Feature Selection for Cross-Scene Hyperspectral Image Classification via Improved Ant Colony Optimization Algorithm

YOUHUA YU, XIAOLAN XIE, (Member, IEEE), YIGANG TANG, AND YARONG LIU

1College of Information Science and Engineering, Guilin University of Technology, Guilin 541000, China
2Guangxi Key Laboratory of Embedded Technology and Intelligent System, Guilin 541000, China
Corresponding author: Xiaolan Xie (xie_xiao_lan@foxmail.com)

This work was supported in part by the National Natural Science Foundation of China under Grant 62262011, in part by the Science and Technology Major Project of Guangxi Province under Grant AA19046004, and in part by the Natural Science Foundation of Guangxi under Grant 2021JJA170130.

ABSTRACT  Hyperspectral images (HSIs) generally contain a large amount of spectral bands (features), and the redundant information in them will cause the Hughes phenomenon in the classification process. And feature extraction and feature selection are the two main existing methods to effectively reduce the redundancy of spectral information in the field of HSIs classification. Compared with feature extraction methods, feature selection methods can preserve most of the features of the original HSIs data without losing their valuable details. However, most existing feature selection methods based on single scene (domain) perform poorly in some scenes (domains) with insufficient labeled samples. Therefore, how to adopt an efficient feature selection method to select the optimal feature subsets of source scene and target scene and use the sample information of source scene to assist in the classification of target scene so as to improve the classification accuracy of images in the target scene as much as possible is still very challenging.

In order to solve the above problem, this paper proposes a new cross-scene algorithm: Improved Ant Colony Optimization Algorithm-Based Cross-Scene Feature Selection Algorithm (IMACO-CSFS). In order to obtain more accurate feature subsets of the two scenes, IMACO-CSFS proposes a priority sorting-based ant colony strategy to make the subsequent search process focus on the global optimal solution (optimal feature subset) found in the previous iteration. In addition, in order to further accelerate the convergence speed of the global optimal solution, an ant colony strategy based on elite ants is proposed in IMACO-CSFS to more efficiently obtain the optimal feature subsets of the two scenes for training the classifier. Furthermore, this paper simultaneously considers overall classification accuracies of the optimal feature subsets for both scenes and dynamically adjusts their scale to ensure the consistency of the selected features between the two scenes, attenuating the effect of spectral shift and achieving the higher image classification accuracy in the target scene. Experimental results on three cross-scene HSI data pairs demonstrate that IMACO-CSFS is superior in cross-scene feature selection and cross-scene HSIs classification.

INDEX TERMS  Cross-scene IMACO-FS, hyperspectral image, cross-scene feature selection, algorithm optimization.

I. INTRODUCTION

HYPERSPECTRAL images (HSIs) typically comprise dozens or hundreds of spectral bands, which contain a lot of spatial and spectral information about ground objects [1], so they have been widely used in agriculture, forestry, environment, ecology and other fields [2], [3]. However, problems of redundant features and high data dimensions still exist in HSIs [4], which may cause Hughes phenomenon in the process of HSIs classification (Hughes phenomenon refers to the phenomenon that the performance of the classifier “increases first and then decreases” with the increase of the feature/band dimension involved in the operation, during the analysis and processing of HSIs). Therefore,
dimensionality reduction must be carried out to better analyze and classify HSIs [5], [6]. Feature extraction and feature selection are two popular technologies for data reduction and elimination of irrelevant features such as redundancy and noise [7]. At present, the mainstream feature extraction methods include unsupervised learning principal component analysis (PCA, also known as KL (Karhunen-Loève) transform) and supervised learning linear discriminant analysis (LDA) [8], [9], [10], etc (et cetera). These methods compress and extract data by using complex mathematical transformations. Although they can extract useful features required from HSIs datasets, it is easy to lose the original physical meaning that can be explained in the data during extraction [11]. In contrast, feature selection method is increasingly favored by researchers because it can maintain the original physical meaning of hyperspectral data [12]. According to the types of learning models, feature selection methods can be divided into three categories: embedded method, filter method and wrapper method [13]. Embedded feature selection method is to carry out feature selection and classification simultaneously in the classifier model, typical embedded methods include weighted Gini index [14] and Sparse Rescaled Linear Square regression feature selection [15]. Filter feature selection method typically adopts a measure based on the distance or information between samples as the evaluation standard and obtains the required feature subset by filtering, thereby reducing computational complexity. Common filter methods include I-RelieFF (Iterative-RelieFF; ReliefF is an improved version of Relief) [16], [17], mutual information [18], information gain [19], etc (et cetera). The wrapper feature selection method is mainly used to search for feature subsets suitable for classifiers and to judge the effectiveness of the selected feature subsets through classifiers such as k-nearest neighbors (KNN) [20] and support vector machine (SVM) [21]. Common wrapper methods include Mixed Whale Optimization Algorithm with Simulated Annealing (WOASA) [22], modified ant lion optimizer [23], etc.

In actual HSIs classification tasks, some scenes are required to use other scenes with similar tag classes for auxiliary classification due to insufficient number of tag samples or unsatisfactory classification effect in a single scene, which is called cross-scene classification [24]. In the cross-scene classification task, the similar marker data in the source scene can be used as training samples to classify the newly acquired target scene lacking marker samples, in order to improve the classification performance of the target scene by transferring the similar information of the source scene [25]. However, due to the large differences between the two scenes in lighting, atmosphere, seasonal influence and adjacency effect, the spectral characteristics of ground objects of the same kind may be different in source scene and target scene, which is the phenomenon of spectral offset in cross-scene classification task [26]. If the information of the source scene is directly used for auxiliary classification of the target scene, the problem of spectral offset will basically be encountered because of the different material composition of the same type of ground object in different time and size and the different nonlinearity of sensors to obtain the data of the two scenes. Therefore, it is particularly important to reduce spectral offset. At present, some methods such as cross-perspective learning [27] and transfer learning [28] have been proposed to solve such problems. For example, Literature [29] calculates the distance of conditional distribution between two scenes in a regenerated kernel hilbert space based on a new dataset shift measure. Based on deep adaptive network (DAN), multi-kernel maximum mean discrepancies (MK-MMD) is used in [30] to align the deep features of source and target scenes. In [31], I-RelieFF (Iterative-RelieFF) was applied in the field of cross-scene classification, and the experiments on the newly proposed CDIRF (Cross-Domain Iterative ReliefF) method on three cross-scene datasets validate its superiority in cross-scene feature selection and reduction of spectral shift between two scenes.

Ant colony optimization algorithm (ACO) is a classical heuristic random search algorithm inspired by nature, because its positive feedback mechanism can help ant colony find the advantage of optimal solution in a short time, ACO has been widely used in image processing, feature/band selection for HSIs data and other fields [32], [33]. However, it is not suitable to be directly used to solve the problem of feature/band selection due to its shortcomings such as slow convergence rate and significant decrease of population diversity in the process of cyclic iteration [34], [35]. Just as the Improved Ant Colony Optimization Algorithm-Based Band Selection Algorithm (IMACO-BS) newly proposed in literature [34] is based on ACO and uses a new pre-filter to reduce the number of candidate bands(n), accelerating the convergence of the standard ant colony optimization algorithm. In addition, the adaptive information updating strategy was introduced in IMACO-BS to avoid the ant colony falling into local optimum. Although it has the advantages of quick ant colony convergence speed and difficulty in falling into local optimal solution, this method is only suitable for the selection of the optimal band subset of a single scene. With the development of HSIs classification technology, it is becoming more and more important to use the information of two scenes to deal with the problem of high-dimensional cross-scene HSIs classification. As Zhang et al. [31] extended the iterative ReliefF method to a cross-scene HSIs classification method (CDIRF), CDIRF (Cross-Domain Iterative ReliefF) extracts effective features from redundant spectral bands for subset evaluation, improving the the classification accuracy of the target scene. Therefore, IMACO-BS is applied and improved in this paper to solve the problem of feature selection and image classification of cross-scene HSIs, and the new algorithm is called cross-scene IMACO-FS (IMACO-CSFS).

At present, there are also some related works that apply ACO to the HSIs classification and feature selection, and some achievements have been achieved. For example: Literature [36] proposes a remote sensing image classification...
technology based on the optimal support vector machine and modified binary encoded ant colony optimization algorithm, which mainly focuses on the classification accuracy of the optimal SVM and the optimization ability of the modified binary encoded ant colony optimization algorithm. The initial pheromone density in this method is set according to the binary string generated by the Genetic Algorithm. Reference [36] simultaneously considers the two mutually influencing problems of SVM parameter adjustment and selection of optimal feature subsets (both are also combinatorial optimization problems) to further improve the performance of the classifier and the classification accuracy of HSIs. Literature [37] proposed a feature selection and classification method for hyperspectral remote sensing images based on ant colony optimization algorithm. It first randomly projects all features onto a plane, and let each ant (instance object) randomly select a feature on the plane; then let the ants decide to choose an appropriate path according to the standard function between the features to form the combined features; finally, the combined features are utilized and the HSIs are classified by a maximum likelihood classifier. The experiment also verifies that the method can finally obtain a better global optimal solution than the sequential floating forward selection algorithm (SFFS), and is more suitable for the selection and classification of hyperspectral images. Literature [38] uses the ant colony optimization algorithm to classify hyperspectral images while performing feature selection and SVM parameter determination, so that it can jump out of the local optimum in the high-dimensional feature space, and then obtain a near-global optimal solution. Because the method proposed in [38] simultaneously optimizes the SVM parameters and a subset of input features, it has better performance than other traditional optimization algorithms such as simulated annealing algorithm and tabu search algorithm in reducing the size of the selected feature subset and improving the classification accuracy of HSIs. Literature [39] proposes a two-stage hybrid ACO algorithm for high-dimensional feature selection, which uses the interval strategy to determine the OFS size of the optimal feature subset (OFS) search, and detects the performance of a part of the feature number endpoints in advance to reduce the complexity of the algorithm and avoid the algorithm from falling into a local optimum. In addition, the algorithm embeds a mixture model that uses the feature’s inherent correlation property to guide OFS search. The experimental results show that the algorithm is more suitable for high-dimensional feature selection and has a shorter running time than other ACO-based feature selection methods. Its next step will also consider how to effectively combine deep learning with itself.

And the main contributions of the IMACO-CSFS algorithm proposed in this paper include the following points:

1) In order to make the subsequent search process focus on the global optimal solution (optimal feature subset) found in the previous iteration, this paper proposes a priority sorting-based ant colony strategy so that the subsequent search process can focus on the global optimal solution found so far. The strategy can obtain more accurate feature subsets of source and target scenes than other cross-scene feature selection methods;

2) In view of the shortcomings of the original IMACO-FS that the optimal solution converges too slowly in each iteration process, this paper proposes an ant colony strategy based on elite ants: in order to make the current optimal solution more attractive to the ants in the next iteration, this strategy gives the optimal solution an extra pheromone after each iteration. The ant colony strategy based on elite ants can further accelerate the convergence speed of the global optimal solution of IMACO-FS, so as to obtain the optimal feature subsets in the source scene and the target scene more effectively for training the classifier;

3) In this paper, the feature selection method IMACO-FS based only on a single scene is successfully applied to the fields of cross-scene HSIs feature selection and cross-scene image classification, which further improves the update strategy of pheromone in the ant colony algorithm, weakens the influence of spectral drift on cross-scene image classification, and greatly improves the feature selection accuracy and image classification accuracy of the target scene.

The rest of this paper is as follows: Section II introduces basic knowledge of ACO and IMACO-FS. Section III introduces a cross-scene feature optimal selection algorithm, IMACO-CSFS, which extends the original single-scene-based IMACO-FS algorithm to make it suitable for cross-scene feature selection. In Section IV, several algorithms are compared on three cross-scene HSIs data pairs and the experimental results are analyzed, proving the superiority of the proposed IMACO-CSFS algorithm. Finally, Section V summarizes the work of this paper.

II. RELATED WORKS

A. ANT COLONY OPTIMIZATION ALGORITHM (ACO)

ACO is a heuristic global optimization algorithm used to find the optimal path. It was first proposed by Marco Dorigo in his doctoral dissertation in 1992. This algorithm has the characteristics of distributed computing, positive feedback of information and heuristic search [49]. ACO imitates the behavior of ants looking for the optimal path in the process of finding food. This principle of finding the optimal path is shown in Fig. 1. While ants travel, they leave behind a volatile secretion called pheromone. Ants can sense the presence of this substance during foraging and walk along paths with higher pheromone concentrations. And each passing ant continues to leave pheromones on the way, forming a mechanism similar to positive feedback. Through this mechanism, ants can finally find the best action path [50].

B. IMPROVED ANT COLONY OPTIMIZATION ALGORITHM-BASED FEATURE SELECTION ALGORITHM (IMACO-FS)

In IMACO-FS, the set \( A = \{A_1, A_2, \ldots, A_n\} \) represents \( n \) graph nodes (a node represents a HSI feature),
E = \{A_iA_j \} (A_i, A_j \in A) represents the edge connecting each HSI feature, \( V_A \) (a subset of A) = \{A_1, A_2, \ldots, A_b \} (b < n) represents a subset of features selected by the ants. Each ant (instance object \( k, k = 1, 2, \ldots, P \)) is allocated memory, which records some control parameters (pheromone \( \alpha \), heuristic information \( \beta \), pheromone volatilization rate \( \rho \), initial value \( Q \) of pheromone secreted by ants (constant), etc.). Besides, a tabu table \( \Gamma \) is also created in memory to store the features that the ants have been visited.

IMACO-FS algorithm first adopts a new pre-filtering method 1), taking the original heuristic expectation \( \eta_{ij}(t) \) (equation (1)) between each pair of nodes \( A_i, A_j \) as a foundation, select the top \( o (o < n) \) features with the highest OA (the overall accuracy) to form the candidate feature subset \( F^k \), reducing the number of candidate features \( n \) and accelerating the convergence of the standard ant colony algorithm. Then IMACO-FS iteratively searches the feature subset \( F^k \). In each iteration, the ants use the correlation information between the selected feature and the unselected feature to perform state transition according to the pseudo-random rule 2) and select the next unvisited feature node in the subset \( F^k \) until m nodes are added to the feature subset \( M(m < o) \). Finally, the OA of \( M \) is calculated by SVM. After one cycle (iteration) ends, the pheromone concentration on the path is updated according to the update strategy 3). Then each ant continues to iterate according to the same method as in the last cycle and selects or updates the optimal feature subset \( M \) that has been traversed until the maximum number of iterations \( I_M \) is reached, at which time the feature subset \( M \) with the optimal OA is output.

The pre-filtering method, pseudo-random rule and pheromone update strategy in IMACO-FS are described in detail below:

1) PRE-FILTERING METHOD

Initialize the pheromone on each path based on the original heuristic expectation \( \eta_{ij}(t) \) (1) between each pair of nodes \( A_i \) and \( A_j \), and the OA between feature \( i \) and other \( n-1 \) features is calculated by SVM, and the features with the highest OA among the first \( o (o < n) \) features are selected to form feature subset \( F^k \)(the value of \( o \) can be determined according to the statistical results [40], which is set to \( n/2 \) in this paper).

The calculation method of the heuristic expectation \( \eta_{ij}(t) \) is shown in formula (1), where \( O_{ij} \) is the OA calculated by the SVM with feature \( i \) and feature \( j \).

\[
\eta_{ij}(t) = O_{ij}
\]  

The pheromone calculation formula between feature \( i \) and other \( n-1 \) features is shown in formula (2).

\[
\tau_{ij}(0) = \begin{cases} \frac{O_{ij}}{O_{\text{max}}}, & j \in F^k \\ \frac{O_{\min}}{O_{\text{max}}}, & \text{otherwise} \end{cases} \]

where \( O_{ij} \) is the OA achieved by SVM with feature \( i \) and feature \( j \), \( O_{\text{max}} \) and \( O_{\min} \) are the maximum and minimum OA achieved by SVM between feature \( i \) and the remaining \( n-1 \) features.

2) PSEUDO-RANDOM RULE

The transformation mode of ants’ states is as follows:

\[
S = \begin{cases} \text{arg max}\{\tau_{ij}(t)\}, & q < q_0 \\ p_{ij}^0(t), & q \geq q_0 \end{cases}
\]

where \( q_0 = 1 - e^{-1/s} \) (\( s = 1, 2, \ldots, I_M \), \( s \) is the number of iterations), \( q \) is a random decimal between 0 and 1, \( p_{ij}^0(t) \) represents the probability that ant \( k \) selects the next feature node \( j \) at time \( t \), which can be calculated as follows:

\[
p_{ij}^k(t) = \begin{cases} \frac{\tau_{ij}(t)^{\alpha \beta/T}}{\sum_{j \in F^k} \tau_{ij}(t)^{\alpha \beta/T}}, & j \in F^k \\ 0, & \text{otherwise} \end{cases}
\]

where \( \alpha \) is the pheromone, \( \beta \) is the heuristic information, \( T \) is the sum of \( I_M \) iteration times, \( \tau_{ij}(t) \) and \( \eta_{ij}(t) \) represent the pheromone concentration and heuristic expectation on the edge \( (A_i, A_j) \) at time \( t \), respectively, and tabu table \( \Gamma \) stores the feature nodes that ant \( k \) has visited.

3) UPDATE STRATEGY FOR PHEROMONES

After each ant completes one cycle (iteration), the pheromone concentration on the path is updated as follows:

\[
\tau_{ij}(t+1) = \begin{cases} (1-\rho) \cdot \tau_{ij}(t) + \sum_{k=1}^{P} \Delta \tau_{ij}^k, & j \in F^k \\ (1-\rho) \cdot \tau_{ij}(t), & \text{otherwise} \end{cases}
\]

where \( \tau_{ij}(t+1) \) and \( \tau_{ij}(t) \) respectively represent the pheromone concentration on the edge \( (A_i, A_j) \) at time \( t+1 \) and time \( t \), and \( \Delta \tau_{ij}^k \) is the pheromone newly left by ant \( k \) on the edge \( (A_i, A_j) \), the calculation method is as follows:

\[
\Delta \tau_{ij}^k = \begin{cases} (Q \cdot \frac{O_{ij}}{O_{\text{max}}})^2 + 1, & j \in F^k \\ 0, & \text{otherwise} \end{cases}
\]

where \( Q \) is the initial value of the pheromone secreted by the ants at the beginning of an iteration, which is a constant, \( O_{\text{max}} \) is the maximum OA achieved by the SVM between feature \( i \) and other \( o-1 \) features in feature subset \( F^k \), \( t \) represents time \( t \), and \( T \) is the sum of \( I_M \) iteration times.
Y. Yu et al.: Feature Selection for Cross-Scene HSI Classification via Improved Ant Colony Optimization Algorithm

III. PROPOSED APPROACH

The common HSIs feature selection method generally only selects the optimal feature based on a single scene. The IMACO-FS method based only on the target scene (single scene) can solve such problems, but it cannot solve the problem of feature selection and classification across scenes. Therefore, this paper introduces the cross-scene IMACO-FS method, the newly proposed method is named Improved Ant Colony Optimization Algorithm-Based Cross-Scene Feature Selection Algorithm (IMACO-CSFS). Basic framework of the proposed IMACO-CSFS for cross-scene HSIs classification is shown in Fig. 2.

In IMACO-CSFS, all initial information carried by each ant is the same as that of ants in IMACO-FS. However, the same as Ant Colony Optimization Algorithm, IMACO-FS has a disadvantage: in the evolution process, because the total quality of the solution increases and the difference between the solution elements decreases, the difference in the selection probability among ants decreases, which makes the subsequent search process not focus on the optimal solution \( \delta \) (optimal feature subset \( M \)) found so far, thus preventing further search for a better solution \( \delta^* \). Inspired by the use of sorting selection mechanism in genetic algorithm to solve the selection pressure [41], [42], IMACO-CSFS proposes a priority sorting-based ant colony strategy, which can make the subsequent search process focus on \( \delta \) found so far. When all ants generate their own path, the ant colony strategy sorts the paths according to their length \( (L_1 \leq L_2 \leq \ldots \leq L_P) \).

The optimal path found by ant \( k \) so far will obtain additional pheromone \( \Delta \tau_{ij}^* \), and the edges passed by the remaining \( P - 1 \) ants can also obtain a certain amount of pheromone \( \Delta \tau_{ij}^k \) \((k = 1, 2, \ldots, P - 1)\). The ant colony strategy based on priority sorting can further search around the current optimal solution \( \delta \), thus obtaining the more accurate optimal feature subsets \( M_s \) and \( M_t \) in the source and target scenes.

In addition, since original IMACO-FS has the disadvantage that the convergence speed of optimal solution \( \delta \) is too slow in each iteration process, this paper proposes an ant colony strategy based on elite ants: in order to make the current optimal solution \( \delta \) more attractive to ants in the next cycle, this strategy will give the optimal solution an additional pheromone \( \Delta \tau_{ij}^{**} \) after each iteration, and the optimal solution \( \delta^* \) is completed after \( I_M \) iterations is the global optimal solution, and the ant that finds this solution is an elite ant. The elite ants-based ant colony strategy further accelerates the convergence speed of the global optimal solution \( \delta^* \) in the original IMACO-FS, thereby obtaining the optimal feature subsets \( M_s \) and \( M_t \) in the source and target scenes more efficiently.

The newly proposed pheromone update formula can be calculated as follows:

\[
\tau_{ij}(t + 1) = \begin{cases} 
(1 - \rho) \cdot \tau_{ij}(t) + \sum_{k=1}^{P-1} \Delta \tau_{ij}^k + \Delta \tau_{ij}^* \\
+ \Delta \tau_{ij}^{**}, & j \in F^k \\
(1 - \rho) \cdot \tau_{ij}(t), & \text{otherwise} 
\end{cases}
\]  

\[
(7)
\]
where \( \sum_{k=1}^{P-1} \Delta \tau_{ij}^k \) is the update of the pheromone content on the path by the \( P - 1 \) ants according to their rank \( k \), the calculation method of \( \Delta \tau_{ij}^k \) is shown in formula (8), the calculation method of \( \Delta \tau_{ij}^* \) is shown in formula (9), and the calculation method of \( \Delta \tau_{ij}^{**} \) is shown in formula (10).

\[
\Delta \tau_{ij}^k = \begin{cases} 
\frac{Q(P-k)}{L_k}, & \text{if } j \in F^k \\
0, & \text{otherwise}
\end{cases} \quad (8)
\]

\[
\Delta \tau_{ij}^* = \begin{cases} 
\frac{Q}{L_{best}}, & \text{if } j \in F^k \\
0, & \text{otherwise}
\end{cases} \quad (9)
\]

\[
\Delta \tau_{ij}^{**} = \frac{\delta^k Q}{L_k} \quad (10)
\]

Among them, \( Q \) is the initial value of pheromone secreted by each ant at the beginning of each iteration, \( L_k \) is the path length that ant \( k \) travels in one iteration, \( L_k = \sum_{i=1}^{n-1} \sum_{j=2}^{n} d_{ij} \), \( d_{ij} \) is the length between feature node \( i \) and feature node \( j \), and \( i, j \in \Gamma^k \), the tabu table \( \Gamma^k \) stores the feature nodes visited by ant \( k \), and each feature node is visited only once in an iteration. \( L_{best} \) is the length of the optimal path found by ant \( k \) so far, and \( \delta^k \) represents the optimal solution \( \delta^* \) that ant \( k \) searches for in one iteration.

In order to decrease the influence of spectral shift and improve the classification accuracy of target scene, IMACO-CSFS considers two aspects at the same time: 1) The consistency of selected feature between source scene and target scene, that is, the connection between the respective optimal feature subsets of the two scenes, and its accuracy is represented by \( \mu O_s + (1 - \mu) O_t \); where \( O_s \) represents the OA of optimal feature subset \( M_s \) in the source scene, \( O_t \) represents the OA of optimal feature subset \( M_t \) in the target scene, and \( \mu \) is a parameter used to adjust the ratio between \( O_s \) and \( O_t \); 2) The discriminability of different types of ground objects in the target scene, and its accuracy is represented by \( D_t \).

Therefore, the newly proposed objective function \( \delta^* \) (the optimal solution) in IMACO-CSFS comprehensively considers the three accuracies of \( O_s \), \( O_t \) and \( D_t \), and \( \delta^* \) can be defined as:

\[
\delta^* = D_t + \mu O_s + (1 - \mu) O_t \quad (11)
\]

where \( O_s \) and \( O_t \) are the OAs calculated by SVM with the optimal feature subset \( M_s \) in the source scene and the optimal feature subset \( M_t \) in the target scene, respectively. The initial value of \( \mu \) is selected as a decimal between the interval \((0,1)\), initialized according to the ratio of the number of labeled samples in the source scene and the target scene. If the number of labeled samples in the source scene is less that in the target scene, \( \mu \in (0.5, 1) \), otherwise \( \mu \in (0, 0.5) \), and the subsequent experiments will iteratively optimize \( \mu \) (adaptive adjustment). The calculation method of \( D_t \) is shown in formula (12):

\[
D_t = \frac{1}{I_M} \sum_{i=1}^{I_M} A_i O_t
\]

where \( I_M \) is the number of iterations, \( m \) is the dimension of the optimal feature subset \( M_t \) of the target scene, \( A_i \) is the average classification accuracy (AA) of \( M_t \), \( O_t \) is the overall classification accuracy (OA) of \( M_t \), and \( O_t \) is the overall classification accuracy (OA) of \( M_t \).

The pseudo code of calculation process of \( \delta^* \) is shown in Algorithm 1: three classifiers are trained and the value of the objective function \( \delta^* \) is calculated. According to definition (11), a larger value of \( \delta^* \) function will bring a better classification effect of target scene images. The training and testing of \( D_t \), \( O_s \) and \( O_t \) are carried out according to the

**Algorithm 1 The Calculation of Objective Function in IMACO-CSFS**

**Input:**
- The labeled samples of source scene \( T_s \).
- The labeled samples of target scene \( T_t \).

**Output:**
- Objective function \( \delta^* \). Simulation steps are as follows:

1. Train a classifier using randomly selected 1/10 of the target scene labeled samples and calculate \( D_t \) by using equation (12) (see Table 1).
2. Train a classifier using randomly selected 1/10 the number of labeled samples from the source scene and calculate \( O_s \) (see Table 1).
3. Train a classifier using randomly selected 1/10 the number of labeled samples from the target scene and calculate \( O_t \) (see Table 1).
4. Use (11) to compute \( \delta^* \).
5. return \( \delta^* \)

**FIGURE 3. Flow chart of IMACO-CSFS.**
division method of the number of labeled samples in Table 1. Note that the final classifier is trained by 1/10 of the labeled samples (training samples) from the source scene and the target scene, and tested on the remaining labeled samples (test samples) in the target scene to obtain the final classification result. The flowchart of IMACO-CSFS is shown in Fig. 3 and the pseudo code of IMACO-CSFS is shown in Algorithm 2.

IV. EXPERIMENTS

A. DATASETS

To test the effectiveness of our proposed method (IMACO-CSFS), we conduct experiments on three representative cross-scene HSI data pairs. The first data pair is Huston 2018 (source scene) and Huston 2013 (target scene), both acquired in different years by different sensors at Huston University and its vicinity. The former includes 48 spectral bands with an image size of 955 × 209 pixels, and the latter includes 144 spectral bands with an image size of 1905 × 349 pixels. Although their pixel sizes and number of bands are inconsistent, they have 7 same land cover classes, and the quantity distribution of their labeled samples is shown in Table 2, which is very suitable for the experiment for classification of cross-scene HSIs. For better comparison, we select Huston 2018 images and Huston 2013 images with a size of 935 × 209 pixels in our experiments. Fig. 4 shows their false color map and ground truth map (the color of each type of ground object is marked next to the ground truth map).

The second data pair is University of Pavia (source scene) and Center of Pavia (target scene). The Pavia dataset was acquired by the DAIS hyperspectral sensor. The size of the PaviaU (University of Pavia) image is 243 × 243 pixels and the number of bands is 72, while the size of the PaviaC (Center of Pavia) image is 400 × 400 pixels and the number of bands is also 72. Although they have different pixels, they have 6 identical land cover classes, and the number distribution of their labeled samples is shown in Table 3, which is in line with the experimental requirements of cross-scene HSIs classification. The false color map and the ground truth map of PaviaU and PaviaC are shown in Fig. 5.

The third HSI data pair is Shanghai (source scene) and Hangzhou (target scene), both of which are obtained by EO-1 Hyperion sensor. Both Shanghai and Hangzhou have

### Table 1: Calculation of \( D_t, O_s \) and \( O_t \).

| \( D_t \) | Randomly selected 1/10 of labeled samples from the target scene |
| --- | --- |
| \( O_s \) | Randomly selected 1/10 of labeled samples from the source scene |
| \( O_t \) | Randomly selected 1/10 of labeled samples from the target scene |

### Table 2: Number of labeled samples for Huston data pair.

| Class | Huston 2018 (Source) | Huston 2013 (Target) |
| --- | --- | --- |
| Trees | 2766 | 363 |
| Road | 6365 | 443 |
| Residential building | 5347 | 319 |
| Water | 22 | 285 |
| Grass stressed | 4888 | 365 |
| Non-residential buildings | 32459 | 408 |
| Grass healthy | 1353 | 345 |
| Total | 53200 | 2530 |

### Algorithm 2 IMACO-CSFS

**Input:**
- Parameter \( \mu \),
- Pheromone \( \alpha \),
- Number of iterations \( I_M \),
- The labeled samples of target scene \( T_t \),
- The labeled samples of source scene \( T_s \).

**Output:**
- Objective function \( \delta^* \).

1. Initialize pheromone \( \alpha \) using (1).
2. Place \( P \) ants randomly on the graph nodes.
3. for ant \( k \in [1, P] \) do
4. Calculate \( OA \) between feature \( i \) and the remaining \( n - 1 \) features in \( T_s \).
5. Select the first \( o \) features with the highest \( OAs \) to add to feature subset \( F_s \).
6. Use the same method as above to get \( F_t \) in \( T_t \).
7. end for
8. while (Number of iterations \( \neq I_M \)) do
9. for ant \( k \in [1, P] \) do
10. for \( i = 1, 2, \ldots, o \) do
11. Search feature subset \( F_s \) and \( F_t \).
12. for \( j = 1, 2, \ldots, m \) do
13. select a feature from \( F_s \) to add it to \( M_s \) and select a feature from \( F_t \) to add it to \( M_t \).
14. end for
15. Calculate \( M_s' \)s OA (\( O_s \)) and \( M_t' \)s OA (\( O_t \)) according to Algorithm 1.
16. Once a better \( O_s \) appears, update \( M_s \).
17. Once a better \( O_t \) appears, update \( M_t \).
18. Use \( O_s \), \( O_t \), and \( \mu \) to calculate \( \mu O_s + (1 - \mu)O_t \).
19. Once a better \( \mu O_s + (1 - \mu)O_t \) appears, update \( \mu \).
20. end for
21. end for
22. Update \( \alpha \) using (7).
23. Reset the \( P \) ants.
24. end while
25. Calculate \( M_t' \)s AA (\( A_t \)).
26. Calculate \( D_t \) using (12) by applying Algorithm 1.
27. Calculate \( \delta^* \) using (11).
28. return \( \delta^* \).
Y. Yu et al.: Feature Selection for Cross-Scene HSI Classification via Improved Ant Colony Optimization Algorithm

FIGURE 4. False-color image and ground truth of Huston 2018-Huston 2013 data pair. (a) False-color image of Huston 2018 (source scene). (b) False-color image of Huston 2013 (target scene). (c) Ground truth of Huston 2018 (source scene). (d) Ground truth of Huston 2013 (target scene).

TABLE 3. Number of labeled samples for Pavia data pair.

| Class     | PaviaU(Source) | PaviaC(Target) |
|-----------|----------------|----------------|
| 1 Trees   | 266            | 2424           |
| 2 Meadow  | 273            | 1251           |
| 3 Asphalt | 266            | 1704           |
| 4 Bitumen | 206            | 685            |
| 5 Parking lot | 265        | 287            |
| 6 Soil    | 213            | 1475           |
| **Total** | **1489**       | **7826**       |

TABLE 4. Number of labeled samples for Shanghai-Hangzhou data pair.

| Class     | Shanghai (Source) | Hangzhou (Target) |
|-----------|-------------------|-------------------|
| 1 Plant   | 83188             | 40207             |
| 2 Land/Building | 161689         | 77450             |
| 3 Water   | 123123            | 18043             |
| **Total** | **368000**        | **138700**        |

FIGURE 5. False-color image and ground truth of PaviaU-PaviaC data pair. (a) False-color image. (b) Ground truth. The upper one is PaviaU (source scene), while the lower one is PaviaC (target scene).

198 spectral bands, consisting of 1600 × 230 pixels and 590 × 230 pixels, respectively. Although their pixel sizes are inconsistent, they share 3 common land cover classes in total, and the number distribution of their labeled samples is shown in Table 4, which is very helpful for experiments on classification of cross-scene HSIs. The false color map and ground truth map of Shanghai and Hangzhou are shown in Fig. 6.

B. ALGORITHMS COMPARISON AND PARAMETER SETTINGS

On these data pairs, we compare the original IMACO-FS algorithm, CDIRF, CDWOASA (Cross-Domain Hybrid Whale Optimization Algorithm With Simulated Annealing), SS-DDNMF (Semisupervised Dual-Dictionary Non-Negative Matrix Factorization), SFA-MSSN (Spectral Feature Adaptation–Multiscale Spectral-Spatial Unified Network), PSO-OBFS (Particle Swarm Optimization for Object-Based Feature Selection) and DRB-RESNET (Diverse Region Dual-Branch–Deep Residual Network) with the newly proposed IMACO-CSFS method to verify IMACO-CSFS’s superiority. Their details and parameter settings are as follows:

1) CDIRF (CROSS-DOMAIN ITERATIVE ReliefF)

It performs I-ReliefF on the source scene and the target scene respectively. CDIRF proposes a cross-scene feature weight update rule, which uses two distance measures in experiments. They are CDIRF1 using the absolute distance and CDIRF2 using the squared Euclidean distance, and the latter CDIRF2 performs better in the experiments as a whole. To verify the performance of feature selection, CDIRF chooses SVM with radial basis function (RBF) as the classifier. The parameters $\sigma \in \{0.1, 0.2, \ldots, 1, 1.5, 2, 10, 100\}$, $T = 100$ are set in CDIRF ($\sigma$ is the kernel width and $T$ is the number of iterations). In addition, CDIRF analyzes the influence of the parameter $\sigma$ on the classification accuracy of the target scene under different feature dimensions [31].

2) CDWOASA (CROSS-DOMAIN WOASA)

It performs the hybrid whale optimization algorithm with simulated annealing (WOASA) on the source and target scenes, respectively. CDWOASA selects feature subsets that are discriminative and scene-invariant by using information
from two HSI scenes simultaneously, and it also proposes a ranking strategy based on the fitness function values, so that the dimension of output features is precisely controlled. The algorithm has two main parameters: the number of iterations M and the population size n. Its M is set to 100 and n is set to 5 [13].

3) SS-DDNMF (SEMISUPERVISED DUAL-DICTIONARY NON-NEGATIVE MATRIX FACTORIZATION)
It is a heterogeneous transfer learning algorithm. SS-DDNMF trains a dictionary each for the source and target scenes, and projects the features of the two scenes into a shared low-dimensional subspace to eliminate the differences between the feature subspaces. SS-DDNMF chooses SVM with RBF as the classifier. The parameter $C \in \{10^{-1}, 10^{0}, \ldots, 10^{4}\}$, $\gamma \in \{2^{-15}, 2^{-14}, \ldots, 2^{-5}\}$ are set in SS-DDNMF ($C$ is the penalty parameter and $\gamma$ is the parameter of the kernel) [43].

4) SFA-MSSN (SPECTRAL FEATURE ADAPTATION–MULTISCALE SPECTRAL-SPATIAL UNIFIED NETWORK)
It reduces the difference of spectral dimension distribution between source scene and target scene through an adaptive method of joint probability distribution and designs a multi-scale spectral space unified network with a dual-branch architecture and a multi-scale bank to fully extract the discriminative features of HSI. In SFA-MSSN, the number of neurons in each FC (fully-connected) layer of the multiscale bank is set to 1024, the two FC layers before soft-max classifier both contains 400 neurons [44].

5) PSO-OBFS (PARTICLE SWARM OPTIMIZATION FOR OBJECT-BASED FEATURE SELECTION)
It is executed on a single scene. It first uses a multi-resolution segmentation method to segment the image; secondly, it extracts object-based feature (OBF) for each object; finally, it uses a combination of particle swarm algorithm and minimum distance classifier to select the optimal OBFs. PSO-OBFS also uses the chaotic random inertia weight which was regarded as one of the best strategies [51]. Furthermore, PSO-OBFS linearly reduces the acceleration coefficient $c_1$ from 2.5 to 0.5 and linearly increases the acceleration coefficient $c_2$ from 0.5 to 2.5 to avoid being trapped in local optima [52].

6) DRB-RESNET (DIVERSE REGION DUAL-BRANCH–DEEP RESIDUAL NETWORK)
It performs DRB-RESNET on a single scene. DRB-RESNET uses an improved deep residual network to extract the spatial and spectral information of the hyperspectral image after dimensionality reduction, and forms a fully connected two-branch feature extraction network. It mainly uses two networks of different depths: the upper network arranges the processed images of the same size into the DRB-ResNet network of a specified depth (11 layers) for training, and sum the feature extraction data of each region; the lower network first reduces the dimension of the spectral information of each pixel to be classified, and then expands it to $11 \times 11$ pictures, and then inputs it into the DRB-ResNet network (7 layers) for feature extraction. Finally, combined with the feature extraction results of the upper and lower layers, the fully connected network is used to obtain the final classification result [53].

7) IMACO-FS (Feature Select Based on Improved ACO): It selects the optimal feature only based on the labeled samples of a single scene (target scene) to achieve the highest possible overall classification accuracy. A new pre-filtering method is introduced into IMACO-FS to optimize the pheromone initialization of the ant colony system, thereby speeding up the convergence speed of the ant colony system, and IMACO-FS also adopts pseudo-random rules and adaptive information update strategy to maintain the diversity of the ant colony. In IMACO-FS, $P$ (user-defined number) features with the highest $OA$ value including feature $i$ are selected, and $P$ is set to $n/2$ ($n$ is the total number of bands) [40]. IMACO-FS initializes the control parameters...
of ant colony system according to previous research experience: $\beta$ (heuristic information), $\rho$ (pheromone volatilization rate) [45], [46], and the initial values of $\beta$ and $\rho$ are set to 1 and 0.2, respectively. $\beta$ is an integer between 1 and 10, and the step size is 1 each time it’s updated, while $\rho$ is a decimal between 0.1 and 0.9, and the step size of each change is 0.1 [34].

8) Cross-Scene IMACO-FS (IMACO-CSFS): IMACO-CSFS is a newly proposed algorithm in this paper, which is executed on the source scene and the target scene respectively. IMACO-CSFS simultaneously considers the consistency of the optimal feature subsets of the two scenarios and the discriminability of different kinds of ground objects in the target scene. In the experiment, we use SVM with RBF and squared Euclidean distance as the classifier. The maximum number of iterations $I_M$ in equation (12) is set to 100; the ant colony size $n$ on each data pair is set to 5; the ant quantity $P$ of the ant colony is set as about 1.5 times of the number of bands (The initial value of $P$ is set to 0.1 times of the number of bands, and it is gradually increased by 0.1 times until it
TABLE 5. The value of objective function $\delta^s$ on the three Data Pairs.

| Data Pairs          | $\delta^s$ | $D_s$ | $\mu O_s + (1 - \mu)O_t$ | $\mu$ | $O_s$ | $O_t$ |
|---------------------|------------|-------|--------------------------|-------|-------|-------|
| Huston 2018–Huston 2013 | 0.1051     | 0.0000| 0.1051                   | 0.0454| 0.1865| 0.1012|
|                     | 0.1012     | 0.0000| 0.1012                   | 0.0000| 0.1865| 0.1012|
|                     | 0.1439     | 0.0000| 0.1439                   | 0.5000| 0.1865| 0.1012|
|                     | 0.1865     | 0.0000| 0.1865                   | 1.0000| 0.1865| 0.1012|
|                     | 0.8668     | 0.0687| 0.7994                   | 0.0755| 0.7737| 0.8015|
|                     | 0.8331     | 0.0316| 0.8015                   | 0.0000| 0.7737| 0.8015|
|                     | 0.8171     | 0.0295| 0.7876                   | 0.5000| 0.7737| 0.8015|
|                     | 0.7980     | 0.0243| 0.7737                   | 1.0000| 0.7737| 0.8015|
| PaviaU–PaviaC       | 0.1064     | 0.0000| 0.1064                   | 0.8402| 0.1015| 0.1324|
|                     | 0.1324     | 0.0000| 0.1324                   | 0.0000| 0.1015| 0.1324|
|                     | 0.1170     | 0.0000| 0.1170                   | 0.5000| 0.1015| 0.1324|
|                     | 0.1015     | 0.0000| 0.1015                   | 1.0000| 0.1015| 0.1324|
|                     | 0.8363     | 0.1015| 0.7348                   | 0.7567| 0.7195| 0.7823|
|                     | 0.8235     | 0.0412| 0.7823                   | 0.0000| 0.7195| 0.7823|
|                     | 0.7903     | 0.0394| 0.7509                   | 0.5000| 0.7195| 0.7823|
|                     | 0.7571     | 0.0376| 0.7195                   | 1.0000| 0.7195| 0.7823|
| Shanghai–Hangzhou   | 0.1164     | 0.0000| 0.1164                   | 0.2694| 0.1328| 0.1104|
|                     | 0.1104     | 0.0000| 0.1104                   | 0.0000| 0.1328| 0.1104|
|                     | 0.1216     | 0.0000| 0.1216                   | 0.5000| 0.1328| 0.1104|
|                     | 0.1328     | 0.0000| 0.1328                   | 1.0000| 0.1328| 0.1104|
|                     | 0.8627     | 0.0671| 0.7956                   | 0.2810| 0.7411| 0.8169|
|                     | 0.8554     | 0.0385| 0.8169                   | 0.0000| 0.7411| 0.8169|
|                     | 0.8096     | 0.0306| 0.7790                   | 0.5000| 0.7411| 0.8169|
|                     | 0.7709     | 0.0298| 0.7411                   | 1.0000| 0.7411| 0.8169|

The values in bold are the highest accuracy.

Reaches 3.0 times. In the interval of 0.1 times to 1.5 times, the ant colony convergence speed and global optimality of the solution continue to rise; in the interval of 1.5 times to 3.0 times, the convergence speed of the ant colony keeps decreasing, and the global optimality of the solution basically remains unchanged. Finally, combining the convergence speed of the ant colony and global optimality of the solution, it is concluded that 1.5 times the number of bands is the best); if the number of ants is too large, the pheromone concentration on each path tends to average eventually, and the positive feedback effect is weakened, resulting in a slow convergence rate; if it is too small, the concentration of some path pheromones that have never been searched may decrease to 0, which will eventually lead to premature convergence and lower global optimality of the solution [47]. Therefore, the $P$ value of Huston’s group experiment is set to 216 (144 × 1.5); in the Pavia group, $P$ is set to 108 (72 × 1.5); the $P$ of Shanghai group is set to 297 (198 × 1.5). In addition, the initial values of $\beta$ and $\rho$ in the three scenes are set to 1 and 0.2, respectively. The step size of $\beta$ and $\rho$ for each update is the same as that in IMACO-FS.

The initial value of $\mu$ in formula (11) is a decimal selected between the interval (0,1). Because the number of labeled samples of most ground object categories in the source scene of Huston data pair and Shanghai data pair is larger than that in the target scene, the initial value of $\mu$ in the Huston group experiment is set to 0.0454 (Number of target scene labeling samples 2530/Sum of two scene labeling samples 55730); $\mu$ in the Shanghai group is set to 0.2694 (135700/503700); $\mu$ in the Pavia group is set to 0.8402 (78269315), and they are iteratively optimized in the experiment. The tuning process of the key parameter $\mu$ in the objective function $\delta^s$ is shown in Fig. 7 and Fig. 8. In Fig. 7, (a), (b) and (c) are 3D histograms of $\mu O_s + (1 - \mu)O_t$ changing with $O_s$ and $O_t$ in all three experimental data pairs, (a) is the Huston group experiment; (b) is the Pavia group experiment; (c) is the Shanghai group experiment. Fig. 8 shows the curve of $\mu O_s + (1 - \mu)O_t$ versus $\mu$ where $\mu_1 \in$ Huston, $\mu_2 \in$ Pavia, $\mu_3 \in$ Shanghai.

Combining Fig. 7 and Fig. 8, it can be seen that when $\mu_1=0.0454$ (initial value), $O_2=0.1865$, and $O_3=0.1012$, $\mu O_s + (1 - \mu)O_t=0.1051$ in the Huston group experiment, which is at the lowest value. As the number of iterations increases, $\mu O_s + (1 - \mu)O_t$ generally increases with the increase of $\mu_1$, but sometimes it will decrease, mainly because the ratio between $O_s$ and $O_t$ is not coordinated, that is, the value of $\mu_1$ has not been iterated to the optimal value. In the last few iterations, the decryption of $\mu_1$ is gathered around the optimal solution, and finally $\mu O_s + (1 - \mu)O_t$ reaches the maximum value of 0.7994 when $\mu_1=0.0755$, $O_3=0.7737$, $O_2=0.8015$; In the Shanghai group experiment, when $\mu_3=0.2694$ (initial value), $O_2=0.1328$ and $O_3=0.1104$, $\mu O_s + (1 - \mu)O_t$ achieve the minimum value of 0.1164. The change trend of $\mu_3$ is similar to that of $\mu_1$, but its fluctuation is larger, and it is more common for $\mu O_s + (1 - \mu)O_t$ to decrease with the update of $\mu_3$, which may be due to the smaller feature fit of the two scenes of the Shanghai-Hangzhou data pair. In the end, $\mu O_s + (1 - \mu)O_t$ reaches the maximum value of 0.7956 when $\mu_3=0.2810$ (the optimal solution of $\mu_3$), $O_3=0.7411$, and $O_2=0.8169$; In the Pavia group experiment, $\mu O_s + (1 - \mu)O_t$ achieve the minimum value of 0.1064 when $\mu_2=0.8402$ (initial value), $O_2=0.1015$ and $O_3=0.1324$. In the iterative process, $\mu_3=0.2810$ fluctuates greatly, and it basically increases as $\mu_2$ decreases. Finally, $\mu O_s + (1 - \mu)O_t$ reaches...
the maximum value of 0.7348 when $\mu_2=0.7567$, $O_s=0.7195$, and $O_t=0.7823$.

In addition, we also evaluate the effect of extreme values of $\mu$ on the objective function $\delta^*$, and the experimental results
are shown in Table 5. Table 5 shows that the $D_t$ in the first four rows experimental data on these three data pairs are all 0.0000, because the iteration has not yet started and the parameters have just been initialized. In the Huston data pair, although $\mu O_s + (1 - \mu)O_t$ reaches its maximum value of 0.8015 when $\mu$ is 0.0000, $O_s$ is 0.7737 and $O_t$ is 0.8015, the $D_t$ value of 0.0316 is relatively small at this time, resulting in the value of the objective function $\delta^*$, 0.8331, is smaller than the value of $\delta^*$, 0.8681, obtained at the end of the iteration.

In addition, the value of $\delta^*$ obtained when $\mu$ takes 0.5000 or 1.0000 is also smaller than the value of $\delta^*$ obtained after all iterations are completed. And the experimental results on the Pavia and Shanghai data pairs are similar to those on the Huston data pair.

In order to preliminarily attenuate the influence of weather and light on the classification results of the two scenes, we call the FLAASH atmospheric correction command in ENVI software to perform rapid atmospheric correction and analysis on the three data pairs required by the experiment. To reduce the impact on the experimental accuracy caused by the randomly selected training sets and test sets, this paper executes each algorithm 10 times and then averages all the accuracy to ensure the accuracy of the experimental results. After 10 repeated experiments, $\beta$ and $\rho$ of Huston data pair in IMACO-CSFS are updated to 7 and 0.4 respectively; $\beta$ and $\rho$ of Pavia data pair are updated to 4 and 0.6 respectively; $\beta$ and $\rho$ of Shanghai data pair are updated to 6 and 0.5 respectively. In the end, the overall classification accuracy (OA), the average classification accuracy (AA) and the Kappa coefficient ($\kappa$) are used as evaluation criteria to evaluate the accuracy of each method [48]. OA is the overall classification accuracy of all classes in the test sample of a single scene (the ratio of the number of correctly predicted samples to the total number of samples), AA is the average of each class classification accuracy in a single scene, $\kappa$ is the mutual information about the strong consistency between the ground truth map and the classification map for a single scene. The OA, AA and $\kappa$ obtained in this experiment are all the accuracy of the target scene. Note that OA, AA and $\kappa$ are the performance metrics obtained by computer calculation in the experiment, and the tuning process of the parameters $O_s$ ($M_s$'s OA), $O_t$ ($M_t$'s OA) and $A_t$ ($M_t$'s AA) includes a sampling strategy for the parameters (see Algorithm 1, Algorithm 2 and Table 1 for details).

C. HUSTON 2018-HUSTON 2013 EXPERIMENT

Experiment compares the performance of IMACO-CSFS with the original IMACO-FS and CDIRF, CDWOASA, SS-DDNMF, SFA-MSSN, PSO-OBFS and DRB-RESNET under different feature dimensions.

Fig. 9 shows the classification accuracy of eight algorithms under different feature dimensions on the Huston 2013 dataset. Experimental results show that IMACO-CSFS (the newly proposed method) achieves the maximum value of the three accuracies of OA, AA and $\kappa$ when the feature dimension $F_n=40$ is selected in the Huston 2013 dataset, while IMACO-FS performs the worst when $F_n=40$. Furthermore, when the feature dimension is small, such as $F_n=15$, the OA of IMACO-CSFS is even lower than that of CDIRF, while all the accuracies of IMACO-CSFS are basically higher than
those of the other seven algorithms when $F_n \geq 20$, which indicates that the classifier with Euclidean distance has a better effect on the performance of the IMACO-CSFS algorithm as the feature dimension increases. Fig. 10 shows the subset of features selected by the eight algorithms on Huston 2013 dataset when $F_n=20$. It can be found from Fig. 10 that the features selected by the new method IMACO-CSFS are more dispersed than those selected by the methods (b), (c), (d), (e), (f), (g), which means that the features selected by IMACO-CSFS have lower low redundancy. Further more, the number of features (colored stripes) concentrated between the 0-25 bands in IMACO-CSFS is smaller than all other methods. And because in the subsequent iterative process, IMACO-CSFS will iteratively search around the features in the full-bands range selected in the previous iteration (It’s purpose is to find a better solution), which finally makes the features selected in the right part of Figure 10 show a shape similar to “heap”. Therefore, the features selected by IMACO-CSFS around the optimal solution are more evenly distributed than those selected by the original IMACO-FS, which also verifies the success of the priority sorting-based ant colony strategy proposed in this paper. Fig. 11 shows the classification graphs of the eight algorithms on the Huston 2013 dataset when $F_n=40$ and its original ground truth. Combining the annotations of each class in Fig. 4 with Fig. 11, we can find that the new method is more advantageous in the land cover classes Trees, Grass stressed and Grass healthy, especially Grass stressed class and Grass healthy class with similar spectral characteristics. It is because the original IMACO-FS method on which IMACO-CSFS is based outperforms the benchmark in land use/cover class distinction [34]. Therefore, compared with the other seven algorithms, the new method has unique advantages in distinguishing different land cover categories in the target scene, especially in complex categories with similar spectral characteristics.

D. PAVIAU-PAVIAC EXPERIMENT

Fig. 12 shows the classification accuracy of several algorithms under different feature dimensions on PaviaC dataset. As shown in Figure 12, the OA and $\kappa$ of IMACO-FS and IMACO-CSFS all reach their maximum when $F_n=40$, and their AA rise to the maximum at $F_n=30$. The overall performance of IMACO-CSFS in the PaviaC dataset is relatively stable, and it can almost surpass the other seven algorithms in each feature dimension. It is worth noting that compared with Huston 2013 dataset, three curves of CDIRF perform better than the six methods except IMACO-CSFS by a large margin on PaviaC dataset, probably because it has a stronger ability to identify each class in PaviaC dataset. The results of the features selected by the eight algorithms when $F_n=20$ are shown in Figure 13. It can be found from Fig. 13 that the number of features concentrated between the 0-15 bands in IMACO-CSFS is much less than other methods and these features are more dispersed. Since in the subsequent iterative process, IMACO-CSFS will iteratively search around the features selected in the previous iteration, which finally makes the features selected in the right part of Figure 13 show a shape similar to “pile”. Therefore, the features selected by IMACO-CSFS around the optimal solution are more evenly distributed than those selected by other methods, IMACO-CSFS tends to select features that are more dispersed and uniform, which benefits from the priority sorting-based ant colony strategy. Fig. 14 contains the classification graphs of several algorithms on PaviaC dataset when $F_n=40$. Combining Fig. 5 and Fig. 14, we can see that IMACO-CSFS has better classification performance on classes Meadow and Soil than several other algorithms.

E. SHANGHAI-HANGZHOU EXPERIMENT

Fig. 15 shows the classification accuracy of several algorithms under different feature dimensions on Hangzhou dataset. As shown in Fig. 15, the OA, $\kappa$ and $\kappa$ of IMACO-FS and IMACO-CSFS almost all reached their maximum at $F_n=8$. On the whole, OA and $\kappa$ of IMACO-CSFS is almost always larger than that of other several algorithms, however, AA of IMACO-CSFS is inferior to that of them when $F_n$ is big, its disadvantages is revealed with the increase of $F_n$. Note that when $F_n \in [8, 10, 12]$, the $\kappa$ of CDWOASA is slightly higher than that of CDIRF, which corresponds to the slightly higher $\kappa$ of CDWOASA than CDIRF in the Hangzhou dataset in Table 7 below. It may be because CDWOASA provides slightly more mutual information about the strong agreement between the ground truth map and the classification map of the Hangzhou dataset than CDIRF. The results of the features selected by the eight algorithms when $F_n=20$ are shown in Figure 16. It can be seen from Fig. 16 that IMACO-CSFS tends to choose more uniform and dispersed features. Fig. 17 shows the classification graphs of the eight algorithms on the Hangzhou dataset when $F_n=8$ and its original ground truth. Combining Fig. 6 and Fig. 17, we can find that IMACO-CSFS has advantages in the land cover classes Water and Land/Building, but the classification effect in the class Plant is not as good as the other seven methods, the reason may be that IMACO-CSFS has a weaker recognition ability for this class, resulting in lower classification performance on this class.

F. ANALYSIS OF EXPERIMENTAL RESULTS

Firstly, this paper compares IMACO-CSFS with the original IMACO-FS. Table 6 shows OA, $\kappa$ and $\kappa$ of Huston 2013, PaviaC and Hangzhou datasets obtained by IMACO-FS and IMACO-CSFS when the feature dimension $F_n$ is 40, 40 and 8, respectively (IMACO-CSFS adds source scene information and considers discriminability of different kinds of ground objects in target scene). In the Huston 2013 dataset, the OA, $\kappa$ and $\kappa$ of IMACO-CSFS and IMACO-FS both reach the maximum value when the selected feature dimension $F_n=40$, and the OA, $\kappa$ and $\kappa$ of IMACO-CSFS are $0.0456$, $0.0501$ and $0.0447$ higher than that of IMACO-FS, respectively. Compared with IMACO-FS, the OA, $\kappa$ and $\kappa$ of IMACO-CSFS are respectively improved by $0.0303$, $0.0363$ and $0.0551$ in the PaviaC dataset. In the Hangzhou
FIGURE 12. Accuracies on PaviaC dataset achieved by the eight algorithms under different features. (a) OA. (b) AA. (c) \( \kappa \).

FIGURE 13. Features selected by the eight algorithms on PaviaC dataset. (a) IMACO-FS. (b) SS-DDNMF. (c) SFA-MSSN. (d) CDWOASA. (e) CDIRF. (f) PSO-OBFS. (g) DRB-RESET. (h) IMACO-CSFS.

dataset, the OA, AA and \( \kappa \) of IMACO-CSFS and IMACO-FS reached the maximum value when \( F_n=8 \). Compared with IMACO-FS, the OA, AA and \( \kappa \) of IMACO-CSFS are respectively improved by 0.0473, 0.0462 and 0.0507. It seems that the effect on the Shanghai-Hangzhou data pair is better, because the target scene Hangzhou has a large number of labeled samples, which is more conducive to IMACO-CSFS, a cross-scene classification method that not only adds source
scene information but also considers the discriminability of different classes of target scenes in the classification process. Experimental results illustrate that the new feature selection method (IMACO-CSFS) applied to cross-scene classification performs better.

Secondly, this paper compares the average accuracy of the eight algorithms across different feature dimensions on the Houston 2013, PaviaC and Hangzhou datasets. As shown in Table 7, the OA, AA and $\kappa$ of IMACO-CSFS are 0.0222, 0.0079 and 0.0218 higher than the second highest CDIRF on the Houston 2013 dataset, respectively. On the PaviaC dataset, OA, AA and $\kappa$ of IMACO-CSFS are 0.0093, 0.0008 and 0.0231 higher than the second highest CDIRF, respectively. The improvement of OA of IMACO-CSFS in this group of experiments is smaller than that in the Houston group experiment, because IMACO-CSFS has a unique advantage in distinguishing different classes of ground objects in the target scene compared with the other several cross-scene algorithms, and the Huston 2013 dataset has 7 different classes, while the PaviaC dataset has only 6 classes. It can be concluded that the method proposed in this paper has better classification effect in Huston 2013 dataset with more classes. And on the Hangzhou dataset, $\kappa$ of IMACO-CSFS are 0.0278 higher than the second highest CDWOASA, OA of IMACO-CSFS are 0.0184 higher than the second highest CDIRF, while its AA is 0.0016 lower than CDIRF. It seems that the improvement of AA by the new method (IMACO-CSFS) is not as large as that of OA and $\kappa$, and even the AA of IMACO-CSFS on the Hangzhou dataset is smaller than that of CDIRF, mainly because the original IMACO-FS method is almost used to improve the overall classification accuracy and Kappa coefficients of feature subsets at different dimensions in the three datasets.
FIGURE 16. Features selected by the eight algorithms on Hangzhou dataset. (a) IMACO-FS. (b) SS-DDNMF. (c) SFA-MSSN. (d) CDWOASA. (e) CDIRF. (f) PSO-OBFS. (g) DRB-RESET. (h) IMACO-CSFS.

TABLE 6. Accuracies on Huston 2013, PaviaC and Hangzhou Datasets Obtained by IMACO-FS and IMACO-CSFS.

| Datasets   | Feature Selection | OA  | AA  | k   | Feature Dimension |
|------------|-------------------|-----|-----|-----|-------------------|
| Huston2013 | IMACO-FS          | 0.9098 | 0.9091 | 0.8972 | 40                |
|            | IMACO-CSFS        | **0.9554** | **0.9592** | **0.9419** | 40                |
| PaviaC     | IMACO-FS          | 0.8559 | 0.8551 | 0.8060 | 40                |
|            | IMACO-CSFS        | **0.8862** | **0.8914** | **0.8611** | 40                |
| Hangzhou   | IMACO-FS          | 0.9110 | 0.9101 | 0.8992 | 8                 |
|            | IMACO-CSFS        | **0.9583** | **0.9563** | **0.9199** | 8                 |

The values in bold are the highest accuracy among compared algorithms.

In addition, in order to more intuitively understand the computational complexity of each algorithm, we compared several algorithms on the three data pairs under the same conditions, and the experimental results are shown in Fig. 18. All experiments are performed with MATLAB R2020 on an Intel(R) Xeon(R) CPU E5-1620 V4 with 32 GB of RAM. It can be seen from Fig. 18 that the computation time of the newly proposed IMACO-CSFS method on all three data pairs is greatly reduced compared with the original IMACO-FS method: for example, the total training and
testing time spent by IMACO-CSFS on the Huston data pair is 72.4 seconds, which is only about 66% of that spent by IMACO-FS of 109.5 seconds; the calculation time of 62 seconds for IMACO-CSFS on the Pavia data pair is only about 71% of the 86.9 seconds for IMACO-FS; the calculation time of 64.3 seconds for IMACO-CSFS on the Shanghai data pair is only about 72% of the 89.6 seconds for IMACO-FS. That verifies the success of the newly proposed ant colony strategy based on elite ants in IMACO-CSFS. However, IMACO-CSFS is an iterative method, and the iterative process increases the computational burden, so its computation time is longer than DRB-RESNET which is not an iterative method: the computation time of IMACO-CSFS on the Huston data pair is 5.9 seconds longer than that of DRB-RESNET; the computation time of IMACO-CSFS on the Pavia data pair is 7.9 seconds longer than that of DRB-RESNET; the computation time of IMACO-CSFS on the Shanghai data pair is 4.7 seconds longer than that of DRB-RESNET. On the three data pairs, the computation time of DRB-RESNET is always the shortest, the second is IMACO-CSFS (the method proposed in this paper), and the last one is the original IMACO-FS method: the computational complexity of all eight methods is shown above the column in Fig. 18. Furthermore, we can find that the execution time of these algorithms on the Pavia data pair is shorter than that on the Huston data pair and the Shanghai-Hangzhou data pair on the whole. It may be because the number of labeled samples of the former is smaller than that of the latter. And the reason why the running time of the algorithms on Huston data pair is greater than that
TABLE 7. Accuracies on houston 2013, PaviaC and Hangzhou datasets.

| Datasets     | Algorithm | OA  | AA   | \(\kappa\)   |
|--------------|-----------|-----|------|---------------|
| Huston 2013  | IMACO-FS  | 0.8706 | 0.8913 | 0.8655 |
|              | SS-DDNMF  | 0.8796 | 0.9112 | 0.8891 |
|              | SFA-MSSN  | 0.8664 | 0.9087 | 0.8974 |
|              | CDOASA    | 0.9184 | 0.9209 | 0.9093 |
|              | CDIF      | 0.9189 | 0.9372 | 0.9106 |
|              | PSO-OBFS  | 0.8785 | 0.9024 | 0.8901 |
|              | DRB-RESET | 0.9177 | 0.9259 | 0.9056 |
|              | IMACO-CSFS| 0.9411 | 0.9451 | 0.9324 |
| PaviaC       | IMACO-FS  | 0.8203 | 0.8296 | 0.7864 |
|              | SS-DDNMF  | 0.8473 | 0.8379 | 0.8014 |
|              | SFA-MSSN  | 0.8524 | 0.8405 | 0.8119 |
|              | CDOASA    | 0.8592 | 0.8573 | 0.8211 |
|              | CDIF      | 0.8617 | 0.8735 | 0.8262 |
|              | PSO-OBFS  | 0.8499 | 0.8376 | 0.8005 |
|              | DRB-RESET | 0.8547 | 0.8605 | 0.8186 |
|              | IMACO-CSFS| 0.8710 | 0.8743 | 0.8493 |
| Hangzhou     | IMACO-FS  | 0.8933 | 0.9015 | 0.8601 |
|              | SS-DDNMF  | 0.9160 | 0.9275 | 0.8817 |
|              | SFA-MSSN  | 0.9194 | 0.9278 | 0.8805 |
|              | CDOASA    | 0.9247 | 0.9305 | 0.8876 |
|              | CDIF      | 0.9339 | 0.9426 | 0.8848 |
|              | PSO-OBFS  | 0.9126 | 0.9205 | 0.8793 |
|              | DRB-RESET | 0.9275 | 0.9258 | 0.8856 |
|              | IMACO-CSFS| 0.9523 | 0.9410 | 0.9154 |

The values in bold are the highest accuracy among compared algorithms.

FIGURE 18. The total time to train and test (in Seconds) of each algorithms in the three data pairs.

V. CONCLUSION

In this paper, a new cross-scene feature selection algorithm IMACO-CSFS is proposed. In order to make the subsequent search process focus on the global optimal solution (optimal feature subset) found in the previous iteration, IMACO-CSFS proposes a priority sorting-based ant colony strategy to obtain more accurate feature subsets of the two scenes than other cross-scene feature selection methods. Furthermore, in order to further accelerate the convergence speed of the global optimal solution, IMACO-CSFS introduces an ant colony strategy based on elite ants to obtain the optimal feature subsets of two scenes more efficiently. Finally, the original IMACO-FS method based only on a single scene is successfully applied to the fields of cross-scene HSIs classification and image classification in this paper, which further improves the update strategy of pheromone in the ant colony algorithm and greatly improves the feature selection accuracy and image classification accuracy of the target scene. IMACO-CSFS’s superiority is demonstrated on three public HSIs datasets—Huston, Pavia data pair, and Shanghai data pair. However, this method is not perfect, as the improvement of AA in this research was small, and we might incorporate improving their AA into the scope of the follow-up work, such as selecting some feature subsets \( M \) from the candidate feature subset \( F \) to pick the best combination of AA, OA and \( \mu \).

REFERENCES

[1] Y. Peng, Y. Zhang, B. Tu, C. Zhou, and Q. Li, “Multiview hierarchical network for hyperspectral and LiDAR data classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 15, pp. 1454–1469, 2022, doi: 10.1109/JSTARS.2022.3144312.

[2] L. Pang, S. Men, L. Yan, and J. Xiao, “Rapid vitality estimation and prediction of corn seeds based on spectra and images using deep learning and hyperspectral imaging techniques,” IEEE Access, vol. 8, pp. 123026–123036, 2020.

[3] X. Yang and Y. Yu, “Estimating soil salinity under various moisture conditions: An experimental study,” IEEE Trans. Geosci. Remote Sens., vol. 55, no. 5, pp. 2525–2533, May 2017.

[4] Y. Qian, F. Yao, and S. Jia, “Band selection for hyperspectral imagery using affinity propagation,” IET Comput. Vis., vol. 3, no. 4, pp. 213–222, Dec. 2009.

[5] X. Yang, W.-D. Xu, H. Liu, and L.-Y. Zhu, “Research on dimensionality reduction of hyperspectral image under close range,” in Proc. Int. Conf. Commun., Inf. Syst. Comput. Eng. (CISCE), Jul. 2019, pp. 171–174.

[6] E. Myasnikov, “Using UMAP for dimensionality reduction of hyperspectral data,” in Proc. Int. Multi-Conf. Ind. Eng. Mod. Technol., Oct. 2020, pp. 1–5.

[7] J. Yin, Y. Wang, and J. Hu, “A new dimensionality reduction algorithm for hyperspectral image using evolutionary strategy,” IEEE Trans. Ind. Inf., vol. 8, no. 4, pp. 935–943, Nov. 2012.

[8] J. Jiang, J. Ma, C. Chen, Z. Wang, Z. Cai, and L. Wang, “SuperPCA: A superpixelwise PCA approach for unsupervised feature extraction of hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens., vol. 56, no. 8, pp. 4581–4593, Aug. 2018.

[9] S. D. Fabiyi, P. Murray, J. Zabalza, and J. Ren, “Folded LDA: Extending the linear discriminant analysis algorithm for feature extraction and data reduction in hyperspectral remote sensing,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 12312–12331, 2021.

[10] Q. Hou, Y. Wang, L. Jing, and H. Chen, “Linear discriminant analysis based on kernel-based possibilistic C-means for hyperspectral images,” IEEE Geosci. Remote Sens. Lett., vol. 16, no. 8, pp. 1259–1263, Aug. 2019.
[11] I. Dopedio, M. Zoretta, A. Villa, A. Plaza, and P. Gamba. “Unmixing prior to supervised classification of remotely sensed hyperspectral images,” IEEE Geosci. Remote Sens. Lett., vol. 8, no. 4, pp. 760–764, Jul. 2011.

[12] G. Taşkın, H. Kaya, and L. Bruzzone. “Feature selection based on high dimensional model representation for hyperspectral images,” IEEE Trans. Image Process., vol. 26, no. 6, pp. 2918–2928, Jun. 2017.

[13] J. Wang, M. Ye, F. Xiong, and Y. Qian, “Cross-scene hyperspectral feature selection via hybrid whale optimization algorithm with simulated annealing,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 2473–2483, 2021.

[14] H. Liu, M. Zhou, and Q. Liu, “An embedded feature selection method for imbalanced data classification,” IEEE/CAA J. Autom. Sinica, vol. 6, no. 3, pp. 703–715, May 2019.

[15] X. Chen, G. Yuan, F. Nie, and Z. Ming. “Semi-supervised feature selection via sparse rescaled linear square regression,” IEEE Trans. Knowl. Data Eng., vol. 32, no. 1, pp. 165–176, Jan. 2020.

[16] M. Robnik-Sikonja and I. Kononenko. “Theoretical and empirical analysis of ReliefF and RReliefF,” Mach. Learn., vol. 53, nos. 1–2, pp. 23–69, Oct. 2003.

[17] Z. Huang, C. Yang, X. Zhou, and T. Huang, “A hybrid feature selection method based on binary state transition algorithm and ReliefF,” IEEE J. Biomed. Health Inform., vol. 23, no. 5, pp. 1888–1898, Sep. 2019.

[18] S. Gao, G. V. Steeg, and A. Galstyan, “Variational information maximization for feature selection,” in Proc. Neural Inf. Process. Syst., 2016, pp. 487–495.

[19] C. Lee and G. G. Lee. “Information gain and divergence-based feature selection for machine learning-based text categorization,” Inf. Process. Manag., vol. 42, no. 1, pp. 155–165, 2006.

[20] B. Tu, J. Wang, X. Kang, G. Zhang, X. Ou, and L. Guo, “KNN-based representation of superpixels for hyperspectral image classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 11, no. 11, pp. 4032–4047, Nov. 2018.

[21] J. Huang, Y. Cai, and X. Xu, “A hybrid genetic algorithm for feature selection wrapper based on mutual information,” Pattern Recognit. Lett., vol. 28, no. 13, pp. 1825–1844, May 2007.

[22] M. M. Mafarja and S. Mirjalili, “Hybrid whale optimization algorithm with simulated annealing for feature selection,” Neurocomputing, vol. 260, pp. 302–312, Oct. 2017.

[23] M. Wang, C. Wu, L. Wang, D. Xiang, and X. Huang, “A feature selection approach for hyperspectral image based on modified ant lion optimizer,” Knowl.-Based Syst., vol. 168, no. 15, pp. 39–48, Mar. 2019.

[24] H. Liu, W. Li, X.-G. Xia, M. Zhang, C.-Z. Gao, and R. Tao, “Spectral shift mitigation for cross-scene hyperspectral imagery classification,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 6624–6638, 2021.

[25] Y. Zhang, W. Li, R. Tao, J. Peng, Q. Du, and Z. Cai, “Cross-scene hyperspectral image classification with discriminative cooperative alignment,” IEEE Trans. Geosci. Remote Sens., vol. 59, no. 1, pp. 9646–9660, Nov. 2021.

[26] G. A. Shaw and H.-K. K. Burke, “Spectral imaging for remote sensing,” Lincoln Lab. J., vol. 14, no. 1, pp. 3–28, 2003.

[27] Y. Xu, J. Hu, K. Wattanachote, K. Zeng, and Y. Gong, “Sketch-based shape retrieval via best view selection and a cross-domain similarity measure,” IEEE Trans. Multimedia, vol. 22, no. 11, pp. 2950–2962, Nov. 2020.

[28] B. Du, L. Zhang, D. Tao, and D. Zhang, “Unsupervised transfer learning for target detection from hyperspectral images,” Neurocomputing, vol. 120, pp. 72–82, Nov. 2013.

[29] C. Persello and L. Bruzzone, “Kernel-based domain-invariant feature selection in hyperspectral images for transfer learning,” IEEE Trans. Geosci. Remote Sens., vol. 54, no. 5, pp. 2615–2626, May 2016.

[30] M. Long, Y. Cao, J. Wang, and M. Jordan, “Learning transferable features with deep adaptation networks,” in Proc. Int. Conf. Mach. Learn., 2015, pp. 97–105.

[31] C. Zhang, M. Ye, L. Lei, and Y. Qian, “Feature selection for cross-scene hyperspectral image classification using cross-domain I-ReliefF,” IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens., vol. 14, pp. 5932–5949, 2021.

[32] X. Liu, X. Li, L. Liu, J. He, and B. Ai, “An innovative method to classify remote-sensing images using ant colony optimization,” IEEE Trans. Geosci. Remote Sens., vol. 46, no. 12, pp. 4198–4208, Dec. 2008.
YOUHUA YU received the B.E. degree in digital media technology from the Guilin University of Electronic Technology, Guilin, China, in 2019, where he is currently pursuing the master’s degree in computer science and technology.

His main research interests include hyperspectral image processing and cross-scene hyperspectral image classification.

XIAOLAN XIE (Member, IEEE) received the M.S. degree in computer science from Shanghai Maritime University, in 2001, and the Ph.D. degree in mechanical manufacturing and automation from Xidian University, China, in 2009.

She is currently working at the College of Information Science and Engineering, Guilin University of Technology, and the Guangxi Key Laboratory of Embedded Technology and Intelligent System. She is also a Senior Visiting Scholar with Middlesex University, U.K. She has published more than 100 scientific research articles, including more than 50 SCI/EI articles and more than 20 Chinese core journals. Her main research interests include cloud computing, big data, geophysical exploration and information technology, intelligent computing, and manufacturing informatization. She is currently the Director of the International Institute of Engineering and Technology (IETI) and a fellow of the American Computer Society (ACM), the Chinese Computer Society (CCF), the Cloud Computing Expert Committee of the Chinese Institute of Communications (CIC), the High Performance Computing Expert Committee of CCF, and the Collaborative Computing, Distributed Computing and Processing Expert Committee of CCF.

YIGANG TANG received the B.E. degree from the Guilin University of Technology, Guiling, China, in 2020, where he is currently pursuing the M.S. degree.

His main research interests include remote sensing image process and hyperspectral image classification.

YARONG LIU received the M.S. degree in computer application technology from the Guilin University of Technology, Guilin, China, in 2007.

Since 2002, she has been with the School of Information Science and Engineering, Guilin University of Technology. Her main research interests include new communication technologies and computer applications.

---

VOLUME 10, 2022