

Multiview Spatial–Spectral Active Learning for Hyperspectral Image Classification

Meng Xu^{ID}, *Member, IEEE*, Qingqing Zhao^{ID}, and Sen Jia^{ID}, *Senior Member, IEEE*

Abstract—Supervised classification algorithms on the intricate ground object information of hyperspectral images (HSIs) require a large number of training samples that are annotated manually for model learning. To reduce the labeling cost and improve training sample effectiveness, a multiview spatial–spectral active learning (MVSS-AL) model is proposed in this study. First, a committee model composed of collaborative representation classification is introduced to form a leave-one-class-out (LOCO) multiview strategy, which explores more effective information in the limited training data. Second, the sample query strategy is designed from the perspective of classification confidence (CC) and training contribution (TC). The most inconsistent high-quality samples are screened by making full use of iterative prediction information and spatial–spectral features contained in hyperspectral imagery. Finally, the spatial–spectral LOCO active learning (AL) model obtains target samples through two-layer screening in each iteration and utilizes a support vector machine to obtain the final classification results. The proposed method is tested on three real-world hyperspectral datasets, and the comparison with several novel methods shows that the proposed method is better in the classification performance of restricted sample training.

Index Terms—Active learning (AL), hyperspectral image (HSI) classification, multiview learning, representation learning.

NOMENCLATURE

Symbol	Meaning
H	Hyperspectral image.
h, w, d	Three dimensions of H .
K, N	Labels and number of classes of H .
T, D, C	Training set, unlabeled sample set and candidate set.
R, L	Residuals and predicted labels of D .
y, c	a sample of D and C .
ω^+	The most frequently assigned category.

Manuscript received April 22, 2021; revised June 18, 2021; accepted July 2, 2021. Date of publication July 20, 2021; date of current version January 17, 2022. This work was supported in part by the Research and Development Program of Guangdong Province under Grant 2020B1111020005; in part by the Natural Science Foundation of Guangdong Province under Grant 2021A1515011413; in part by the National Natural Science Foundation of China under Grant 41971300, Grant 61901278, and Grant 62001303; in part by the Key Project of Department of Education of Guangdong Province under Grant 2020ZDZX3045; and in part by the Shenzhen Scientific Research and Development Funding Program under Grant JCYJ20180305124802421, Grant JCYJ20180305125902403, and Grant 20200803152531004. (Corresponding author: Sen Jia.)

The authors are with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen 518060, China, and also with the Key Laboratory for Geo-Environmental Monitoring of Coastal Zone, Ministry of Natural Resources, Shenzhen University, Shenzhen 518060, China (e-mail: m.xu@szu.edu.cn; 18826054283@163.com; senjia@szu.edu.cn).

Digital Object Identifier 10.1109/TGRS.2021.3095292

I	Subscripts of the training subset that predicts y into the ω^+ class.
J	Subscripts of the training subset that predicts y into the second most easily assigned class.
x_T, y_T	The horizontal and vertical coordinate values of samples in T .
x_c, y_c	The horizontal and vertical coordinate values of c .
p	Number of samples which added to the volume calculation.
r	Number of samples which added to C .
z	Number of samples which added to T in each iteration.
b	Number of samples which selected from each class to form initial T .
β	Number of AL iteration.

I. INTRODUCTION

IN IMAGE classification methods with supervised or semi-supervised learning, manually labeled training samples are needed [1], [2]. Due to the complex spatial and spectral characteristics of hyperspectral images (HSIs), more resources and time need to be consumed in training and classification [3], [4]. Therefore, it is particularly important to intelligently construct a training set that is as small as possible and process a large or complex feature space through limited training information. In machine learning, the method that actively selects training samples to reduce computational cost is called active learning (AL) [5]–[7]. AL addresses the interaction between experts and the classification model by setting a reasonable classification model and sample selection strategy and assigning the samples selected for inclusion in the training set to manual calibration. For example, Samat *et al.* [8] used the pixel purity index (PPI) algorithm to distinguish pure pixel points and mixed points of hyperspectral data after dimensionality reduction by minimum noise fraction (MNF) and obtained the candidate sample set as the input for AL. Liu *et al.* [9] extracted the original hyperspectral features by Gabor filtering and extended morphological attribute profiles (EMAPs) and used the global error probability and Fisher ratio to screen out the features for AL. Sun *et al.* [10] utilized the posterior probability of the Gaussian process (GP) classifier to set up the heuristic active selection algorithm and used AL to implement the progressive updating of the GP. Bai *et al.* [11] used manual intervention to sort the training samples, selected samples and set stopping condition through the entropy function, and performed object detection on HSIs by ranking SVM model.

In the training stage of AL, the above methods only carry out single prediction of a single classifier, which belongs to the single-view AL method. For small-scale sample training, the classification accuracy of training samples is greatly limited if only a single-view classification method is utilized. To improve the model, scholars proposed multiview active learning (MVAL) [12]. In the MVAL method, a series of prediction results are obtained by designing multiple classifiers with multiple views in the same dataset. After integrating the information of each pixel, samples to be classified with high classification inconsistency and abundant information are selected for the training set [13], [14]. Hu *et al.* [15] used the Fisher ratio and conditional mutual information method on data processed by a 3-D Gabor kernel to select highly distinguishable and low redundancy data, but the experiments required the Gabor filter to be chosen by experienced selection in different datasets. Li *et al.* [16] generated multiview hyperspectral data from the three perspectives of pixels, subpixels, and superpixels and used an extreme learning machine and Markov random field (MRF) to optimize classification results. Liu *et al.* [17] proposed a semisupervised AL method based on superpixels for classifying small-scale hyperspectral samples. A dynamic clustering enhancement strategy was introduced to generate superpixels, and multinomial logistic regression (MLR) was used with MRF for sample prediction.

In addition, combining the ample spatial and spectral information of HSIs can significantly improve the analysis accuracy and guide the AL model [18]–[22]. Many scholars have carried out scientific research and exploration in this regard. In [23], a pruning strategy with a supervised hierarchical segmentation (HSEG) tree was designed to obtain spatial information, which was combined with original features to enhance spectral data to improve the AL classification performance. Zhang *et al.* [24] also used HSEG to extract spatial features and integrated spectral information to generate multiple views and then selected features according to feature importance (FI) to form dynamic views. In this method, a random forest (RF) classifier was adopted. Li *et al.* [25] presented an AL-based supervised Bayesian HSI segmentation method, which used a multilevel logistic (MLL) prior to obtaining spatial information and MLR to calculate the posterior probabilities of the categories. In addition, a modified breaking ties (MBT) method was proposed, in which category samples were selected iteratively to achieve unbiased sample selection. Afterward, an AL framework based on MRF was proposed based on the acquisition of spatial–spectral information by using isotropic MLL prior and MLR [26]. The confidence of the prediction was assessed according to whether the unlabeled samples changed before and after MRF operation. For the AL strategy itself, the work [27] uses the marginal probability to calculate the fuzzy mapping between the posterior probability and the output and utilizes the spectral angle information to select samples with low interclass redundancy and high spatial heterogeneity. Most of the above methods considered spatial–spectral information in the feature extraction step but ignored the remote sensing information combined with the spectrum in the AL framework.

With the rise of deep learning (DL), many studies have begun to combine AL with DL to explore more possibilities by forming a variety of deep AL models [28]–[31]. In the HSI field, the deep belief network (DBN) was combined with a weighted incremental dictionary learning (WI-DL) algorithm, and the selection of samples was considered in terms of both representativeness and uncertainty [32]. Haut *et al.* [33] proposed a Bayesian convolutional neural network (B-CNN) using spatial–spectral information, which realized network fitting under limited training samples. Deng *et al.* [34] combined transfer learning with AL and used three-layered sparse autoencoders to fuse spatial and spectral domain features. Furthermore, more reviews on AL combined with DL methods such as reinforcement learning can be found in [35] and [36]. A variety of studies, such as few-shot learning [37], are exploring DL classification methods under small-sample training sets. However, the fitting of deep neural networks [38] usually requires more labeled sample data than traditional machine learning methods.

In this article, we focus on optimizing the AL algorithm individually and make the following improvements to the issues that exist in HSI AL as follows.

- 1) In the AL method of HSIs, when the size of the training sample set is very small, the classification model is often limited by insufficient information, leading to inaccurate prediction results. Under this condition, the subsequent sample query strategy selects nonoptimal candidate samples based on the misclassification information. Therefore, to address this problem, a multiview leave-one-class-out (LOCO) strategy is proposed in this article. The limited training samples are divided into multiple subsets according to the ground object category, and the collaborative representation classification algorithm with low time consumption is selected as the classification model to output a series of multiview classification results. Afterward, the classification confidence (CC) of the samples to be classified is quantified through the classification label and residual, and the optimal candidate sample set is selected. By exploring the classification of the limited training samples from multiple perspectives, this method alleviates the overconfidence of the classification method on the mislabeled samples to a certain extent.
- 2) To make sufficient use of the ample spatial and spectral domain information of HSIs, a sample selection strategy combining space and spectrum is introduced in this article, which focuses on the use of space spectrum information by the AL strategy itself. The training contribution (TC) of candidate samples is ranked by analyzing the spatial and spectral relationship between candidate samples and training sample sets. Specifically, spatial selection uses the spatial information of the extracted pixels, whereas spectral selection calculates spectral volume to obtain the endmember distribution. The two processes form the TC model that addresses the uncertainty of samples and reduces the information redundancy among selected samples.

3) Based on the proposed strategies, a multiview active learning framework with spatial-spectral sample query (MVSS-AL) is constructed. Twofold sample screening is carried out in the model. In the first step, the sample is initially screened by the LOCO method, and the candidate set is composed of the samples that are difficult to predict by the classification model, namely, the samples with low CC. The second step is based on the candidate set selecting the samples with considerable differences from the effective information of the training samples and high TC for labeling. The proposed method gradually decreases the target set to realize the sample query and reduces the time cost of the sample query based on the effective selection.

By comparing a series of state-of-the-art methods in three real hyperspectral datasets, the results show that the proposed MVSS-AL model is superior in hyperspectral classification with a minimal training set and has a certain algorithm stability for datasets with different spatial resolutions and features.

The organizational structure of this article is as follows. Section II describes the collaborative representation classification and spectral volume studies related to our framework. Section III introduces the proposed MVSS-AL method in detail. The results of parametric analysis, ablation, and comparative experiments are presented in Section IV. The final summary is given in Section V.

II. RELATED WORKS

A. Collaborative Representation-Based Classification

Collaborative representation-based classification (CRC) is an improved supervised classification method based on sparse representation [39], [40]. Assume an image contains N categories, and an equal number of samples are selected from each class to join the training set. The training set can be expressed as $\mathbf{X} = [\mathbf{X}_1, \mathbf{X}_2, \dots, \mathbf{X}_N]$, where \mathbf{X}_i is the training dataset of class i . For a given test image \mathbf{Y} , the expression is

$$\mathbf{Y} = \mathbf{X} \cdot \hat{\boldsymbol{\rho}} \quad (1)$$

where $\hat{\boldsymbol{\rho}} = [\hat{\boldsymbol{\rho}}_1; \dots; \hat{\boldsymbol{\rho}}_i; \dots; \hat{\boldsymbol{\rho}}_N]$ and $\hat{\boldsymbol{\rho}}_i$ is the coefficient vector of class i . ℓ_2 -norm is adopted for the CRC method. The objective function ($\hat{\boldsymbol{\rho}}$) [41] is

$$(\hat{\boldsymbol{\rho}}) = \arg \min_{\hat{\boldsymbol{\rho}}} \{ \|\mathbf{Y} - \mathbf{X} \cdot \hat{\boldsymbol{\rho}}\|_2^2 + \lambda \|\hat{\boldsymbol{\rho}}\|_2^2 \}. \quad (2)$$

To make the solution $\hat{\boldsymbol{\rho}}$ stable, the regularization parameter λ is introduced as a constant with a very small value. The further solution of (2) can be derived by using regularized least squares

$$\hat{\boldsymbol{\rho}} = (\mathbf{X}^T \mathbf{X} + \lambda \cdot \mathbf{I})^{-1} \mathbf{X}^T \mathbf{Y} \quad (3)$$

where \mathbf{I} is the identity matrix. Combining (1) and (3), we can obtain the predicted sample residuals. Thus, the regularization residual r_i represented by a query sample $\mathbf{y} \in \mathbf{Y}$ and the i th class training samples \mathbf{X}_i can be obtained by

$$r_i = \frac{\|\mathbf{y} - \mathbf{X}_i \cdot \hat{\boldsymbol{\rho}}_i\|_2}{\|\hat{\boldsymbol{\rho}}_i\|_2}. \quad (4)$$

By comparing the residual values of each class, CRC results for \mathbf{y} can be output

$$\text{identity}(\mathbf{y}) = \arg \min_i \{r_i\}. \quad (5)$$

The samples to be classified in CRC are represented by the linear combination of training samples, and the labels of samples to be classified are determined by the residual results calculated for each class. Initially used in natural images, CRC was gradually widely introduced in HSI classification [42]–[44]. Due to the advantages of fast and efficient operation, CRC is selected as the multiview classifier in the AL model iteration process.

B. Simplex Volume

For a set of data points, there exists a simplex \mathcal{S} external containing them in the spectral eigenspace. Different points are selected as vertices from the data, resulting in different shapes and volumes of \mathcal{S} . Assume that p sample points $\{\mathbf{a}_1, \dots, \mathbf{a}_p\} \in \mathbb{R}^d$ ($d \geq p$) are selected, and they are affine independent. The corresponding volume V can be obtained by using the determinant [45]

$$V = \frac{1}{(p-1)!} \left| \det \left(\begin{bmatrix} 1 & \cdots & 1 \\ \mathbf{a}_1 & \cdots & \mathbf{a}_p \end{bmatrix} \right) \right|. \quad (6)$$

As the determinant operator, $\det(\cdot)$ needs to ensure that the row and column numbers of the matrix are equal. Therefore, the dimensionality of $\{\mathbf{a}_1, \dots, \mathbf{a}_p\}$ needs to be reduced to $p-1$.

Simplex volume is often calculated in hyperspectral spectral unmixing to extract specific class features [46], [47]. In this article, we apply it to the AL sample selection, which is specifically introduced in Section III-B.

III. PROPOSED METHOD

The proposed MVSS-AL framework is improved in three ways. First, a LOCO method is presented to process the original HSI by bagging to generate multiple views and utilize a set of CRC committee classifiers to obtain the multiview prediction results of the samples to be classified. Since there are sufficient differences and complementarities in the classification information between the committees, the validity and richness of the pixel information are guaranteed. Second, the multiview results are used to calculate the CC and TC of the samples to be classified, during which the spectral and spatial features are fully utilized. Finally, a dual sample query strategy is designed to perfect the AL model, gradually reducing the size of the candidate sample set and adding the samples to the training set after manual calibration. The algorithm expands the training sample pool of AL through several iterations, and the operation process of each iteration is shown in Fig. 1, which is introduced in more detail in the following. In addition, the mathematical variables used in the proposed method are shown in the Nomenclature.

A. View Generation by the LOCO Method

In small-sample training, there are a few training data available for the classifier, which often leads to unsatisfactory

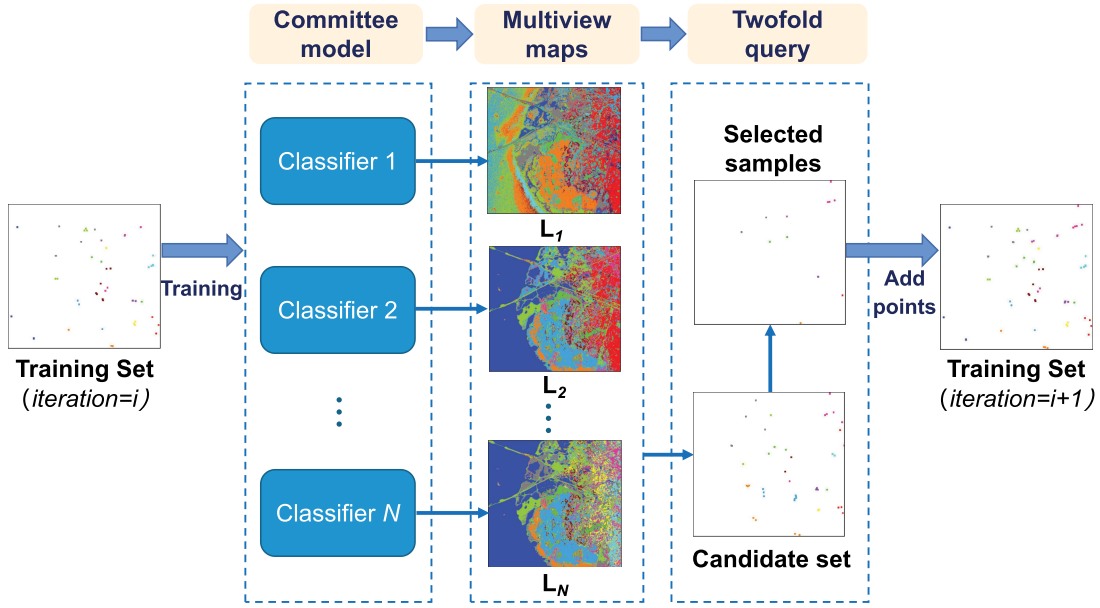


Fig. 1. System block diagram of the proposed MVSS-AL method for HSI classification in the i th iteration.

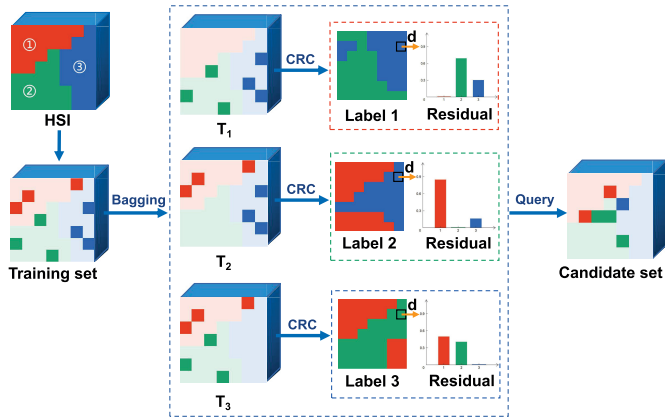


Fig. 2. AL process of HSIs based on the LOCO strategy.

classification results. The mismarked samples are the ones that are least likely to be marked correctly in the current classification condition. In other words, these unlabeled points are the samples that have the best effect of improving the classification accuracy. Therefore, it is significant to identify misclassified samples and determine how easily they are mismarked. Moreover, obtaining as much effective information as possible from the limited training samples is also helpful to improving model classification performance. In this article, the LOCO method is proposed to perform multiview prediction with differentiated and diverse classification signals.

For an HSI $\mathbf{H} \in \mathbb{R}^{h \times w \times d}$, where $h \times w$ is the spatial pixel number and d is the spectral size, assume that \mathbf{H} contains N kinds of ground objects and the labels are represented as $\mathbf{K} = \{1, \dots, N\}$. In the LOCO method, the same number of samples is randomly selected from each class to form the initial training set \mathbf{T} . As shown in Fig. 2, \mathbf{T} is divided into N subsets, and each subset retains $N - 1$ class samples with one

class of ground objects removed, denoted as $[\mathbf{T}_1, \mathbf{T}_2, \dots, \mathbf{T}_N]$. To achieve a rapid and efficient training process, a committee consisting of a set of CRC classifiers is introduced to train each subset separately. First, (3) is utilized to train each subset \mathbf{T}_i separately and obtain its corresponding coefficient

$$\hat{\mathbf{y}}_i = (\mathbf{T}_i^T \mathbf{T}_i + \lambda \cdot \mathbf{I})^{-1} \mathbf{T}_i^T \mathbf{D} \quad (7)$$

where $\mathbf{D} \in \mathbb{R}^{d \times m}$ is the set of m samples to be classified.

Second, each classifier trained by the training subset \mathbf{T}_i is inputted with all unlabeled samples to obtain the corresponding residual $\mathbf{R}_i \in \mathbb{R}^{1 \times m}$ and the predicted labels $\mathbf{L}_i \in \mathbb{R}^{1 \times m}$ in this view. The results are calculated by (4) and (5)

$$\mathbf{R}_i = \mathbf{e}_{i, \mathbf{L}_i} \quad (8)$$

$$\mathbf{L}_i = \arg \min_k \{\mathbf{e}_{i, k}\}, \quad i, k = 1, \dots, N \quad (9)$$

where $\mathbf{e}_{i, k}$ represents the residual value of \mathbf{D} predicted by the training subset \mathbf{T}_i in class k . Since each training subset only contains $N - 1$ classes, although each $\mathbf{e}_{i, k}$ has residual results for N categories, in practice, the residual values of the excluded class are infinite.

In this way, all samples to be classified are divided into $N - 1$ classes n times. The stability and CC of samples can be quantified by reasonable analysis of the classification label and residual information obtained. For example, as shown in Fig. 2, the unlabeled sample $\mathbf{d} \in \mathbb{R}^{d \times 1}$ has high CC, whose true category is 3 and the total class number is $N = 3$. Using the LOCO method to classify \mathbf{d} obtains the result that it is predicted to the third class $N - 1$ times, and only the training subset \mathbf{T}_3 that eliminates class 3 is predicted to be a non-3 class. Moreover, since \mathbf{d} is more likely to be assigned to class 3, the residual value of class 3 will be significantly smaller than that of other classes. To design the CC of samples based on this embodied information, Section III-B describes the proposed query method in detail.

B. Sample Selection

The setting of the sample query rule is one of the most important steps in AL. It determines the improvement effect of the selected sample on model classification accuracy and reflects its ability to integrate and utilize effective information. Therefore, we model the query strategy from two aspects: CC and TC. The strategy integrates the multiview prediction results of the committee model mentioned above and then calculates the CC of the unlabeled samples from multiple perspectives. It also analyzes the TC value of the samples to be classified based on the spatial-spectral characteristics of the original HSI. The design of the CC and TC is described as follows.

1) *Classification Confidence*: As mentioned above, the output results of the LOCO method can be used to screen high-uncertainty samples and provide more comprehensive and complementary information for the correct classification of the algorithm. Therefore, it is necessary to design a query strategy from which to extract effective content as the sample judgment. The CC value of a pixel $\mathbf{y} \in \mathbf{D}$ from the unlabeled sample pool is calculated as follows:

$$CC(\mathbf{y}) = \frac{\Delta G(\mathbf{y}) \times \Delta Q(\mathbf{y})}{A(\mathbf{y})}. \quad (10)$$

The calculation of the $CC(\mathbf{y})$ value can be analyzed from three aspects. In the first part, the $G(\cdot)$ function counts the number of times that \mathbf{y} is assigned to class $k = 1, \dots, N$

$$G(\mathbf{y}, k) = \sum_{i=1}^N l_i, \quad \text{s.t. } l_i = \begin{cases} 1, & \mathbf{L}_i(\mathbf{y}) = k \\ 0, & \mathbf{L}_i(\mathbf{y}) \neq k \end{cases} \quad (11)$$

$$\Delta G(\mathbf{y}) = \max_{\omega \in \mathbf{K}} G(\mathbf{y}, \omega) - \max_{\omega \in \mathbf{K} \setminus \omega^+} G(\mathbf{y}, \omega) \quad (12)$$

where ω^+ is the most frequently assigned class of all training subsets. The $\Delta G(\cdot)$ function can determine the two categories in which samples are most easily allocated and calculate the frequency difference of these two categories by counting the predicted label of different views. In other words, a sample with high CC is more likely to be predicted in the same category, that is, the number of times that it is classified into the ω^+ category should be much larger than other categories, so the value of ΔG is greater. Similarly, the difference in the residual values between the ω^+ category and other categories should be greater. This leads to the second part, the calculation of ΔQ

$$\Delta Q(\mathbf{y}) = \left| \min_{i \in \mathbf{I}} \mathbf{R}_i(\mathbf{y}) - \min_{j \in \mathbf{J}} \mathbf{R}_j(\mathbf{y}) \right| \quad (13)$$

and

$$\begin{aligned} \mathbf{I} &= \{i | \mathbf{L}_i(\mathbf{y}) = \arg \max_{\omega \in \mathbf{K}} G(\mathbf{y}, \omega), i = 1, \dots, N\} \\ \mathbf{J} &= \{j | \mathbf{L}_j(\mathbf{y}) = \arg \max_{\omega \in \mathbf{K} \setminus \omega^+} G(\mathbf{y}, \omega), j = 1, \dots, N\}. \end{aligned} \quad (14)$$

The sets \mathbf{I} and \mathbf{J} record the subscripts of the training subset that predicts \mathbf{y} into the ω^+ class and the second most easily assigned class, respectively, and the smallest residuals are selected for comparison. When sample \mathbf{y} is difficult to misclassify, the residual value corresponding to the correct class is more different from the corresponding value of other

classes, so the ΔQ value is higher. In addition, samples with high CC are more likely to maintain the same prediction results under the training of different views, so the third part considers the number A of different categories of \mathbf{y} classified by a group of committee models

$$A(\mathbf{y}) = |\text{Unique}(\mathbf{L}(\mathbf{y}))|. \quad (15)$$

The $|\text{Unique}(\cdot)|$ function is used to obtain the number of nonrepeating categories. By sorting the samples in the unlabeled sample pool from small to large, a specific number of samples with low CC can be selected and added to the training pool. Then, the heuristic strategy is set as

$$\alpha^{\text{LOCO}} = \arg \min_{\mathbf{y} \in \mathbf{D}} CC(\mathbf{y}). \quad (16)$$

This query strategy is combined with the view generation method proposed above to form the LOCO AL strategy.

2) *Training Contribution*: In the previous section, we analyzed and found samples that are easily misclassified by the model from the multiview prediction results, but there is no guarantee of the information difference between these samples and the samples in the training pool. To reduce the redundancy of effective content among these pixels, we propose a TC calculation method, which is analyzed from spatial and spectral dimension perspectives.

On the one hand, by observing the pseudocolor image of the hyperspectral data, it can be found that the various types of features tend to appear in patches, and the pixels of the same category usually have a short spatial distance. Therefore, for the candidate set $\mathbf{C} \in \mathbb{R}^{d \times r}$ with r samples waiting to be screened, the closest spatial distance exists between the candidate sample $\mathbf{c} \in \mathbf{C}$ and a training sample, and the label of the training sample affects the TC of the candidate sample \mathbf{c} to a certain extent. The Euclidean spatial relationship is calculated as

$$S(\mathbf{c}) = \begin{cases} \infty, & k = \varepsilon \\ \min \sqrt{(\mathbf{x}_T - \mathbf{x}_c)^2 + (\mathbf{y}_T - \mathbf{y}_c)^2}, & \text{else} \end{cases} \quad (17)$$

where \mathbf{x}_T and \mathbf{y}_T represent the horizontal and vertical coordinate values, respectively, of samples in training set \mathbf{T} and \mathbf{x}_c and \mathbf{y}_c are the horizontal and vertical coordinate values of candidate sample \mathbf{c} . In addition, k is the class of sample \mathbf{c} predicted by the LOCO method, ε represents the label of the training sample with the closest spatial distance from \mathbf{c} , and the label set of the training set is represented by \mathbf{L}_T

$$\begin{aligned} k &= \arg \max_{\omega \in \mathbf{K}} G(\mathbf{c}, \omega) \\ \varepsilon &= \mathbf{L}_T(\arg \min_{i \in \mathbf{T}} \sqrt{(x_i - x_c)^2 + (y_i - y_c)^2}). \end{aligned} \quad (18)$$

If the label of the training sample that is closest to \mathbf{c} is the same as the predicted category of \mathbf{c} , it is considered that from the spatial perspective, the candidate sample \mathbf{c} added to the training set can provide less difference information. Moreover, it is easier to classify correctly, so an infinite TC value is given. Conversely, the predicted class of the sample \mathbf{c} is different from the label of the closest training sample. Whether it is due to a prediction error or the distance between the two is close but the category is different, it can reflect

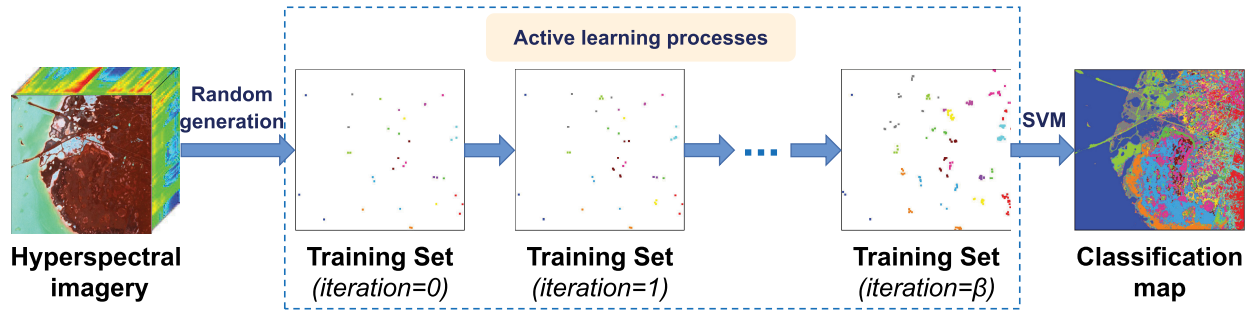


Fig. 3. Flowchart of the proposed AL algorithm for HSIs.

that candidate sample \mathbf{c} added to the training set has a great contribution to the ε category (if the categories are the same, the predicted accuracy of the ε category can be modified; if the categories are not same, the difference between the two categories increases). The minimum spatial distance $S(\mathbf{c})$ between \mathbf{c} and the training pool is recorded as a TC measure. The lower the value of $S(\mathbf{c})$ is, the higher the contribution of \mathbf{c} to the training.

On the other hand, effective information can be extracted from the affluent spectral dimensions of the HSI as a reference for the TC of candidate samples. Thus, the spectral volume strategy is introduced to measure the spectral relationship between candidate samples and the training set. As shown in (6), the total p samples are added to the volume calculation, \mathbf{a}_1 represents the candidate sample \mathbf{c} , and $\mathbf{a}_2, \dots, \mathbf{a}_p$ are the last $p - 1$ samples selected from the training set. It is worth mentioning that when the number of samples in the training set is less than $p - 1$, all training samples are selected for calculation. When the spectral volume value is large, the spectral difference between samples is large, this indicates that the candidate samples involved in the calculation and the training samples are highly complementary to each other in spectral information, and this candidate sample has a higher TC. Therefore, the TC calculation formula for the sample selection can be obtained

$$TC(\mathbf{c}) = V(\mathbf{c}) - S(\mathbf{c}). \quad (19)$$

In summary, the lower the spatial distance, the larger the spectral volume value of the candidate sample, and the larger the corresponding TC value, and the higher the corresponding TC.

C. AL Model Establishment

The LOCO method and the spatial-spectral query strategy are proposed above. To avoid all unlabeled samples from participating in the query, the proposed framework uses a two-step screening strategy to reduce the computational time consumption. First, all unlabeled samples that were calculated by (16) are sorted, and r samples with the smallest CC are selected to form a candidate set \mathbf{C} . Second, the sample query method combined with spatial and spectral consideration is introduced to select the candidate samples α with low

Algorithm 1 MVSS-AL Algorithm for HSI Classification

Input: The HSI \mathbf{H} , the number of candidate samples to be selected z , and the total iteration number β .

Output: The classification result of \mathbf{H} .

- 1: **BEGIN**
- 2: Initialize the training set \mathbf{T} , and the unlabeled set \mathbf{D} contains the left samples;
- 3: **for** $i = 1$ to β **do**
- 4: Using equations (8) and (9) to obtain the predicted labels \mathbf{L} and residuals \mathbf{R} ;
- 5: First query: Using equations (10) and (16) to obtain the candidate set \mathbf{C} ;
- 6: Using equation (17) to obtain the spatial distance \mathbf{S} ;
- 7: Using equation (6) to obtain the spectral volume \mathbf{V} ;
- 8: Second query: Using equations (19) and (20) to obtain the target candidate set \mathbf{C}' ;
- 9: Update $\mathbf{T} = \mathbf{T} \cup \mathbf{C}'$ and $\mathbf{D} = \mathbf{D} \setminus \mathbf{C}'$;
- 10: **end for**
- 11: $\mathbf{labels} = \text{SVM}(\mathbf{T}, \mathbf{D})$
- 12: **End**

CC and high TC

$$\alpha = \arg \min_{\mathbf{c} \in \mathbf{C}} \{CC(\mathbf{c}) - TC(\mathbf{c})\}. \quad (20)$$

z samples with the lowest α value are selected from the candidate set to join the training set. Then, a new round of prediction is conducted. After multiple iterations to obtain the target training set, the multiclass support vector machine method is utilized to predict the final classification result. The detailed steps of the proposed MVSS-AL method are summarized in Algorithm 1.

\mathbf{C}' is the candidate set that was selected to join the training set. For the initial training set, b points are randomly selected from each category to form \mathbf{T} , and all the remaining points automatically constitute the unmarked set \mathbf{D} . As shown in Fig. 3, regarding steps 4–9 in Algorithm 1 as an AL process, β is the number of AL runs, which can be adjusted according to needs. It is worth mentioning that the SVM classifier can determine the decision surface and support vectors through a small-scale training sample and has the ability to calculate complex feature space problems, so SVM is chosen as the classifier of the proposed small-samples training AL model.

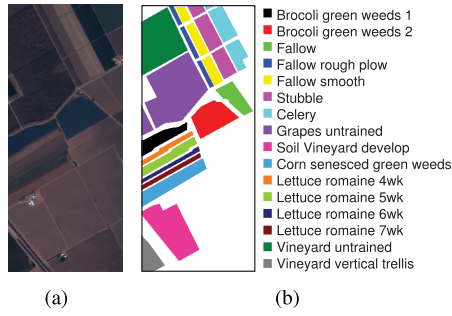


Fig. 4. (a) Pseudocolor illustration and (b) ground-truth map of the Salinas HSI (16 land-cover classes).

TABLE I
LAND-COVER CLASSES WITH THE NUMBER OF LABELED
SAMPLES FOR THE SALINAS HSI

Class	Land-Cover Type	No. of Samples
C1	Brocoli green weeds 1	2,009
C2	Brocoli green weeds 2	3,726
C3	Fallow	1,976
C4	Fallow rough plow	1,394
C5	Fallow smooth	2,678
C6	Stubble	3,959
C7	Celery	3,579
C8	Grapes untrained	11,271
C9	Soil vineyard develop	6,203
C10	Corn senesced green weeds	3,278
C11	Lettuce romaine 4wk	1,068
C12	Lettuce romaine 5wk	1,927
C13	Lettuce romaine 6wk	916
C14	Lettuce romaine 7wk	1,070
C15	Vineyard untrained	7,268
C16	Vineyard vertical trellis	1,807
	Total	54,129

IV. EXPERIMENTAL RESULTS

In this section, the main parameters of the MVSS-AL algorithm are analyzed to obtain better classification results. To prove the effectiveness of the proposed query method, an ablation experiment is carried out. In addition, our MVSS-AL method is trained on three real hyperspectral datasets, and the experimental results are compared with other state-of-the-art feature extraction and AL methods.

A. Hyperspectral Dataset and Parameter Setting

1) *Salinas HSI*: The first real hyperspectral dataset used in the experiment was collected by the AVIRIS sensor in Salinas Valley, CA, USA. The spatial dimension of the Salinas image is 512×217 , so there are 111 104 sample points in total, and the spatial resolution is 3.7 m. In addition, there are a total of 224 spectral bands, the spectral range is 400–2500 nm, and the bandwidth of each band is 10 nm. To reduce the influence of image noise, a total of 20 water absorption bands [108–112], [154–167], and 224 are removed. The dataset contains 16 kinds of ground objects, of which 54 129 points are labeled into these 16 classes. The pseudocolor image and the distribution of labels are shown in Fig. 4, and additional details about each class are shown in Table I.

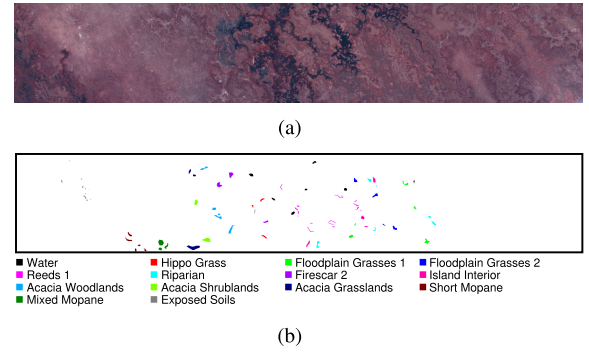


Fig. 5. (a) Pseudocolor illustration and (b) ground-truth map of Botswana HSI (14 land-cover classes).

TABLE II
LAND-COVER CLASSES WITH THE NUMBER OF LABELED
SAMPLES FOR BOTSWANA HSI

Class	Land-Cover Type	No. of Samples
C1	Water	270
C2	Hippo Grass	101
C3	Floodplain Grasses 1	251
C4	Floodplain Grasses 2	215
C5	Reeds	269
C6	Riparian	269
C7	Firescar	259
C8	Island Interior	203
C9	Acacia Woodlands	314
C10	Acacia Shrublands	248
C11	Acacia Grasslands	305
C12	Short Mopane	181
C13	Mixed Mopane	268
C14	Exposed Soils	95
	Total	3,248

2) *Botswana HSI*: The Botswana dataset is the second real hyperspectral dataset used in the experiment, which was captured by the NASA EO-1 satellite over the Okavango Delta, Botswana, in 2001–2004. The spatial resolution of the data is 30 m per pixel, and the spatial dimensions are 1476×256 . A total of 242 spectral bands were collected initially. After removing the uncalibrated and noisy bands, the remaining 145 bands (10–55, 82–97, 102–119, 134–164, and 187–220) are used for classification. A total of 3248 samples were labeled in the Botswana dataset, which was divided into 14 kinds of ground features. Specific class information and image labels are shown in Fig. 5 and Table II.

3) *KSC Hyperspectral Image*: The last real hyperspectral dataset used in the experiment is the Kennedy Space Center (KSC) image, which was acquired by a NASA AVIRIS (Airborne Visible/Infrared Imaging Spectrometer) sensor over the Kennedy Space Center in Florida. The spatial size of the original data is 614×512 , and the spatial resolution is 18 m. In this experiment, the cropped 421×444 image data are used, and the corresponding pseudocolor image and sample distribution map are shown in Fig. 6. The spectral range of the KSC image is 400–2500 nm, and the spectral resolution is 10 nm. There are 224 spectral bands in total. After removing the bands with low water absorption and signal-to-noise ratios, 176 feature bands are retained. The image contains a total

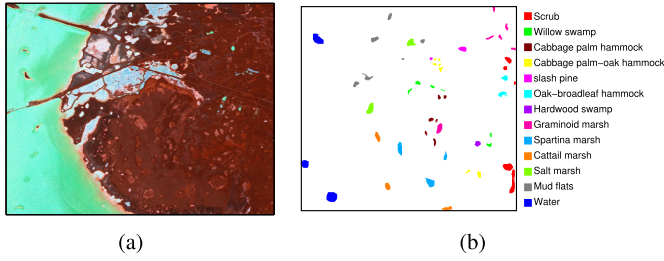


Fig. 6. (a) Pseudocolor illustration and (b) ground-truth map of the KSC HSI (13 land-cover classes).

TABLE III
LAND-COVER CLASSES WITH THE NUMBER OF LABELED
SAMPLES FOR THE KSC HSI

Class	Land-Cover Type	No. of Samples
C1	Scrub	761
C2	Willow swamp	243
C3	Cabbage palm hammock	255
C4	Cabbage palm-oak hammock	252
C5	slash pine	161
C6	Oak-broadleaf hammock	229
C7	Hardwood swamp	105
C8	Graminoid marsh	431
C9	Spartina marsh	520
C10	Cattail marsh	404
C11	Salt marsh	419
C12	Mud flats	503
C13	Water	927
	Total	5,211

of 5211 labeled ground object samples, which are classified into 13 categories, as detailed in Table III.

4) *Experimental Setting*: In all experiments in this article, three samples are randomly selected from each class to form the initial training set, that is, the total number of initial training samples of the datasets Salinas, Botswana, and KSC are 48 (16×3), 42 (14×3), and 39 (13×3), respectively. In each iteration, N sample points to be classified are selected and added into the training set, where N is the total number of classes in the dataset. As mentioned earlier, the values of N in the Salinas, Botswana, and KSC datasets are 16, 14, and 13, respectively. The total number of iterations $\beta = 15$. Thus, after 15 iterations, there are a total of 288 ($48 + 15 \times 16$), 252 ($42 + 15 \times 14$), and 234 ($39 + 15 \times 13$) training samples in the Salinas, Botswana, and KSC datasets.

B. Parameters Analysis

The MVSS-AL algorithm proposed in this article contains multiple parameter values. Due to the length limitation of this article, two main parameters are analyzed in this section: the number r of candidate samples and the number p of samples, which were added to the calculation of spectral volume. The changes in the overall classification accuracy of the three hyperspectral datasets under different values of the two parameters are shown in Fig. 7. The number of candidate samples in the experiment is set as $r = [20, 50, 250, 500, 1000]$, ranging from close to the number z of samples selected in each iteration to much greater than z , where the value

of z is the total number of feature categories N of the hyperspectral dataset. It can be seen in the figure that the overall classification accuracy of the three datasets increases at first and then decreases with the increase in r value, among which the OA value of Salinas and KSC datasets rise again in the later period.

Furthermore, the number of samples participating in the calculation of spectral volume is set as $p = [10, 30, 50, 70, 90]$. From the perspective of the r -axis, the OA values of the three datasets are high when the values of $r = [20, 50]$, and the results are better when $r = 50$. From the perspective of p -axis, when the p value is large (50, 70, 90), the OA results of the three datasets are relatively good. However, due to the large number of samples added to the calculation, the time efficiency of the algorithm is significantly reduced. Thus, choosing $p = 50$ can improve the computational efficiency better. In addition, when $r = [20, 50]$, the OA value of Botswana is relatively high under all the values of p . However, in Salinas and KSC datasets, the overall classification accuracy is significantly higher when $r = 50$ ($p = 50$). In summary, $r = 50$ and $p = 50$ are the most ideal group of parameters, which is selected for the experimental analysis in the ablation and comparison experiments.

C. Ablation Experiments

To demonstrate the effectiveness of the MVSS-AL sample query strategy proposed in (20), ablation experiments are conducted on the Salinas, Botswana, and KSC datasets in this section. Four ablation methods are designed and compared with the MVSS-AL sample query method. They differ from the proposed approach in the second step of the query strategy, including single LOCO, LOCO combined with spatial distance (LOCO + S), LOCO combined with spectral volume (LOCO + V), and spatial distance with spectral volume (S + V). Table IV shows the classification performance of the five ablation methods on the three hyperspectral datasets. The combined strategy of the MVSS-AL framework obtained the highest OA and kappa evaluation standard values after 15 iterations.

In addition, the OA and kappa values of the ablation comparison method with increasing iteration times are shown in line graphs. Fig. 8 shows that the LOCO + S method (gray line) in the Salinas dataset displays relatively poor performance results. As shown in Fig. 9, the LOCO method (red line) has the worst OA result in the Botswana dataset, while the classification accuracy of the LOCO + S method is higher. In the KSC dataset of Fig. 10, the accuracy of LOCO + S is closest to that of the proposed method in this article after several iterations. Moreover, the method that does not use the LOCO strategy in the second query is unstable. Observing the S + V strategy (blue-purple line), it can be seen that the accuracy of KSC and Botswana is not very satisfactory. Besides, the red broken lines in Figs. 8–10 represent the results of LOCO. It can be seen intuitively that in all ablation methods, the position of the red broken is almost at the bottom. This shows that the LOCO method as the benchmark has relatively poor classification performance in all datasets.

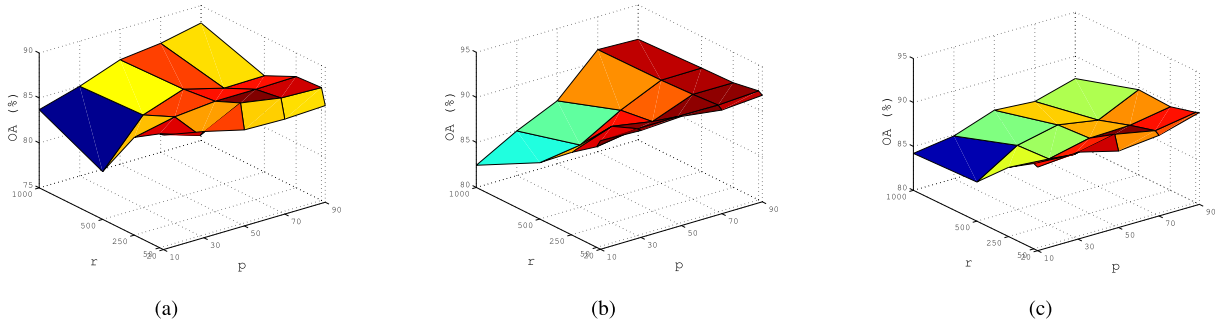


Fig. 7. Changes in classification accuracy of (a) Salinas, (b) Botswana, and (c) KSC datasets with two parameters: the number r of candidate samples and the number p of samples participating in volume calculation.

TABLE IV

CLASSIFICATION PERFORMANCE OF THE FIVE ABLATION APPROACHES ON SALINAS, BOTSWANA, AND KSC HSI AFTER 15 ITERATIONS

Dataset	Method	OA	AA	Kappa
Salinas	LOCO	86.47	91.74	84.85
	LOCO+S	85.74	91.89	84.09
	LOCO+V	87.38	92.35	85.94
	S+V	87.74	91.88	86.33
	MVSS-AL	88.47	92.28	87.13
Botswana	LOCO	91.10	91.61	90.36
	LOCO+S	91.30	91.40	90.66
	LOCO+V	91.24	90.96	90.51
	S+V	90.36	90.56	89.63
	MVSS-AL	92.14	92.25	91.49
KSC	LOCO	88.77	83.98	87.56
	LOCO+S	89.75	85.67	88.56
	LOCO+V	88.63	84.72	87.48
	S+V	88.64	84.41	87.43
	MVSS-AL	90.61	86.22	89.52

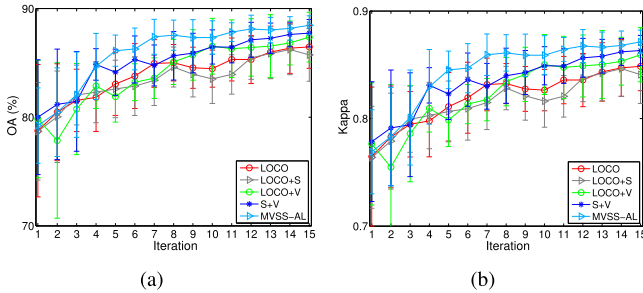


Fig. 8. Salinas dataset: (a) OA and (b) Kappa of five ablation methods.

In conclusion, both the CC and TC strategies proposed in this article can improve the classification performance of the AL algorithm to varying degrees. Although the contribution of different combined strategies to the overall classification performance varies from dataset to dataset, the proposed strategy achieves the best classification results in all three datasets.

D. Compared Experimental Results and Analysis

To evaluate the effectiveness of the proposed MVSS-AL method in the classification and selection strategy of HSIs, several classical and state-of-the-art methods are selected for comparison in this section. First, three feature extraction

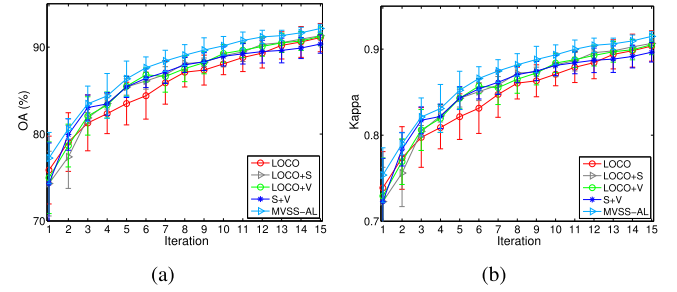


Fig. 9. Botswana dataset: (a) OA and (b) Kappa of five ablation methods.

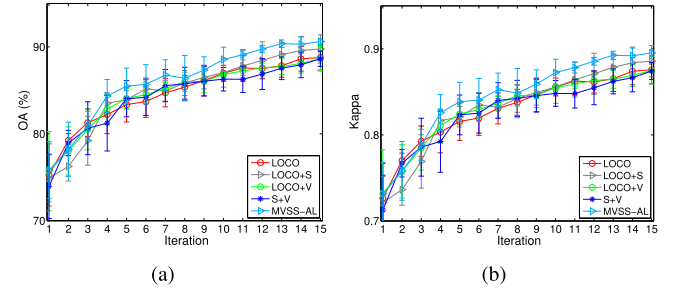


Fig. 10. KSC dataset: (a) OA and (b) Kappa of five ablation methods.

and classification methods are introduced: kernel principal component analysis (KPCA) [48], nonlinear multiple feature learning-based classification (NMFL) [49], and 3-D full convolutional neural network (3-D FCN) [50], [51]. Second, there are five AL methods: random selection as a benchmark, breaking ties (BT) [52], multilayer extreme learning machine-based autoencoder (GELM-AE-AL) [53], [54], superpixel and neighborhood assumption-based semisupervised active learning (SSAL-SN) [17], and deep active learning (CNN-AL-MRF) [55].

In the experiment, all AL methods use the same initial training set, which is composed of three randomly selected sample points from each class of ground objects, and the final training sample set is obtained after 15 iterations. According to [55], it should be noted that the CNN-AL-MRF method does not use multiple iterations. Therefore, according to the method introduced in that article, the total number of training samples obtained in the CNN-AL-MRF method is the same

TABLE V
RUNNING TIME (IN SECONDS) FOR DIFFERENT METHODS

Method	Random	KPCA	NMFL	3-D FCN	BT	GELM-AE-AL	SSAL-SN	CNN-AL-MRF	Ours
Salinas	50.75	13.68	7.71	114.57	89.01	276.09	14141	3882.6	2660.6
Botswana	59.54	15.95	17.31	209.29	109.45	603.55	35194	11529	10592.7
KSC	53.32	13.98	10.02	139.43	90.68	301.73	20227	5085.5	3376.9

TABLE VI
CLASSIFICATION PERFORMANCE OF THE NINE COMPARED APPROACHES ON THE SALINAS HSI AFTER 15 ITERATIONS (TOTAL 288 TRAINING SAMPLES)

Class	Random		KPCA		NMFL		3-D FCN		BT		GELM-AE-AL		SSAL-SN		CNN-AL-MRF		MVSS-AL	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
C1	97.78	3.08	92.34	3.87	99.52	0.36	60.55	46.27	98.57	0.48	69.45	4.61	96.68	4.57	75.56	32.87	96.35	3.44
C2	98.40	0.81	92.40	2.13	99.20	0.32	78.43	18.71	95.29	4.16	61.06	7.37	99.19	0.62	99.44	0.80	96.33	4.76
C3	75.50	14.71	87.07	2.65	98.43	0.98	70.05	27.99	64.99	27.76	17.22	9.01	89.14	4.38	75.10	14.96	89.62	6.51
C4	98.31	3.74	99.53	0.14	98.11	0.73	96.07	5.69	96.09	4.89	0.76	0.54	84.84	18.93	94.51	6.85	97.25	4.51
C5	98.55	1.51	87.89	1.65	98.22	0.40	41.05	36.66	90.81	10.56	4.56	4.93	88.55	9.30	97.74	1.82	95.54	2.56
C6	99.56	0.10	98.79	0.63	98.10	1.46	64.09	37.16	99.30	0.12	42.90	19.94	100.0	0.00	98.81	1.57	99.55	0.27
C7	99.45	0.13	96.69	1.86	99.10	0.63	27.30	21.95	89.81	18.48	47.88	21.92	94.44	4.57	99.40	0.85	96.72	4.11
C8	89.83	4.02	48.63	3.18	61.27	6.53	30.66	19.04	58.98	36.62	28.63	4.36	72.92	15.25	96.74	1.66	85.42	4.79
C9	98.63	1.25	92.24	2.91	98.74	0.94	32.87	34.89	94.17	2.29	16.22	6.23	99.03	0.41	99.75	0.56	98.22	1.18
C10	77.27	4.78	69.72	5.03	88.39	1.20	37.61	22.35	66.76	5.73	41.82	7.02	89.85	8.72	91.06	5.22	87.06	5.01
C11	42.55	33.69	91.52	1.83	91.33	2.21	51.00	41.89	91.67	3.76	45.71	17.65	40.56	18.62	91.90	11.45	92.26	5.16
C12	94.66	19.11	94.58	3.20	91.27	1.78	28.11	17.14	97.78	3.26	54.70	7.81	91.10	7.52	97.98	0.15	98.67	1.88
C13	96.98	0.84	94.40	3.35	95.49	2.04	70.92	33.75	97.43	0.84	63.07	11.08	92.40	5.97	49.18	39.09	97.82	0.68
C14	85.43	4.40	90.60	2.68	89.66	1.66	53.15	21.48	93.32	0.57	57.52	19.54	95.42	4.43	96.26	4.76	92.90	3.61
C15	44.14	12.94	62.25	4.46	57.92	3.86	32.58	28.76	52.68	27.19	77.35	4.07	39.52	1.21	17.85	23.88	57.77	7.98
C16	61.99	27.41	62.52	2.30	98.03	0.64	49.58	17.55	79.93	20.31	63.01	8.15	94.76	4.11	88.07	7.94	94.98	2.11
OA	84.65	2.10	77.37	1.18	84.15	0.96	43.91	4.89	78.36	5.35	41.78	2.85	73.77	1.60	83.92	2.64	88.47	0.87
AA	85.00	1.77	85.14	0.60	91.55	0.36	51.50	4.36	85.47	2.25	43.24	3.36	85.77	1.72	85.58	1.46	92.28	0.80
Kappa	0.83	0.02	0.75	0.01	0.82	0.01	0.40	0.05	0.76	0.06	0.33	0.04	0.71	0.04	0.82	0.03	0.87	0.01

as that of other methods after four training iterations. In the KPCA and 3-D FCN classification methods, 4–18 samples are randomly selected from each category as the training set, and all the remaining samples are used as the test set. Among all the comparison methods, 3-D FCN, SSAL-SDP, and CNN-AL-MRF use their built-in classifiers, while the other methods use a support vector machine classifier and set parameters through tenfold cross validation. To reflect the stability of the experimental results, ten experiments with different training samples are conducted for all methods, and the mean and standard deviation of each class and three evaluation criteria are obtained.

Furthermore, the running time (in seconds) of each comparative experimental method on three hyperspectral datasets is shown in Table V. Specifically, the training environment of the 3-D FCN DL method consists of a four-core/eight-way multitask Intel Xeon E-5 2699 processor with a processing speed of 2.20 GHz and two Tesla P100 GPUs. All the remaining machine learning methods are trained on a workstation with a 24-core Intel processor of 2.20 GHz and 128-GB RAM to get the calculation time.

1) *Salinas HSI*: Table VI shows the classification results of nine contrasting methods on the Salinas hyperspectral dataset, including the classification accuracy and standard deviation of each class and the whole, and the bold value is the optimal result. In addition, the number of initial training samples of the AL method is 48, and 16 samples are added in each iteration. All methods use a total of 288 training samples, accounting for 0.53% of the total labeled samples and 0.26%

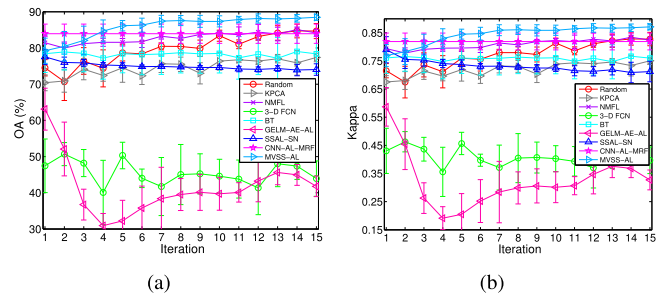


Fig. 11. Classification accuracy versus the number of iterations on the Salinas HSI. (a) OA. (b) Kappa.

of the total number of samples. In the classification accuracy of each class, although the CNN-AL-MRF method achieved the maximum single-class results value many times, the overall classification accuracy is affected due to the low accuracy of other classes (such as class 13 and 15). In addition, the OA value of all AL methods except ours is lower than that of the random method (especially the GELM-AE-AL method), indicating that extreme selection occurs in sample selection, which inhibits the classification accuracy of some classes. It can be seen in the average results of the ten experiments that the three evaluation coefficient values of MVSS-AL all obtain the optimal solution, and the standard deviation value is relatively low. This indicates that the proposed MVSS-AL method has relatively stable classification results in each kind of ground object, maintains the balance of all species of samples

TABLE VII
CLASSIFICATION PERFORMANCE OF THE NINE COMPARED APPROACHES ON THE BOTSWANA HSI
AFTER 15 ITERATIONS (TOTAL 252 TRAINING SAMPLES)

	Random		KPCA		NMFL		3-D FCN		BT		GELM-AE-AL		SSAL-SN		CNN-AL-MRF		MVSS-AL	
Class	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
C1	98.31	1.59	99.90	0.20	98.89	1.24	100.0	0.00	94.15	4.86	47.86	8.52	99.75	0.21	98.15	3.70	98.51	1.40
C2	84.09	13.22	96.39	4.61	92.29	2.35	82.47	19.26	95.34	2.09	25.15	10.93	93.15	10.56	69.31	24.75	97.07	3.56
C3	92.12	2.33	95.60	1.18	74.16	2.65	66.07	33.94	85.69	9.14	65.63	10.56	78.75	26.94	91.24	9.67	95.38	1.11
C4	89.88	5.17	88.96	7.18	90.56	3.26	41.23	22.86	79.13	10.87	82.14	11.07	81.54	19.04	84.65	17.46	99.00	1.06
C5	87.49	2.99	76.39	9.20	59.04	9.20	45.81	22.74	67.36	12.31	26.91	15.39	81.79	12.61	42.19	28.55	88.44	4.78
C6	62.64	11.13	65.54	7.54	52.67	7.12	37.21	18.16	50.75	17.30	58.19	18.45	64.52	20.55	43.49	23.98	82.60	4.53
C7	94.75	4.83	92.01	0.86	94.94	2.95	48.31	40.22	76.66	15.75	85.02	1.52	100.0	0.00	88.32	7.82	98.16	1.15
C8	94.45	2.24	93.92	2.94	93.41	1.64	60.20	26.65	78.63	11.79	86.37	8.61	88.36	16.85	100.0	0.00	96.64	1.50
C9	86.53	5.81	81.93	5.38	68.45	4.11	46.45	23.35	67.60	13.26	68.62	11.71	92.33	6.25	98.41	1.86	90.02	4.39
C10	84.69	8.46	76.85	1.89	68.52	4.22	35.00	26.28	62.65	12.08	50.97	21.33	61.52	18.05	81.35	21.53	87.53	1.70
C11	88.81	3.08	87.72	3.55	72.61	4.13	78.41	11.57	72.35	4.48	79.60	10.65	92.84	6.24	84.51	17.93	90.18	2.61
C12	90.48	5.87	95.86	0.59	72.15	2.66	53.33	16.05	78.79	20.24	56.46	17.55	98.92	1.86	55.39	37.37	88.43	6.72
C13	86.10	10.49	81.90	2.00	55.44	6.93	38.56	33.32	46.79	19.44	58.96	22.33	66.21	16.30	97.57	3.05	92.07	4.27
C14	84.06	3.05	99.35	0.75	83.64	2.85	66.59	36.38	79.77	23.01	72.47	13.60	98.19	3.14	81.58	17.07	88.52	7.99
OA	87.46	0.94	86.45	0.78	75.04	1.60	55.81	14.09	71.04	3.62	62.46	3.03	81.13	3.45	80.54	3.39	92.14	0.49
AA	87.46	1.29	88.02	0.70	76.91	1.47	57.12	13.57	73.98	2.48	61.74	3.31	85.56	2.21	79.73	2.65	92.25	0.82
Kappa	0.86	0.01	0.85	0.01	0.72	0.02	0.52	0.15	0.69	0.04	0.59	0.03	0.80	0.04	0.79	0.04	0.91	0.01

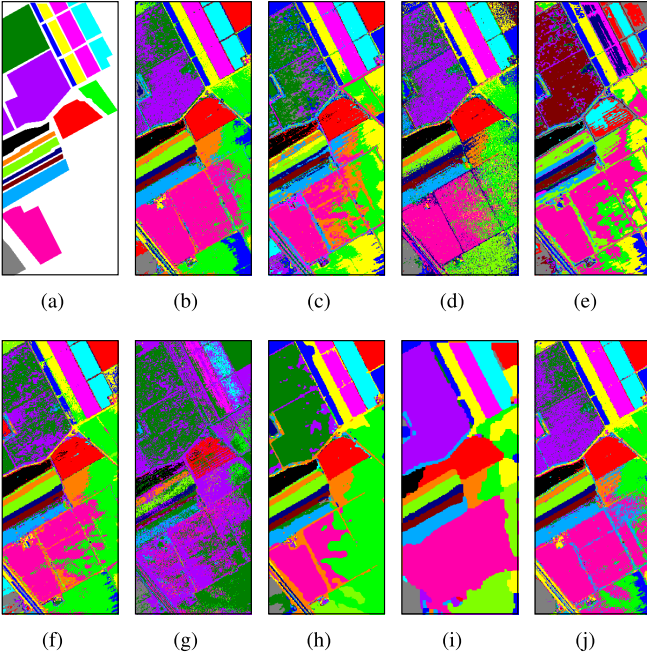


Fig. 12. Salinas dataset: (a) distribution map of all labeled samples and classification maps obtained by (b) random (83.92%), (c) KPCA (78.08%), (d) NMFL (84.72%), (e) 3-D FCN (44.75%), (f) BT (78.09%), (g) GELM-AE-AL (41.70%), (h) SSAL-SN (73.99%), (i) CNN-AL-MRF (82.77%), and (j) MVSS-AL (88.79%) after 15 iteration (the percentage in the brackets is the OA value).

in the training set, and improves the overall classification accuracy.

Moreover, Fig. 11 shows the variation in OA and kappa values for all methods from 1 to 15 iterations. It should be noted that since the CNN-AL-MRF method did not use the iterative method in the experiment, only the result of 288 training samples is shown in the line chart. It can be seen in the figure that the classification accuracy of random, KPCA, NMFL, and MVSS-AL increases with the number of itera-

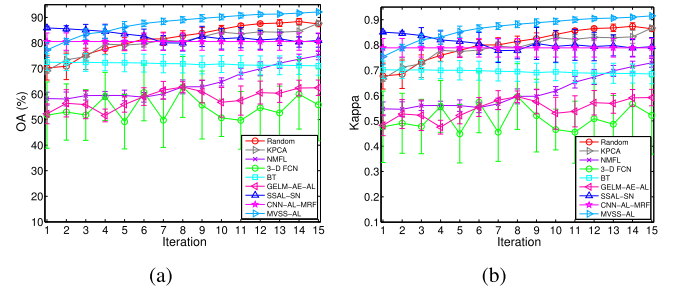


Fig. 13. Classification accuracy versus the number of iterations on the Botswana HSI. (a) OA. (b) Kappa.

tions. The BT, GELM-AE-AL, and SSAL-SN methods show a constant or even decreasing trend with the increase in the number of training samples, which reflects the shortcomings of the sample selection strategy. Correspondingly, Fig. 12 shows the classification maps and OA values obtained by the nine contrast algorithms under the training of 288 samples. Compared with other methods, the classification map [see Fig. 12(j)] of the proposed MVSS-AL algorithm is more similar to the structure of the real label [see Fig. 12(a)], which shows that our method can better retain the ground object structure information based on improving the classification accuracy.

2) *Botswana HSI*: Table VII lists the classification results of the nine compared methods on the Botswana hyperspectral dataset. The classification accuracy obtained by training with 252 samples accounts for 7.8% and 0.07% of the labeled samples and the total number of samples, respectively. It can be seen in the table that the classification results of the 3-D FCN and GELM-AE-AL method are much worse than those of the other methods, indicating that the neural network methods are not adequate for small-sample training. However, the proposed MVSS-AL method achieves the highest classification accuracy under multiple classes and all evaluation criteria,

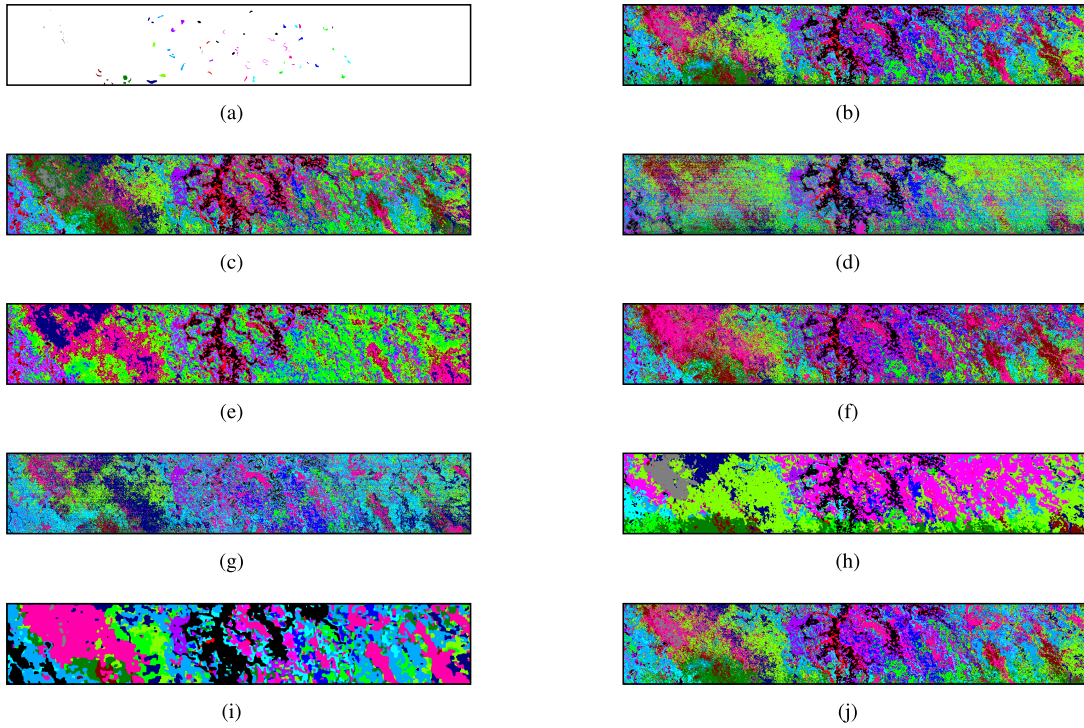


Fig. 14. Botswana dataset: (a) distribution map of all labeled samples and classification maps obtained by (b) random (87.93%), (c) KPCA (86.35%), (d) NMFL (74.70%), (e) 3-D FCN (54.93%), (f) BT (71.21%), (g) GELM-AE-AL (62.84%), (h) SSAL-SN (82.18%), (i) CNN-AL-MRF (69.02%), and (j) MVSS-AL (92.46%) after 15 iteration (the percentage in the brackets is the OA value).

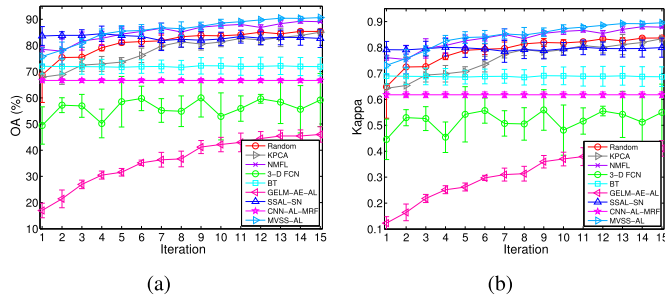


Fig. 15. Classification accuracy versus the number of iterations on the KSC HSI. (a) OA. (b) Kappa.

which reflects the effectiveness of the AL strategy of our method.

OA and kappa values for all compared methods increasing with the number of iterations are shown in Fig. 13. Similarly, the CNN-AL-MRF method only shows the classification result of the final 252 training samples in the figure. It can be seen in the figure that the classification accuracy of the SSAL-SN method decreases gradually with the increase in the number of training samples, while the classification results of MVSS-AL continuously improve due to the selection of samples with a high TC. In addition, Fig. 14 shows the classification maps and the corresponding OA values of the nine methods. All the above data and figures show that even when the spatial distribution of the hyperspectral data is relatively complex, the proposed MVSS-AL method is still stable and effective to a certain extent.

3) *KSC Hyperspectral Image*: When 234 samples (4.5% of the labeled sample and 0.13% of the total samples) are selected as the training set, Table VIII shows the mean and standard deviation of each class and three evaluation criteria trained by the KSC dataset in the nine compared methods. Although the CNN-AL-MRF method achieves the best classification accuracy in some classes, its overall classification accuracy is the lowest among all AL methods, indicating that this method is not suitable for the classification of small samples in KSC datasets. However, the proposed MVSS-AL method obtains the highest classification accuracy among multiple ground features and the best value under all three evaluation criteria, which indicates that the proposed strategy can select valuable samples.

Fig. 15 shows the relationship between the overall classification results of the nine methods and the number of iterations. The CNN-AL-MRF method only showed the classification result of 234 training samples. In this figure, BT and SSAL-SN methods do not significantly improve their classification accuracy with the increase in training samples. Instead, the classification results of MVSS-AL, NMFL, GELM-AE-AL, and the random strategy are constantly optimized. Similarly, in Fig. 16, the classification maps and OA values of the nine methods in a single experiment of 234 training samples are visually shown. It can be seen in the figure that the MVSS-AL method achieves the highest overall classification accuracy while maintaining the distribution of image features, which is better than that of the other comparison methods, demonstrating the classification performance of our method again.

TABLE VIII

CLASSIFICATION PERFORMANCE OF THE NINE COMPARED APPROACHES ON KSC HSI AFTER 15 ITERATIONS (TOTAL 234 TRAINING SAMPLES)

Class	Random		KPCA		NMFL		3-D FCN		BT		GELM-AE-AL		SSAL-SN		CNN-AL-MRF		MVSS-AL	
	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std	mean	std
C1	94.01	0.91	84.86	4.52	76.83	2.37	31.89	30.23	64.87	18.77	56.61	12.05	84.83	5.54	100.0	0.00	89.59	3.35
C2	82.53	10.94	79.89	6.33	88.26	2.99	42.76	12.97	73.12	10.81	43.62	13.51	90.92	6.75	15.91	13.78	84.39	9.14
C3	80.34	11.87	79.62	4.02	86.99	2.29	40.16	26.77	58.30	15.70	65.94	12.30	66.23	12.86	79.30	35.86	82.80	8.23
C4	64.26	19.46	59.62	13.28	56.67	6.29	31.94	27.16	47.29	10.96	35.16	11.22	68.28	25.67	97.09	2.52	72.43	7.09
C5	20.49	12.39	55.94	8.23	79.24	4.53	5.99	11.66	61.55	3.16	40.37	14.95	59.56	6.58	75.78	2.15	66.42	3.72
C6	47.78	12.64	47.99	5.96	70.65	4.38	9.16	11.32	37.28	5.20	23.49	9.48	69.41	15.16	17.90	1.51	73.74	6.87
C7	69.29	13.05	79.89	9.03	94.40	2.73	37.03	29.78	74.51	11.18	77.90	21.08	96.96	6.17	94.92	4.40	82.50	12.88
C8	73.03	11.79	87.89	2.05	92.62	2.62	44.36	19.48	63.26	8.16	24.08	7.49	72.39	14.01	75.10	10.58	85.76	2.20
C9	97.96	0.52	90.94	4.02	93.71	1.94	69.15	34.90	77.85	27.17	53.73	9.77	75.40	18.82	46.79	12.08	95.74	1.37
C10	96.43	1.58	80.96	4.01	97.49	1.34	66.35	15.00	78.86	6.92	4.80	4.14	95.92	5.93	0.00	0.00	97.72	0.74
C11	94.96	1.73	95.14	1.89	94.55	1.03	77.78	28.50	87.25	5.65	1.77	1.72	99.46	1.21	100.0	0.00	95.21	2.59
C12	81.91	2.87	87.22	4.48	94.33	2.90	81.28	2.04	53.37	22.60	24.05	7.12	89.37	13.79	9.74	4.13	94.62	2.08
C13	100.0	0.00	99.83	0.06	100.0	0.00	99.61	0.48	99.95	0.06	99.13	1.06	97.17	3.83	100.0	0.00	100.0	0.00
OA	85.35	0.41	85.07	0.46	89.09	0.31	59.14	10.14	71.91	3.45	46.00	2.95	82.76	2.45	66.63	0.73	90.61	0.31
AA	77.15	1.26	79.21	0.79	86.02	0.37	49.03	7.68	67.50	2.80	42.36	3.46	81.99	2.41	62.50	1.13	86.22	1.69
Kappa	0.84	0.01	0.83	0.01	0.88	0.01	0.55	0.11	0.69	0.04	0.41	0.03	0.81	0.03	0.62	0.01	0.90	0.01

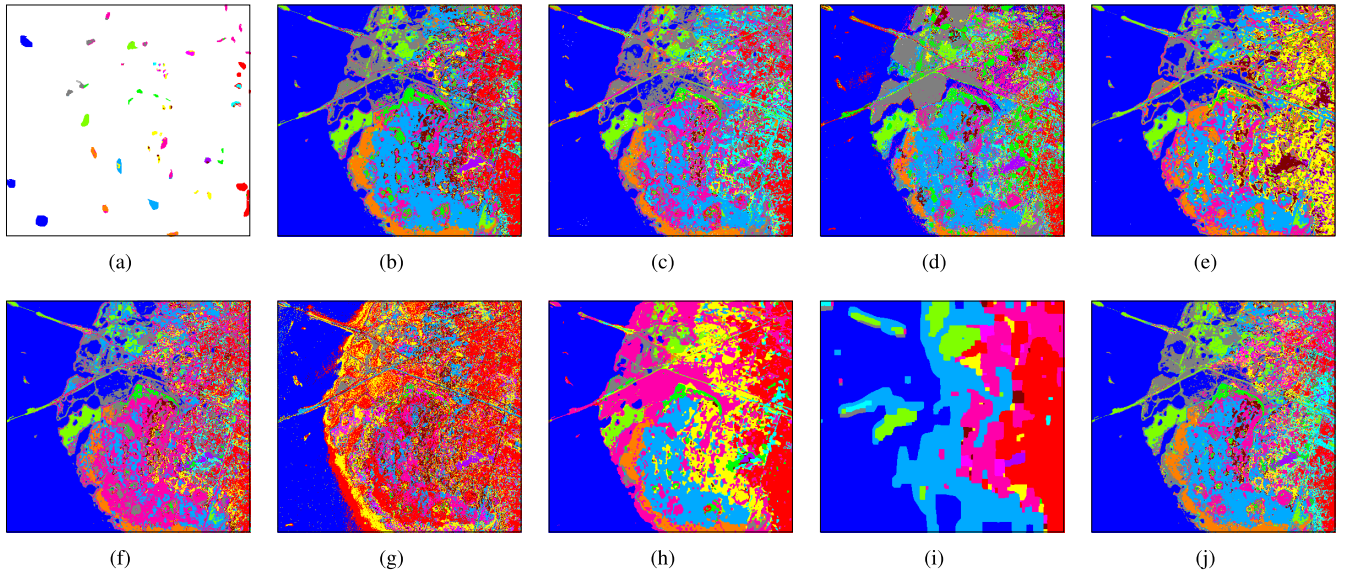


Fig. 16. KSC dataset: (a) distribution map of all labeled samples and classification maps obtained by (b) random (84.89%), (c) KPCA (84.40%), (d) NMFL (58.86%), (e) 3-D FCN (71.55%), (f) BT (46.81%), (g) GELM-AE-AL (82.63%), (h) SSAL-SN (67.12%), (i) CNN-AL-MRF (91.02%), and (j) MVSS-AL (91.02%) after 15 iteration (the percentage in the brackets is the OA value).

V. CONCLUSION

In this article, a multiview LOCO AL framework combined with spatial and spectral features (MVSS-AL) is constructed to obtain comprehensive classification information of HSIs and improve the credibility of the model prediction results. The model trains a group of collaborative representation classifiers after bagging the original hyperspectral dataset according to the category and performs diversified information analysis on the multiview prediction results of the samples to be classified to quantify the CC of the samples. Then, the spatial structure relationship and the spectral volume value between the candidate samples and the training samples are calculated to obtain the TC, thereby generating a spatial-spectral sample query strategy. Finally, the iteratively expanded training sample pool is input into the support vector machine to output the classification results, and the spatial-spectral AL framework of LOCO is constructed. After ablation and comparative

experimental analysis, it is shown that our method effectively utilizes the spatial-spectral characteristics of HSIs and reflects excellent universality and classification performance.

However, there are still some areas that can be optimized to improve the AL framework in this article. In the future, we will attempt to ameliorate the existing sample query model and formulate a feature extraction method to apply effective features instead of original images to enhance the efficiency of the algorithm. Similarly, the introduction of semisupervised learning methods to further reduce sample costs is also a research direction to be considered.

REFERENCES

- [1] W. Li and Q. Du, "A survey on representation-based classification and detection in hyperspectral remote sensing imagery," *Pattern Recognit. Lett.*, vol. 83, no. 2, pp. 115–123, Nov. 2015.
- [2] S. S. Sawant and M. Prabukumar, "Semi-supervised techniques based hyper-spectral image classification: A survey," in *Proc. Innov. Power Adv. Comput. Technol. (i-PACT)*, Apr. 2017, pp. 1–8.

- [3] L. Zhou, X. Bai, X. Liu, J. Zhou, and E. R. Hancock, "Learning binary code for fast nearest subspace search," *Pattern Recognit.*, vol. 98, Feb. 2020, Art. no. 107040.
- [4] L. Zhou *et al.*, "Subspace structure regularized nonnegative matrix factorization for hyperspectral unmixing," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 13, pp. 4257–4270, 2020.
- [5] D. Angluin, "Queries and concept learning," *Mach. Learn.*, vol. 2, no. 4, pp. 319–342, Apr. 1988.
- [6] D. J. C. MacKay, "Information-based objective functions for active data selection," *Neural Comput.*, vol. 4, no. 4, pp. 590–604, Jul. 1992.
- [7] D. Cohn, L. Atlas, and R. Ladner, "Improving generalization with active learning," *Mach. Learn.*, vol. 15, no. 2, pp. 201–221, May 1994.
- [8] A. Samat, J. Li, S. Liu, P. Du, Z. Miao, and J. Luo, "Improved hyperspectral image classification by active learning using pre-designed mixed pixels," *Pattern Recognit.*, vol. 51, pp. 43–58, Mar. 2016.
- [9] C. Liu, L. He, Z. Li, and J. Li, "Feature-driven active learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 1, pp. 341–354, Jan. 2018.
- [10] S. Sun, P. Zhong, H. Xiao, and R. Wang, "Active learning with Gaussian process classifier for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 4, pp. 1746–1760, Apr. 2015.
- [11] X. Bai, H. Zhang, and J. Zhou, "VHR object detection based on structural feature extraction and query expansion," *IEEE Trans. Geosci. Remote Sens.*, vol. 52, no. 10, pp. 6508–6520, Oct. 2014.
- [12] I. Muslea, S. Minton, and C. A. Knoblock, "Active learning with multiple views," *J. Artif. Intell. Res.*, vol. 27, no. 1, pp. 203–233, Oct. 2006.
- [13] W. Di and M. M. Crawford, "Active learning via multi-view and local proximity co-regularization for hyperspectral image classification," *IEEE J. Sel. Topics Signal Process.*, vol. 5, no. 3, pp. 618–628, Jun. 2011.
- [14] X. Zhou, S. Prasad, and M. Crawford, "Wavelet domain multi-view active learning for hyperspectral image analysis," in *Proc. 6th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens. (WHISPERS)*, Jun. 2014, pp. 1–4.
- [15] J. Hu, Z. He, J. Li, L. He, and Y. Wang, "3D-Gabor inspired multiview active learning for spectral-spatial hyperspectral image classification," *Remote Sens.*, vol. 10, no. 7, p. 1070, Jul. 2018.
- [16] Y. Li, T. Lu, and S. Li, "Subpixel-pixel-superpixel-based multiview active learning for hyperspectral images classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4976–4988, Jul. 2020.
- [17] C. Liu, J. Li, and L. He, "Superpixel-based semisupervised active learning for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 12, no. 1, pp. 357–370, Jan. 2019.
- [18] 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.
- [19] S. Patra, K. Bhardwaj, and L. Bruzzone, "A spectral-spatial multicriteria active learning technique for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 12, pp. 5213–5227, Dec. 2017.
- [20] M. Ahmad *et al.*, "Multiclass non-randomized spectral-spatial active learning for hyperspectral image classification," *Appl. Sci.*, vol. 10, no. 14, p. 4739, 2020.
- [21] C. Mu, J. Liu, Y. Liu, and Y. Liu, "Hyperspectral image classification based on active learning and spectral-spatial feature fusion using spatial coordinates," *IEEE Access*, vol. 8, pp. 6768–6781, 2020.
- [22] K. Bhardwaj, A. Das, and S. Patra, "Spectral-spatial active learning in hyperspectral image classification using threshold-free attribute profile," in *Proc. Int. Conf. Frontiers Comput. Syst.*, Singapore, 2021, pp. 45–56.
- [23] Z. Zhang, E. Pasolli, M. M. Crawford, and J. C. Tilton, "An active learning framework for hyperspectral image classification using hierarchical segmentation," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 9, no. 2, pp. 640–654, Feb. 2016.
- [24] Z. Zhang, E. Pasolli, and M. M. Crawford, "An adaptive multiview active learning approach for spectral-spatial classification of hyperspectral images," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 4, pp. 2557–2570, Apr. 2020.
- [25] J. Li, J. M. Bioucas-Dias, and A. Plaza, "Hyperspectral image segmentation using a new Bayesian approach with active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 49, no. 10, pp. 3947–3960, Oct. 2011.
- [26] S. Sun, P. Zhong, H. Xiao, and R. Wang, "An MRF model-based active learning framework for the spectral-spatial classification of hyperspectral imagery," *IEEE J. Sel. Topics Signal Process.*, vol. 9, no. 6, pp. 1074–1088, Sep. 2015.
- [27] M. Ahmad *et al.*, "Spatial prior fuzziness pool-based interactive classification of hyperspectral images," *Remote Sens.*, vol. 11, no. 9, p. 1136, May 2019.
- [28] J. Li, "Active learning for hyperspectral image classification with a stacked autoencoders based neural network," in *Proc. 7th Workshop Hyperspectral Image Signal Process.: Evol. Remote Sens. (WHISPERS)*, Jun. 2015, pp. 1–4.
- [29] H. Ranganathan, H. Venkateswara, S. Chakraborty, and S. Panchanathan, "Deep active learning for image classification," in *Proc. IEEE Int. Conf. Image Process. (ICIP)*, Sep. 2017, pp. 3934–3938.
- [30] K. Wang, D. Zhang, Y. Li, R. Zhang, and L. Lin, "Cost-effective active learning for deep image classification," *IEEE Trans. Circuits Syst. Video Technol.*, vol. 27, no. 12, pp. 2591–2600, Dec. 2017.
- [31] Y. Gal, R. Islam, and Z. Ghahramani, "Deep Bayesian active learning with image data," in *Proc. Int. Conf. Mach. Learn. (PMLR)*, vol. 70, 2017, pp. 1183–1192.
- [32] P. Liu, H. Zhang, and K. B. Eom, "Active deep learning for classification of hyperspectral images," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 10, no. 2, pp. 712–724, Feb. 2017.
- [33] J. M. Haut, M. E. Paoletti, J. Plaza, J. Li, and A. Plaza, "Active learning with convolutional neural networks for hyperspectral image classification using a new Bayesian approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 56, no. 11, pp. 6440–6461, Nov. 2018.
- [34] C. Deng, Y. Xue, X. Liu, C. Li, and D. Tao, "Active transfer learning network: A unified deep joint spectral-spatial feature learning model for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 57, no. 3, pp. 1741–1754, Mar. 2019.
- [35] P. Ren *et al.*, "A survey of deep active learning," 2020, *arXiv:2009.00236*. [Online]. Available: <http://arxiv.org/abs/2009.00236>
- [36] K. Matsushita, K. Matsushita, and Hasebe, *Deep Active Learning*. Springer, 2018.
- [37] J. Lu, P. Gong, J. Ye, and C. Zhang, "Learning from very few samples: A survey," 2020, *arXiv:2009.02653*. [Online]. Available: <http://arxiv.org/abs/2009.02653>
- [38] C. Wang, X. Bai, S. Wang, J. Zhou, and P. Ren, "Multiscale visual attention networks for object detection in VHR remote sensing images," *IEEE Geosci. Remote Sens. Lett.*, vol. 16, no. 2, pp. 310–314, Feb. 2019.
- [39] J. Wright, A. Y. Yang, A. Ganesh, S. S. Sastry, and Y. Ma, "Robust face recognition via sparse representation," *IEEE Trans. Pattern Anal. Mach. Intell.*, vol. 31, no. 2, pp. 210–227, Feb. 2009.
- [40] 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.
- [41] 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.
- [42] W. Li and Q. Du, "Joint within-class collaborative representation for hyperspectral image classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 7, no. 6, pp. 2200–2208, Jun. 2014.
- [43] S. Jia, L. Shen, and Q. Li, "Gabor feature-based collaborative representation for hyperspectral imagery classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 2, pp. 1118–1129, Feb. 2015.
- [44] J. Jiang, C. Chen, Y. Yu, X. Jiang, and J. Ma, "Spatial-aware collaborative representation for hyperspectral remote sensing image classification," *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 404–408, Mar. 2017.
- [45] X. Geng, K. Sun, L. Ji, Y. Zhao, and H. Tang, "Optimizing the endmembers using volume invariant constrained model," *IEEE Trans. Image Process.*, vol. 24, no. 11, pp. 3441–3449, Nov. 2015.
- [46] J. Li, A. Agathos, D. Zaharie, J. M. Bioucas-Dias, A. Plaza, and X. Li, "Minimum volume simplex analysis: A fast algorithm for linear hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 9, pp. 5067–5082, Sep. 2015.
- [47] S. Zhang, A. Agathos, and J. Li, "Robust minimum volume simplex analysis for hyperspectral unmixing," *IEEE Trans. Geosci. Remote Sens.*, vol. 55, no. 11, pp. 6431–6439, Nov. 2017.
- [48] B. Rasti *et al.*, "Feature extraction for hyperspectral imagery: The evolution from shallow to deep (overview and toolbox)," *IEEE Geosci. Remote Sens. Mag.*, vol. 8, no. 4, pp. 60–88, Dec. 2020.
- [49] J. Li *et al.*, "Multiple feature learning for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 3, pp. 1592–1606, Mar. 2015.
- [50] H. Lee and H. Kwon, "Going deeper with contextual CNN for hyperspectral image classification," *IEEE Trans. Image Process.*, vol. 26, no. 10, pp. 4843–4855, Oct. 2017.

- [51] D. Hong, L. Gao, J. Yao, B. Zhang, A. Plaza, and J. Chanussot, "Graph convolutional networks for hyperspectral image classification," *IEEE Trans. Geosci. Remote Sens.*, vol. 59, no. 7, pp. 5966–5978, Jul. 2021.
- [52] T. Luo *et al.*, "Active learning to recognize multiple types of plankton," in *Proc. ICPR*, vol. 3, Aug. 2004, pp. 478–481.
- [53] M. Ahmad, S. Shabbir, D. Oliva, M. Mazzara, and S. Distefano, "Spatial-prior generalized fuzziness extreme learning machine autoencoder-based active learning for hyperspectral image classification," *Optik*, vol. 206, Mar. 2020, Art. no. 163712.
- [54] M. Ahmad, A. M. Khan, M. Mazzara, and S. Distefano, "Multi-layer extreme learning machine-based autoencoder for hyperspectral image classification," in *Proc. 14th Int. Joint Conf. Comput. Vis., Imag. Comput. Graph. Theory Appl. (VISIGRAPP)*, 2019, pp. 75–82.
- [55] X. Cao, J. Yao, Z. Xu, and D. Meng, "Hyperspectral image classification with convolutional neural network and active learning," *IEEE Trans. Geosci. Remote Sens.*, vol. 58, no. 7, pp. 4604–4616, Jul. 2020.



Qingqing Zhao received the B.S. degree from the Guangzhou University of Computer Science and Technology, Guangzhou, China, in 2018. She is pursuing the M.Sc. degree in engineering with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China.

Her research interests include hyperspectral image classification, pattern recognition, and active learning.



Meng Xu (Member, IEEE) received the B.S. and M.E. degrees in electrical engineering from the Ocean University of China, Qingdao, China, in 2011 and 2013, respectively, and the Ph.D. degree from the University of New South Wales, Canberra, ACT, Australia, in 2017.

She is an Assistant Professor with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China. Her research interests include cloud removal and remote sensing image processing.



Sen Jia (Senior Member, IEEE) received the B.E. and Ph.D. degrees from the College of Computer Science, Zhejiang University, Hangzhou, China, in 2002 and 2007, respectively.

Since 2008, he has been with the College of Computer Science and Software Engineering, Shenzhen University, Shenzhen, China, where he is a Full Professor. His research interests include hyperspectral image processing, signal and image processing, and machine learning.