of material classification for surveillance," *IEEE Trans. Multimedia*, vol. 26, pp. 8569–8580, 2024.

[3] N. Kussul, M. Lavreniuk, S. Skakun, and A. Shelestov, "Deep learning classification of land cover and crop types using remote sensing data," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 5, pp. 778–782, May 2017.

[4] Y. Li, B. Dang, Y. Zhang, and Z. Du, "Water body classification from high-resolution optical remote sensing imagery: Achievements and perspectives," *ISPRS J. Photogramm. Remote Sens.*, vol. 187, pp. 306–327, May 2022.

[5] G. Giacinto, F. Roli, and L. Bruzzone, "Combination of neural and statistical algorithms for supervised classification of remote-sensing images," *Pattern Recognit. Lett.*, vol. 21, no. 5, pp. 385–397, May 2000.

[6] X. Qian, C. Wang, W. Wang, X. Yao, and G. Cheng, "Complete and invariant instance classifier refinement for weakly supervised object detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 62, 2024, Art. no. 5627713.

[7] C. Kyba et al., "High-resolution imagery of Earth at night: New sources, opportunities and challenges," *Remote Sens.*, vol. 7, no. 1, pp. 1–23, Dec. 2014.

[8] P. Liu, T. Xu, H. Chen, S. Zhou, H. Qin, and J. Li, "Spectrum-driven mixed-frequency network for hyperspectral salient object detection," *IEEE Trans. Multimedia*, vol. 26, pp. 5296–5310, 2024.

[9] Z. Lv, P. Zhang, W. Sun, J. A. Benediktsson, and T. Lei, "Novel land-cover classification approach with nonparametric sample augmentation for hyperspectral remote-sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 4407613.

[10] Z. Zheng, Y. Zhong, J. Wang, A. Ma, and L. Zhang, "FarSeg++: Foreground-aware relation network for geospatial object segmentation in high spatial resolution remote sensing imagery," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 45, no. 11, pp. 13715–13729, Nov. 2023.

[11] Z. Lv, G. Li, Z. Jin, J. A. Benediktsson, and G. M. Foody, "Iterative training sample expansion to increase and balance the accuracy of land classification from VHR imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 1, pp. 139–150, Jan. 2021.

[12] B. Tu, X. Zhang, X. Kang, G. Zhang, and S. Li, "Density peak-based noisy label detection for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1573–1584, Mar. 2019.

[13] X. Kang, P. Duan, X. Xiang, S. Li, and J. A. Benediktsson, "Detection and correction of mislabeled training samples for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 10, pp. 5673–5686, Oct. 2018.

[14] A. Mellor, S. Boukir, A. Haywood, and S. Jones, "Exploring issues of training data imbalance and mislabelling on random forest performance for large area land cover classification using the ensemble margin," *ISPRS J. Photogramm. Remote Sens.*, vol. 105, pp. 155–168, Jul. 2015.

[15] Z. Lv, P. Zhang, W. Sun, J. A. Benediktsson, J. Li, and W. Wang, "Novel adaptive region spectral–spatial features for land cover classification with high spatial resolution remotely sensed imagery," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5609412.

[16] M. Bayer, M.-A. Kaufhold, and C. Reuter, "A survey on data augmentation for text classification," *ACM Comput. Surv.*, vol. 55, no. 7, pp. 1–39, 2022.

[17] N. V. Chawla, K. W. Bowyer, L. O. Hall, and W. P. Kegelmeyer, "SMOTE: Synthetic minority over-sampling technique," *J. Artif. Intell. Res.*, vol. 16, pp. 321–357, Jun. 2002.

[18] G. Douzas, F. Bacao, and F. Last, "Improving imbalanced learning through a heuristic oversampling method based on k-means and SMOTE," *Inf. Sci.*, vol. 465, pp. 1–20, Oct. 2018.

[19] Q. Hao, S. Li, and X. Kang, "Multilabel sample augmentation-based hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 6, pp. 4263–4278, Jun. 2020.

[20] J. A. Richards and X. Jia, "Using suitable neighbors to augment the training set in hyperspectral maximum likelihood classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 774–777, Oct. 2008.

[21] G. Li, Q. Gao, J. Han, and X. Gao, "A coarse-to-fine cell division approach for hyperspectral remote sensing image classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 34, no. 6, pp. 4928–4941, Jun. 2024.

[22] C. Shorten and T. M. Khoshgoftaar, "A survey on image data augmentation for deep learning," *J. Big Data*, vol. 6, no. 1, pp. 1–48, Dec. 2019.

[23] A. Dabouei, S. Soleymani, F. Taherkhani, and N. M. Nasrabadi, "SuperMix: Supervising the mixing data augmentation," in *Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2021, pp. 13789–13798.

[24] H. K. Choi, J. Choi, and H. J. Kim, "TokenMixup: Efficient attention-guided token-level data augmentation for transformers," in *Proc. Adv. Neural Inf. Process. Syst.*, vol. 35, 2022, pp. 14224–14235.

[25] C. Shi and C.-M. Pun, "Multiscale superpixel-based hyperspectral image classification using recurrent neural networks with stacked autoencoders," *IEEE Trans. Multimedia*, vol. 22, no. 2, pp. 487–501, Feb. 2020.

[26] X. Yu, X. Wu, C. Luo, and P. Ren, "Deep learning in remote sensing scene classification: A data augmentation enhanced convolutional neural network framework," *GIScience Remote Sens.*, vol. 54, no. 5, pp. 741–758, Sep. 2017.

[27] S. C. Wong, A. Gatt, V. Stamatescu, and M. D. McDonnell, "Understanding data augmentation for classification: When to warp?," in *Proc. Int. Conf. Digit. Image Comput., Techn. Appl. (DICTA)*, Nov. 2016, pp. 1–6.

[28] P. Smits, S. Dellepiane, and R. Schowengerdt, "Quality assessment of image classification algorithms for land-cover mapping: A review and a proposal for a cost-based approach," *Int. J. Remote Sens.*, vol. 20, no. 8, pp. 1461–1486, 1999.

[29] X. Qian, Q. Jian, W. Wang, X. Yao, and G. Cheng, "Incorporating multiscale context and task-consistent focal loss into oriented object detection," *IEEE Trans. Geosci. Remote Sens.*, vol. 63, 2025, Art. no. 5628411.

[30] J. Wu, A. Guo, V. S. Sheng, P. Zhao, and Z. Cui, "An active learning approach for multi-label image classification with sample noise," *Int. J. Pattern Recognit. Artif. Intell.*, vol. 32, no. 3, Mar. 2018, Art. no. 1850005.

[31] J. Jiang, J. Ma, and X. Liu, "Multilayer spectral–spatial graphs for label noisy robust hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 33, no. 2, pp. 839–852, Feb. 2022.

[32] G. Algan and I. Ulusoy, "Image classification with deep learning in the presence of noisy labels: A survey," *Knowl.-Based Syst.*, vol. 215, Mar. 2021, Art. no. 106771.

[33] N. Audebert, B. Le Saux, and S. Lefevre, "Deep learning for classification of hyperspectral data: A comparative review," *IEEE Geosci. Remote Sens. Mag.*, vol. 7, no. 2, pp. 159–173, Jun. 2019.

[34] J. Zeng, Z. Xue, L. Zhang, Q. Lan, and M. Zhang, "Multistage relation network with dual-metric for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5510017.

[35] Z. Xue, Y. Zhou, and P. Du, "S3Net: Spectral–spatial Siamese network for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5531219.

[36] Y. Zhang, W. Li, M. Zhang, S. Wang, R. Tao, and Q. Du, "Graph information aggregation cross-domain few-shot learning for hyperspectral image classification," *IEEE Trans. Neural Netw. Learn. Syst.*, vol. 35, no. 2, pp. 1912–1925, Feb. 2024.

[37] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Spectral–spatial classification of hyperspectral data using loopy belief propagation and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 2, pp. 844–856, Feb. 2013.

[38] Q. Liu et al., "Refined prototypical contrastive learning for few-shot hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5506214.

[39] X. Qian, B. Wu, G. Cheng, X. Yao, W. Wang, and J. Han, "Building a bridge of bounding box regression between oriented and horizontal object detection in remote sensing images," *IEEE Trans. Geosci. Remote Sens.*, vol. 61, 2023, Art. no. 5605209.

[40] S. Liu, Q. Shi, and L. Zhang, "Few-shot hyperspectral image classification with unknown classes using multitask deep learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5085–5102, Jun. 2021.

[41] Q. Jin, M. Yuan, H. Wang, M. Wang, and Z. Song, "Deep active learning models for imbalanced image classification," *Knowl.-Based Syst.*, vol. 257, Dec. 2022, Art. no. 109817.

[42] Y. Ren et al., "Full convolutional neural network based on multi-scale feature fusion for the class imbalance remote sensing image classification," *Remote Sens.*, vol. 12, no. 21, p. 3547, Oct. 2020.

[43] A. Özdemir, K. Polat, and A. Alhudhaif, "Classification of imbalanced hyperspectral images using SMOTE-based deep learning methods," *Expert Syst. Appl.*, vol. 178, Sep. 2021, Art. no. 114986.

[44] Q. Zhu et al., "A spectral–spatial-dependent global learning framework for insufficient and imbalanced hyperspectral image classification," *IEEE Trans. Cybern.*, vol. 52, no. 11, pp. 11709–11723, Nov. 2022.

[45] G. M. Foody, "Impacts of ignorance on the accuracy of image classification and thematic mapping," *Remote Sens. Environ.*, vol. 259, Jun. 2021, Art. no. 112367.

[46] D. Li, C. Liu, and W. Gan, "A new cognitive model: Cloud model," *Int. J. Intell. Syst.*, vol. 24, no. 3, pp. 357–375, Mar. 2009.

[47] J. Wang, F. Gao, J. Dong, and Q. Du, "Adaptive dropblock-enhanced generative adversarial networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 6, pp. 5040–5053, Jun. 2021.

[48] S. K. Roy, J. M. Haut, M. E. Paoletti, S. R. Dubey, and A. Plaza, "Generative adversarial minority oversampling for spectral–spatial hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 60, 2022, Art. no. 5500615.

[49] M. A. Hearst, S. T. Dumais, E. Osuna, J. Platt, and B. Scholkopf, "Support vector machines," *IEEE Intell. Syst. Appl.*, vol. 13, no. 4, pp. 18–28, Jul./Aug. 1998.

[50] T. K. Ho, "Random decision forests," in *Proc. 3rd Int. Conf. Document Anal. Recognit.*, vol. 1, Aug. 1995, pp. 278–282.

[51] L. E. Peterson, "K-nearest neighbor," *Scholarpedia*, vol. 4, no. 2, p. 1883, Jan. 2009.

[52] G. Fu, C. Liu, R. Zhou, T. Sun, and Q. Zhang, "Classification for high resolution remote sensing imagery using a fully convolutional network," *Remote Sens.*, vol. 9, no. 5, p. 498, May 2017.

[53] E. Maggiori, Y. Tarabalka, G. Charpiat, and P. Alliez, "Convolutional neural networks for large-scale remote-sensing image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 2, pp. 645–657, Feb. 2017.

[54] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in *Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR)*, Jun. 2016, pp. 770–778.

[55] A. Dosovitskiy et al., "An image is worth 16 × 16 words: Transformers for image recognition at scale," 2020, *arXiv:2010.11929*.

[56] G. M. Foody, "Status of land cover classification accuracy assessment," *Remote Sens. Environ.*, vol. 80, no. 1, pp. 185–201, Apr. 2002.

[57] R. G. Congalton, "A review of assessing the accuracy of classifications of remotely sensed data," *Remote Sens. Environ.*, vol. 37, no. 1, pp. 35–46, Jul. 1991.
[58] G. M. Foody and A. Mathur, "A relative evaluation of multiclass image classification by support vector machines," *IEEE Trans. Geosci. Remote Sens.*, vol. 42, no. 6, pp. 1335–1343, Jun. 2004.
[59] S.-J. Yen and Y.-S. Lee, "Cluster-based under-sampling approaches for imbalanced data distributions," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 5718–5727, Apr. 2009.

**Zhiyong Lv** (Senior Member, IEEE) received the M.S. and Ph.D. degrees from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2008 and 2014, respectively.

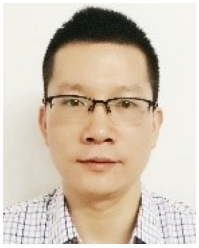
He was an Engineer of surveying at The First Institute of Photogrammetry and Remote Sensing, Xi'an, China, from 2008 to 2011. He is currently with the School of Computer Science and Engineering, Xi'an University of Technology, Xi'an. His research interests include multihyperspectral and high-resolution remotely sensed image processing, spatial feature extraction, neural networks, pattern recognition, deep learning, and remote sensing applications.

**Pengfei Zhang** is currently pursuing the master's degree in computer science with Xi'an University of Technology, Xi'an, Shaanxi, China.

He is interested in spatial–spectral feature extraction, pattern recognition, ground target detection, and land cover/land use change detection, through remote sensing image with high or very high spatial resolution (including satellite imagery and aerial images).

**Pei Wang** received the M.S. degree from the School of Remote Sensing and Information Engineering, Wuhan University, Wuhan, China, in 2008.

She is an Engineer of surveying with Shaanxi Bureau of Surveying and Mapping Geographic Information, Xi'an, China. Her research interests include high-resolution remotely sensed image processing, spatial feature extraction, neural networks, and pattern recognition.



**Weiwei Sun** (Senior Member, IEEE) received the B.S. degree in surveying and mapping and the Ph.D. degree in cartography and geographic information engineering from Tongji University, Shanghai, China, in 2007 and 2013, respectively.

From 2011 to 2012, he studied at the Department of Applied Mathematics, University of Maryland College Park, College Park, MD, USA. From 2014 to 2016, he studied at the State Key Laboratory for Information Engineering in Surveying, Mapping and Remote Sensing (LIESMARS), Wuhan University, Wuhan, China. From 2017 to 2018, he worked at the Department of Electrical and Computer Engineering, Mississippi State University, Starkville, MS, USA. He is currently a Full Professor with Ningbo University, Ningbo, Zhejiang, China. He has published more than 70 journal articles. His research interests include hyperspectral image processing with manifold learning, anomaly detection, and target recognition of remote sensing imagery using compressive sensing.

**Minghua Zhao** received the Ph.D. degree in computer science from Sichuan University, Chengdu, China, in 2006.

After that, she joined Xi'an University of Technology, Xi'an, China, where she is currently a Professor, a Doctoral Supervisor, and the Dean of the Information Management Center. She has published more than 90 academic papers and authorized 20 national invention patents. Her research interests include computer vision, intelligent information processing and visualization, and image recognition.

Dr. Zhao is a Senior Member of China Computer Federation (CCF) and China Society of Image and Graphics (CSIG) and a member of the CCF Virtual Reality and Visualization Technology Professional Committee and the CSIG Multimedia Professional Committee. She won two Second Prizes of Science and Technology in Shaanxi Province.

**Rui Zhu** received the Ph.D. degree from The Hong Kong Polytechnic University, Hong Kong, China, in 2018.

He is a Senior Scientist at the Institute of High Performance Computing, Agency for Science, Technology and Research, Singapore. He was a Research Assistant Professor at The Hong Kong Polytechnic University, Hong Kong, and a Post-Doctoral Associate at MIT Senseable City Laboratory. His study focused on GIScience, urban informatics, and solar energy with a publication of more than 80 SCI articles in journals, such as *Nature Communications*, *The Innovation*, and *Science Bulletin*.

Dr. Zhu is an Associate Editor of *Computer Science* (Springer Nature), an Editor of *Big Earth Data and Advances in Applied Energy*, and a Young Editor of *The Innovation*. He is also the PI/Co-I for several research grants and the Board of Director Member of Chinese Professional in Geographic Information Systems. His study has been reported by international media, such as Singapore TV, Lianhe Zaobao, and MIT News.