IEEE Access
Multidisciplinary | Rapid Review | Open Access Journal

Received January 2, 2020, accepted January 25, 2020, date of publication February 3, 2020, date of current version February 11, 2020.

Digital Object Identifier 10.1109/ACCESS.2020.2971327

An Improved Ant Colony Algorithm for Optimized Band Selection of Hyperspectral Remotely Sensed Imagery

XIAOHUI DING[1,2,3,4,5], HUAPENG LI[2], JI YANG[1,3,4,5], PATRICIA DALE[6], XIANGCONG CHEN[2,7], CHUNLEI JIANG[8], AND SHUQING ZHANG[2]

[1]Guangzhou Institute of Geography, Guangzhou 510070, China
[2]Northeast Institute of Geography and Agroecology, Chinese Academy of Sciences, Changchun 130102, China
[3]Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou), Guangzhou 511458, China
[4]Guangdong Open Laboratory of Geospatial Information Technology and Application, Guangzhou 510070, China
[5]Key Laboratory of Guangdong for Utilization of Remote Sensing and Geographical Information System, Guangzhou 510070, China
[6]Environmental Futures Research Institute, School of Environment, Griffth University, Brisbane, QLD 4122, Australia
[7]Harbin Institute of Geotechnical Investigation and Surveying, Harbin 150010, China
[8]School of Geomatics and Prospecting Engineering, Jilin Jianzhu University, Changchun 130114, China

Corresponding author: Huapeng Li (lihuapeng@iga.ac.cn)

This work was supported with the National Key Research and Development Program of China under Grant 2016YFB0502301, in part by the National Natural Science Foundation of China under Grant 41301465 and Grant 4197060691, in part by the Key Special Project for Introduced Talents Team of Southern Marine Science and Engineering Guangdong Laboratory (Guangzhou) under Grant GML2019ZD0301, in part by the Guangdong Innovative and Entrepreneurial Research Team Program under Grant 2016ZT06D336, and in part by the GDAS Project of Science and Technology Development under Grant 2019GDASYL-0501001.

ABSTRACT The ant colony algorithm (ACA) has been widely used for reducing the dimensionality of hyperspectral remote sensing imagery. However, the ACA suffers from problems of slow convergence and of local optima (caused by loss of population diversity). This paper proposes an improved ant colony algorithm (IMACA) based band selection algorithm (IMACA-BS), to overcome the two shortcomings of the standard ACA. For the former problem, a pre-filter is applied to improve the heuristic desirability of the ant colony system; the Pearson's similarity measurement of the degree of redundancy among the selected bands is taken as one of the terms in the heuristic function, and this further accelerates the convergence of the IMACA-BS. For the latter problem, a pseudo-random rule and an adaptive information update strategy are, respectively, introduced to increase the population diversity of the ant colony system. The effectiveness of the proposed algorithm was evaluated on three public datasets (Indian Pines, Pavia University and Botswana datasets), and compared with a series of benchmarks. Experimental results demonstrated that the IMACA-BS consistently achieved the highest overall classification accuracies and significantly outperformed other benchmarks over all of the three experiments. The proposed IMACA-BS is, therefore, recommended as an effective alternative for band selection of hyperspectral imagery.

INDEX TERMS Hyperspectral remotely sensed imagery, band selection, ant colony algorithm, artificial intelligence.

I. INTRODUCTION

Hyperspectral image (HSI) containing hundreds of spectral bands provides abundant spectral information about on-ground objects. HSI has broad application prospects in a wide range of fields, among, inter alia, agriculture, forestry, environment and ecology. However, the massive bands of HSI pose a great challenge to data processing and analysis [1], leading to the curse of dimensionality phenomena [2]. It is, therefore, essential to reduce data dimensions to facilitate the analysis of hyperspectral data. Feature extraction and feature selection are two kinds of typical data reduction techniques [3]. Feature extraction techniques (e.g., principal components analysis), are developed to compress data by using complex mathematical transformations. Though these methods may extract useful features from the HSI data sets, the interpretable primitive physical significance of the data is often lost in the data transformation process.

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

In recent years, feature selection methods have received increasing attention in light of their reliable performance in preserving the primitive physical interpretability of the original data [4]. These methods aim to select a feature subset with low redundancy and high relevance to the specific task (e.g., classification) [5]. In general, feature selection methods can be categorized into three types: unsupervised, semi-supervised, and supervised [6]. Unsupervised methods select feature subsets without the assistance of labeled classes. However, it is very hard for them to achieve promising results since they may neglect the possible correlation between different features. Semi-supervised methods utilize labeled and unlabeled samples to maximize the margin between data points of different classes, but they rely heavily on the number of labeled samples [7]. In contrast, supervised methods that select bands with powerful classification capacity under the supervision of class labels always achieve promising band selection results [8]. In this paper, we focus on supervised feature selection methods.

Generally, supervised feature selection consists of three types of searching strategies. These are exhaustive searches, sequential searches, and random searches [9]. Exhaustive searches (such as branch and bound [10]) enumerate all the possible combinations of features, which always leads to an unacceptable high time complexity [11]. Sequential searches, such as sequential forward selection (SFS), sequential backward selection (SBS), and sequential floating forward selection (SFFS) [12], are also computationally intensive as well as increasing the number of selected features. In comparison, random searches select features randomly and always deliver promising results with relatively high efficiency. Recently, nature-inspired random search algorithms have been widely used for feature selection from HSI data [13], [14], due to their powerful searching capacity in a high-dimensional space. These include genetic algorithms (GA) [15], clone selection algorithms (CSA) [16], particle swarm optimization (PSO) [17], and ant lion optimizer (ALO) [18].

The ant colony algorithm (ACA) is a typical heuristic search algorithm inspired by natural biological systems. It simulates the behaviors of real ants in the process of seeking food. Through the cooperation between individual ants, ACA has the capacity to solve complicated optimization problems, such as finding the optimal route to food [19]. The positive feedback mechanism of ACA may help an ant colony find the optimal solutions in a short time. ACA has been used in the field of image processing, pattern recognition, and feature selection [20]–[22]. These efforts have demonstrated the ACA as a potentially effective algorithm to solve complex optimization problems. Recently, ACA was also applied in feature selection from HSI data [20], [23]. Nonetheless, some studies illustrated that the standard ACA was incapable of solving the NP-hard feature selection problem due to its two intrinsic flaws [24]. First, the ACA is often slow to converge on optimal solutions owing to its random initialization strategy. Second, the population diversity of the ant colony may decrease significantly during the searching process because of the strong positive feedback via pheromone trails. These two issues easily trap the standard ACA on local optima in the process of feature selection.

In this paper, an improved ant colony algorithm (IMACA) based band selection algorithm (IMACA-BS) is proposed to select the optimal band subset from the HSI data sets. In order to accelerate the convergence of the ant colony system, a pre-filter is introduced to optimize pheromone initialization of the ant colony system. Pearson's similarity, measuring the degree of redundancy among the selected bands, is employed in the heuristic function. To maintain the population diversity of the ant colony, a pseudo-random rule and an adaptive information update strategy are both employed in the IMACA-BS. The effectiveness of the proposed IMACA-BS was tested on three types of public HSI datasets and compared with a series of benchmarks, including the ACA based band selection algorithm (ACA-BS), the genetic algorithm-based band selection algorithm (GA-BS), the particle swarm optimization-based band selection algorithm (PSO-BS), the hybridization of GA and PSO based band selection algorithms (GAPSO-BS), the sequential forward floating selection-based band selection algorithm (SFFS-BS), and the newly proposed ant lion optimizer-based band selection method (ALO-BS) [25].

The key innovations of the proposed method can be summarized as: 1) a novel pre-filter is designed and applied for the first time to accelerate the convergence of the standard ACA-BS, and 2) an adaptive information update strategy is introduced to avoid the ant colony being trapped on local optima. The paper is organized as follows. Section 2 introduces the proposed IMACA-BS in detail. Section 3 presents the HSI datasets. The experimental results are provided in Section 4, followed by a discussion in Section 5. The conclusions are drawn in Section 6.

II. METHODS

A. CRITERION FUNCTION

The criterion function evaluating the contribution of each band to classification accuracy is critical to band selection algorithms. Due to the possible appearance of singular covariance, some statistical-based criterion functions (e.g., Bhattacharya distance, Jeffreys–Matusita distance, and Divergence) are not suitable for band selection. In practice, the overall classification accuracy (OA) is therefore employed as the criterion function in band selection algorithms [26], where the band subsets are selected to maximize the OA. In this research, the support vector machine (SVM) that always achieves accurate classification results with a limited number of training samples, was employed and served as the criterion function for band selection [27].

B. ACA-BASED BAND SELECTION ALGORITHM (ACA-BS)

In the standard ACA, a fully connected undirected weighted graph $G =< B, E >$ is used to represent the solution space of the ant colony, where $B = \{b_1, b_2, \ldots, b_n\}$ indicates the

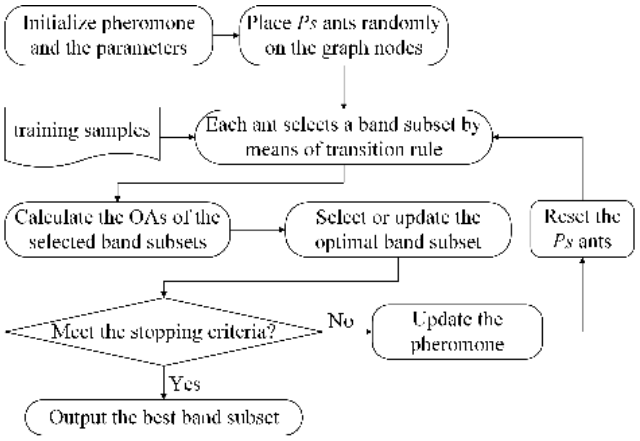


FIGURE 1. Flow of band selection based on the ant colony algorithm.

graph nodes (each node represents a band of hyperspectral image), and $E = \{(b_i, b_j), b_i, b_j \in B\}$ denotes the edges of the graph. The edges traveled by the same ant in one iteration construct a path. $VB = \{vb_1, vb_2, \ldots, vb_m\}$ (a subset of B) represents the selected band subset ($m < n$). The major steps of the ACA for band selection are illustrated in Figure 1. First, initialize the control parameters and pheromone of the ant colony system. Second, release a total of Ps (population size of the ant colony) ants and randomly allocate them to the graph nodes; each ant selects a band subset independently by exploring the nodes in sequence using a transition rule. Third, when selecting m nodes for each ant, calculate the OA of the band subsets and select or update the optimal one (with the best OA). Finally, update the pheromone concentration on the paths, and continue the iterations until the stopping criterion is met. Two major operations of the ACA-BS, the transition rule and the update of pheromone, are elaborated in detail hereafter.

1) TRANSITION RULE

In the ACA-BS, each ant selects the next node by means of a roulette wheel selection with the probability pb_{ij} (Equation (1)), which is directly proportional to the pheromone concentration on the edge (b_i, b_j) and the heuristic desirability $\eta_{ij}(t)$ of the next selected node.

$$pb_{ij} = \frac{\tau_{ij}(t)^{\alpha}\eta_{ij}(t)^{\beta}}{\sum_{j \notin tabu^k} \tau_{ij}(t)^{\alpha}\eta_{ij}(t)^{\beta}} \tag{1}$$

where α denotes the rate of information accumulation during the movement of the ants, and β is the heuristic coefficient; $tabu^k$ represents the tabu list of nodes that have been visited by the ant k; $\tau_{ij}(t)$ and $\eta_{ij}(t)$ denote the pheromone concentration and heuristic desirability on the edge (b_i, b_j) at time t, respectively. The heuristic desirability can be calculated as:

$$\eta_{ij}(t) = O_{ij} \tag{2}$$

where O_{ij} is the overall classification accuracy achieved by SVM with bands i and j.

2) UPDATE OF PHEROMONE

In this study, an ant-quantity system proposed by Dorigo *et al.* (1999) was adopted to simulate ants' secretion behavior in which the amount Q of pheromone secreted by each ant is a constant [28]. The pheromone concentration τ_{ij} on the edge (b_i, b_j) will change with the time as follows:

$$\tau_{ij}(t+1) = (1-\rho)\tau_{ij}(t) + \sum_{k=1}^{Ps} \Delta\tau_{ij}^{k} \tag{3}$$

where ρ is a volatility coefficient that controls the volatilization rate of the pheromone, and $\Delta\tau_{ij}^{k}$ is the newly dropped pheromone by the ant k on the the edge (b_i, b_j) which can be calculated as:

$$\Delta\tau_{ij}^{k} = Q \times \frac{O_{ij}}{O_{max}} \tag{4}$$

where O_{max} denotes the maximum OA provided by the SVM with band i and one of the other $n-1$ bands; Q is a constant.

C. IMPROVED ANT COLONY ALGORITHM-BASED BAND SELECTION ALGORITHM (IMACA-BS)

In the IMACA-BS, the pheromone on each path is initialized according to the heuristic desirability between each pair of nodes by means of a pre-filter. During the searching process, the information between the potential node and all the selected nodes is applied to guide a search ant to make a state transition with the pseudo-random rule. As well, an adaptive information strategy is introduced to avoid leaving pheromone only in certain (rather than all) optimal paths by the searching ants. before presenting the flowchart of the proposed IMACA-BS, we first elaborate below the pre-filter and the searching process of the ant colony.

1) PRE-FILTER

The pheromone on the edge (b_i, b_j) is initialized according to the heuristic desirability O_{ij}. The overall accuracies with band i and one of the other $n-1$ bands are calculated, and the p (a user defined number) bands that produce the highest OAs with band i are selected and denoted as C (Figure 2). Note that the value of p can be determined according to statistical result [29], and it is set as $n/2$ in this research. The pheromone on the edges between node (band)i and other $n-1(j)$ bands can be calculated as follows:

$$\tau_{ij}(0) = \begin{cases} \dfrac{O_{ij}}{O_{max}}, & j \in C \\ \dfrac{O_{min}}{O_{max}}, & \text{otherwise} \end{cases} \tag{5}$$

where O_{ij} is the OA of SVM with band i and band j, O_{max} and O_{min} represent the maximum and minimum OA provided by SVM with band i and one of the other $n-1$ bands.

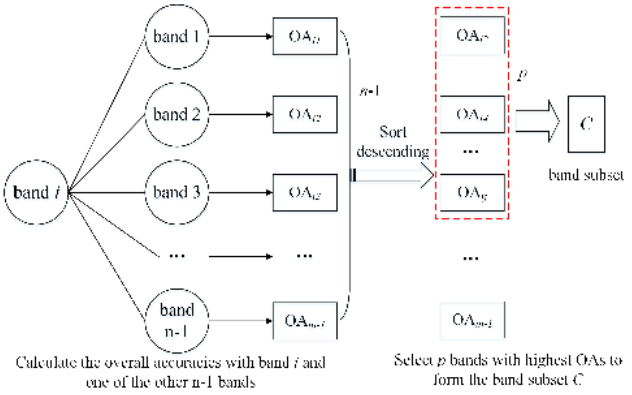


FIGURE 2. The process of pre-filter.

2) SEARCH PROCESS OF ANT COLONY

When the search ant k arrives at band i, it will make a state transition according to a pseudo-random rule as follows:

$$S = \begin{cases} argmax\{\tau_{ij}(t)^{\alpha} || \eta_{ij}(t)^{\beta}\}, & q < q_0 \\ p_{ij}^{k}(t), & q \geq q_0 \end{cases} \tag{6}$$

where q is a random number uniformly distributed in [0, 1], and $q_0 = 1 - e^{-1/t}$ ($t = 1, 2, \ldots, T$, T denotes the number of maximum iterations); $p_{ij}^{k}(t)$ denotes the probability for the k-th ant to select the node j, which can be calculated as follows:

$$p_{ij}^{k}(t) = \begin{cases} \dfrac{\tau_{ij}(t)^{4\alpha t/T} \eta_{ij}(t)^{2\beta t/T}}{\sum_{j \notin \Gamma^k} \tau_{ij}(t)^{4\alpha t/T} \eta_{ij}(t)^{2\beta t/T}}, & j \in C \\ 0, & \text{otherwise} \end{cases} \tag{7}$$

To minimize the redundancy of the selected band subset so as to improve the global search capacity of the proposed algorithm, the $\eta_{ij}(t)$ adopted in the proposed IMACA-BS is determined according to Pearson's correlation as follows:

$$\eta_j(t)^k = \frac{O_{ij}}{1 + \sum_s r(s, j)}, \quad \forall s, i \in V^k \tag{8}$$

where V^k is the band subset visited by the ant k; $r(s, j)$ is the Pearson's correlation between the s-th and j-th bands calculated as:

$$r(s, j) = \frac{\sum_{n_p} (R_{b_s} - \bar{R}_{b_s})(R_{b_j} - \bar{R}_{b_j})}{\sqrt{\sum_{n_p} (R_{b_s} - \bar{R}_{b_s})^2} \sqrt{\sum_{n_p} (R_{b_j} - \bar{R}_{b_j})^2}} \tag{9}$$

where R_{b_s} and R_{b_j} denote the reflectance value of bands s and j, respectively; $\bar{R}_{b_s}$ and $\bar{R}_{b_j}$ are the mean value of bands s and j, respectively; n_p is the total number of pixels of the hyperspectral imagery.

After all ants (with a number of Ps in total) have selected m bands, update the pheromone on edges as follows:

$$\tau_{ij}(t+1) = \begin{cases} (1-\rho) \times \tau_{ij}(t) + \sum_{k=1}^{Ps} \Delta\tau_{ij}^{k}, & j \in C \\ (1-\rho) \times \tau_{ij}(t), & \text{otherwise} \end{cases} \tag{10}$$

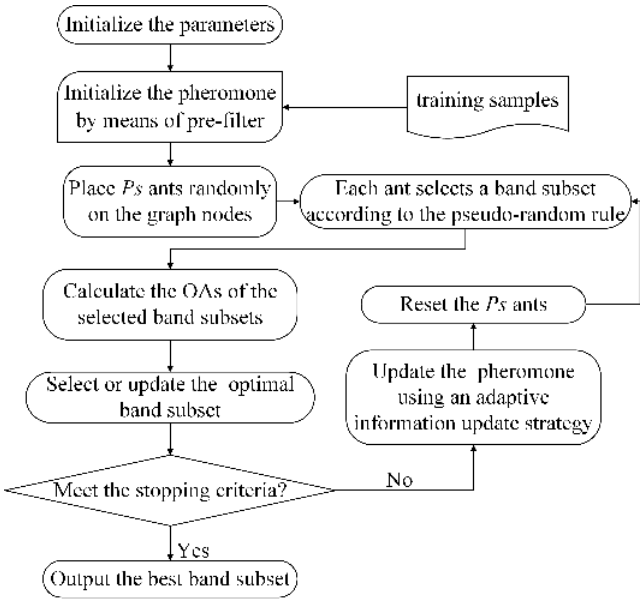


FIGURE 3. Flow of band selection based on the improved ant colony algorithm.

where $\Delta\tau_{ij}^{k}$ denotes the pheromone increment on the edges from the ant k for population diversity maintenance, which can be dynamically calculated as follows:

$$\Delta\tau_{ij}^{k} = \begin{cases} (Q \times \dfrac{O_{ij}}{O_{max}})^{\frac{2t}{T}+1}, & j \in C \\ 0, & \text{otherwise} \end{cases} \tag{11}$$

3) STEPS OF THE IMACA-BS

The flowchart of the proposed IMACA-BS algorithm is demonstrated in Figure 3, with the major steps being detailed as follows:

Step 1: Initialize the control parameters of the IMACA-BS, including Q, ρ, α, β, m, Ps, and the number of maximum iterations T.

Step 2: For each band i, a pre-filter is employed to initialize the pheromone according to Eq. (5), and set the evolution generation t as 0.

Step 3: Place Ps search ants on the nodes and let each ant randomly select an initial position which is subsequently added into the corresponding *tabu* list.

Step 4: Transfer the state of the ant k ($k = 1, \ldots, Ps$) from node i to j with the transition probability, until that the number of elements contained in a solution reaches m.

Step 5: Calculate the OA with the band subsets and memorize the current optimal band subset.

Step 6: Update the pheromone on the paths using the adaptive pheromone update strategy defined by Eq. (10). Continue the iterations by repeating step 3-6 until t is equal to T.

Step 7: Output the band subset with the best OA.

TABLE 1. Land cover classes and the number of samples from the Indian dataset.

Class Number	Land use/cover type	Number of Samples			
		Training	Validation	Test	Total
1	Alfalfa	8	7	31	46
2	Buildings-Grass-Trees-Drives	25	25	336	386
3	Corn	25	25	187	237
4	Corn-mintill	25	25	780	830
5	Corn-notill	25	25	1378	1428
6	Grass-pasture	25	25	433	483
7	Grass-pasture-mowed	8	7	13	28
8	Grass-trees	25	25	680	730
9	Hay-windrowed	25	25	428	478
10	Oats	8	7	5	20
11	Soybean-clean	25	25	543	593
12	Soybean-mintill	25	25	2405	2455
13	Soybean-notill	25	25	922	972
14	Stone-Steel-Towers	25	25	43	93
15	Wheat	25	25	155	205
16	Woods	25	25	1215	1265
Total		349	346	9554	10249

TABLE 2. Land cover classes and the number of samples from the Pavia dataset.

Class Number	Land use/cover type	Number of Samples			
		Training	Validation	Test	Total
1	Asphalt	663	663	5305	6631
2	Bare Soil	502	503	4024	5029
3	Bitumen	133	133	1064	1330
4	Gravel	209	210	1680	2099
5	Meadows	1864	1865	14920	18649
6	Painted metal sheets	134	135	1076	1345
7	Self-Blocking Bricks	368	368	2946	3682
8	Shadows	95	94	758	947
9	Trees	306	306	2452	3064
	Total	4274	4277	34225	42776

III. DATA SOURCE

In our experiment, three public hyperspectral datasets were employed to test the performance of the proposed algorithm (IMACA-BS), including the indian pines, pavia university, and botswana datasets.

Indian Pines Dataset (INDIAN DATASET): The indian dataset is gathered by airborne visible/infrared imaging spectrometer (AVIRIS) sensor. It has a spatial extent of 145 $\times$ 145 pixels with a spatial resolution of 20 m. the dataset consists of 224 original bands over a spectrum range of 0.4-2.5μm. a total of 200 bands were reserved after discarding the bands that were adversely influenced by moisture absorption. there were 16 land use/cover types in the ground truth (Table 1). the number of training, validation and testing samples for each of the 16 classes collected from the ground truth map are shown in Table 1.

Pavia University Dataset (Pavia Dataset): The original pavia dataset, acquired by the reflective optics system imaging spectrometer (ROSIS) sensor, consists of 115 bands in the spectral region of 0.43-0.86 μm. the dataset has a spatial resolution of 1.3 m with a spatial extent of 610 $\times$ 340 pixels. the bands influenced by water absorption were removed, and there were 103 bands remaining in the dataset. in total nine land use/cover types were identified (Table 2). the number of training, validation and testing samples of each class are shown in Table 2.

Botswana Dataset: The original botswana dataset, collected by the hyperion sensor on EO-1 at a 30m pixel resolution, contains 242 bands covering a spectral range of 0.4-2.5 μm. after the elimination of the uncalibrated and noisy bands that were seriously influenced by water absorption, 145 bands were remained. The ground truth of the dataset contained 14 identified classes (Table 3). the number of training, validation and testing samples are shown in Table 3.

TABLE 3. Land cover classes and the number of samples from the Botswana dataset.

Class Number	Land use/cover type	Number of Samples			
		Training	Validation	Test	Total
1	Water	27	27	216	270
2	Hippo grass	10	10	81	101
3	Floodplain grasses1	25	25	201	251
4	Floodplain grasses2	22	21	172	215
5	Reeds1	27	26	216	269
6	Riparian	27	26	216	269
7	Firescar2	25	26	208	259
8	Island interior	20	20	163	203
9	Acacia woodlands	31	31	252	314
10	Acacia shrublands	25	25	198	248
11	Acacia grasslands	31	30	244	305
12	Short mopane	18	18	145	181
13	Mixed mopane	27	27	214	268
14	Exposed soils	10	9	76	95
	Total	325	321	2602	3248

IV. EXPERIMENTAL RESULTS AND ANALYSIS

A. EXPERIMENT SETUP

In this experiment, the effectiveness of the proposed algorithm (IMACA-BS) and benchmark comparators including ACA-BS, GA-BS, PSO-BS, GAPSO-BS, ALO-BS, and SFFS-BS, were investigated on the indian pines, pavia university, and botswana datasets. With the training and validation samples, the SVM classifier was used for the evaluation of the land-cover/use discrimination ability of the band subsets. The parameters of SVM were optimized using a 5-fold cross-validation. the band selection algorithms were coded and executed under the MATLAB environment on a pc (i.e. an intel (r) core (TM) CPU i5-4460 processor (3.20 GHz) and 8 GB RAM). Each algorithm was repeatedly run 10 times on each dataset, and the average OAs of the band subsets were calculated for different numbers of bands (5, 10, 15, 20, 25, 30, 35, and 40). The McNemar test was employed to compare and investigate the difference between the results of the proposed IMACA-BS and of the benchmarks.

Generally, the parameters of the aca-based band selection algorithm were set according to previous experience [30], [31]. The parameters of ACA-BS and IMACA-BS, including the population size ps, the information heuristic factor α, the desired heuristic factor β, and the pheromone evaporation factor ρ, exert a strong effect on the band selection results [32]. Here α and β are two adjustable parameters controlling the relative influence of pheromone τ_{ij} and heuristic desirability η_{ij}, and ρ determines the pheromone evaporation rate [33].

The parameters of the ACA-BS and IMACA-BS were optimized by repeating the experiments 10 times, and the parameter combinations leading to the highest average OAS were used for band selection. In the process of parameter

TABLE 4. Optimized parameters of the ACA-BS and IMACA-BS in the three experiments.

Dataset	Band selection method	α	β	ρ
Indian Pines	ACA-BS	4	6	0.3
	IMACA-BS	5	2	0.3
Pavia University	ACA-BS	4	6	0.4
	IMACA-BS	3	6	0.4
Botswana	ACA-BS	5	6	0.1
	IMACA-BS	1	7	0.2

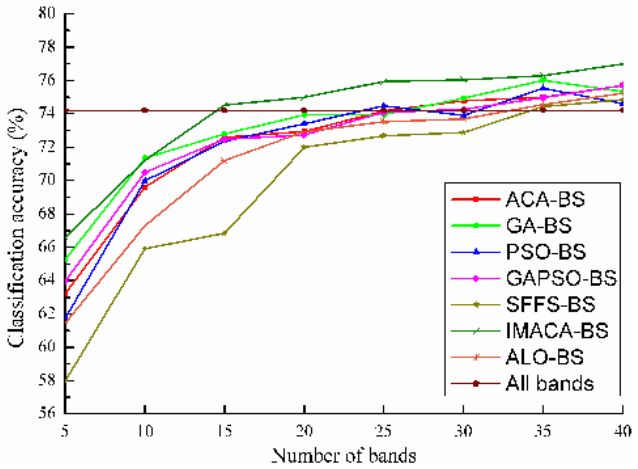


FIGURE 4. The OA of full-bands and the average OAS with different sizes (5, 10, 15, 20, 25, 30, 35, 40) of band subsets selected from indian dataset by IMACA-BS, ACA-BS, GA-BS, PSO-BS, GAPSO-BS, ALO-BS, and SFFS-BS.

optimization, a ''grid search'' approach was adopted: α and β varied from 1 to 10 with a step of 1, while ρ varied in the range of [0.1, 0.9] with a step of 0.1. the optimized parameter combinations of the ACA-BS and IMACA-BS for the indian pines, pavia university, and botswana datasets are listed in Table 4.

B. INDIAN PINES EXPERIMENT

The OA of full-bands and the average OAS of the algorithms using differently sized band subsets (the number of bands m varies from 5 to 40) are shown in figure 4. as shown in the figure, the OAs achieved by the band selection algorithms showed an overall increased accuracy when more bands were selected for classification. The proposed IMACA-BS always achieved the greatest OA, whereas the SFFS-BS generally presented the least value. The other algorithms achieved comparable OA values. The highest OA was acquired by the IMACA-BS, up to 76.99% for the case of $m = 40$; the second highest OA was obtained by the GA-BS excluding the cases of $m = 25$ and $m = 40$. In the cases of $m = 40$, the average OAs of GAPSO-BS and ACA-BS surpassed that of the GA-BS. Similarly, for the case of $m = 35$, the GA-BS and PSO-BS achieved higher OAs than the GAPSO-BS and ALO-BS.

For a more detailed comparison of the algorithms, the average OA AND per-class accuracy in the case of $m = 40$ are reported in Table 5, with the best OAs for each class highlighted in boldface. As illustrated in the table,

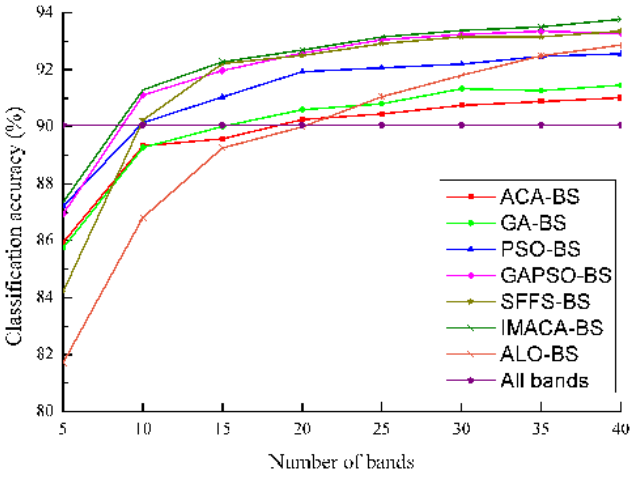


FIGURE 5. The OA of full-bands and the average OAS with different sizes (5, 10, 15, 20, 25, 30, 35, 40) of band subsets selected from pavia dataset by IMACA-BS, ACA-BS, GA-BS, PSO-BS, GAPSO-BS, ALO-BS, and SFFS-BS, respectively.

the IMACA-BS outperforms other band selection algorithms in discriminating complex classes. For instance, the most difficult to distinguish classes, Corn-mintill (class 5) and Soybean-clean (class 11) that were often misidentified as Corn (class 3), were better identified with the band subset selected by the IMACA-BS.

The mcNemar test was further adopted to analyze whether there is a significant difference in classification accuracy between the proposed method and benchmarks. The significance testing was carried out between the trials of the IMACA-BS and that of one benchmark with the same order, for example, the first trial of the IMACA-BS and GA-BS. the results (i.e. average p-values) of the significance testing are presented in Table 5. It can be seen from the table that all of the average p-values were lower than 0.05, demonstrating that the IMACA-BS achieved significantly better classification results in comparison with benchmark comparators.

C. PAVIA UNIVERSITY EXPERIMENT

the classification results of the full-bands and the differently sized band subsets ($m = 5$, 10, 15, 20, 25, 30, 35, 40) chosen by the algorithms are illustrated in Figure 5. In the case of $m = 5$, the classification accuracy of the selected band subsets was inferior to that of the full-bands (90.06%). The average OA achieved by each band selection method tended to increase with the increase of the band subset size, and the proposed IMACA-BS always obtained the best OA. Specifically, in the case of $m = 40$, the IMACA-BS achieved the best classification result, with an average OA of 93.78%. In most cases, the GAPSO-BS produced the second highest average OA. The performance of SFFS-BS was very close to the GAPSO-BS, followed by the PSO-BS and GA-BS. The least OAs were present by ALO-BS and ACA-BS for the case of $m \leq 20$ and $m > 20$, respectively. The average OA of ALO-BS surpassed that of PSO-BS, GA-BS, and ACA-BS for the case of $m = 40$.

TABLE 5. The overall accuracy (%), per-class accuracy (%), Kappa coefficients (K), and the significance testing results (p-value) between the proposed method and other algorithms in terms of classification accuracy with the Indian dataset ($m = 40$); class names are in Table 1.

Class No.	ACA-BS	GA-BS	GAPSO-BS	PSO-BS	IMACA-BS	SFFS-BS	Full-bands	ALO-BS
1	94.52±1.83	94.56±0.18	92.17±4.93	94.13±2.72	**95.65±4.82**	82.61	82.61	86.45±1.04
2	**77.12±3.45**	73.16±2.15	70.15±5.51	74.50±3.45	75.65±1.51	75.13	74.87	72.70±2.01
3	93.16±3.34	93.12±1.44	93.03±2.75	91.81±3.01	**94.09±3.21**	92.83	88.61	91.44±3.05
4	73.12±1.41	**73.30±4.05**	62.20±3.93	69.79±3.80	67.71±3.57	63.73	68.67	69.17±2.88
5	67.22±1.63	66.33±2.54	63.13±3.62	66.04±3.00	68.63±4.40	59.03	62.89	67.81±3.65
6	88.22±1.63	89.68±2.26	89.02±1.35	88.71±1.61	90.73±1.03	91.51	**92.75**	89.49±0.80
7	95.35±1.72	95.71±1.50	97.21±2.52	95.35±1.72	96.43±2.91	96.43	**100.00**	96.92±3.97
8	94.13±1.43	91.57±3.36	94.95±2.18	91.73±2.26	93.53±0.61	95.07	93.70	92.39±1.66
9	95.33±0.95	94.72±2.46	94.87±2.01	93.68±4.32	96.65±0.90	98.12	98.33	96.30±2.12
10	100	100	100	100	100.00	90.00	100.00	98.00±6.32
11	76.07±2.81	73.82±3.95	74.38±4.15	73.67±5.12	81.20±3.21	77.91	75.55	80.33±3.33
12	61.09±3.36	57.25±5.11	64.18±3.19	56.39±3.32	63.43±3.92	66.23	54.70	61.26±4.14
13	77.24±2.19	79.38±2.65	77.76±3.25	77.78±2.85	73.69±4.14	66.15	76.03	74.42±2.49
14	**98.81±1.06**	97.74±1.28	96.12±1.69	98.06±1.50	97.31±1.86	96.77	97.85	95.58±2.31
15	98.48±0.71	98.43±0.75	98.63±0.50	98.48±0.71	**98.90±0.46**	97.56	98.05	98.90±0.68
16	86.98±3.56	86.73±2.08	87.48±3.15	87.93±3.35	88.10±0.99	85.38	87.17	85.40±2.26
OA	75.71±0.93	75.34±1.57	75.71±2.12	74.61±1.60	**76.99±1.09**	74.84	74.20	75.25±1.62
K	0.75±0.01	0.75±0.02	0.75±0.02	0.74±0.01	**0.74±0.01**	0.71	0.71	0.72±0.02
p-value	0.0001±0.0002	0.0001±0.0003	0.007±0.014	0.0002±0.0002	--	0.0325	0.0243	0.005±0.002

TABLE 6. The overall accuracy (%), per-class accuracy (%), Kappa coefficients (K), and the significance testing results (p-value) between the proposed method and other algorithms in terms of classification accuracy with the Pavia university dataset ($m = 40$); class names are in Table 2.

Class No.	ACA-BS	GA-BS	GAPSO-BS	PSO-BS	IMACA-BS	SFFS-BS	Full-bands	ALO-BS
1	91.39±0.37	92.20±0.55	92.26±0.58	92.41±0.46	**92.73 ±0.41**	91.68	91.51	91.39±0.51
2	88.00±0.49	88.55±1.95	88.88±0.89	88.01±1.27	**89.90±0.61**	87.71	83.07	88.29±1.38
3	85.37±1.25	85.06±0.94	85.27±0.67	84.90±1.35	84.25±1.47	**85.49**	77.36	84.64±0.95
4	74.85±1.96	74.52±1.75	75.77±1.06	75.58±1.36	75.62±1.37	**76.04**	67.07	73.08±1.57
5	97.76±0.20	97.76±0.21	97.45±0.30	97.78±0.27	**97.87±0.11**	97.66	95.68	97.39±0.26
6	98.27±0.11	98.36±0.06	99.36±0.06	99.36±0.06	99.32±0.06	**99.41**	89.21	99.24±0.16
7	91.57±0.79	**91.69±0.06**	90.89±0.81	91.36±0.73	90.99±0.11	90.85	86.9	90.08±0.81
8	100	100	100	100	100	99.89	100	99.52±0.32
9	91.21±0.50	92.21±0.70	91.59±0.82	92.33±0.51	**93.14±0.27**	92.62	88.96	91.74±0.95
OA	91.02±0.88	91.46±0.41	93.28±0.21	92.56±0.12	**93.78±0.09**	93.36	90.06	92.87±0.26
K	0.91±0.01	0.91±0.01	0.91±0.01	0.91±0.02	**0.92±0.01**	0.91	0.87	0.90±0.003
p-value	0.0022±0.003	0.0037±0.008	0.0001±0.011	0.0423±0.015	--	0.0189	0.0265	0.0312±0.0092

To further investigate the effectiveness of the band selection algorithms, the average per-class classification accuracy and the corresponding standard deviation with the pavia dataset are presented in Table 6. As shown by the table, the IMACA-BS outperforms the benchmarks in terms of land use/cover class discrimination, especially for the complex classes. For example, the complicated bare soil (class 2) and meadows (class 5) were better identified by the IMACA-BS algorithm. the mcnemar test results were summarized and shown in Table 6. It can be seen that all of the average p-values were lower than 0.05, indicating that the IMACA-BS performed significantly better than other algorithms in terms of classification accuracy.

D. BOTSWANA EXPERIMENT

The classification accuracy with the full-bands and the differently sized (m varies over a range of 5-40) band subsets achieved by the six algorithms were illustrated in Figure 6. As shown in the figure, the classification accuracy (OA $=$ 85.26%) with the full-bands was lower than that of most band selection algorithms for the cases of $m > 5$. The average OA achieved by each band selection algorithm generally increased with the increase in band subset size. The proposed IMACA-BS always obtained the best accuracy, while the GAPSO-BS produced the second-best accuracy for most of the cases ($5 < m \leq 25$). The average OAs of the GA-BS and PSO-BS were close to each other (except for the cases of $m = 25$, and $m = 30$). In contrast, the average OA of the ALO-BS was the lowest for the cases of $m \leq 35$. In the case of $m = 40$, the IMACA-BS achieved the highest average OA (reaching up to 91.32%), whereas the ACA-BS obtained the lowest (only 89.05%).

The average per-class classification accuracy achieved by the algorithms is summarized in Table 7. The outperformance of the IMACA-BS is clearly demonstrated by the classification accuracy at class-level, especially for the Reeds1

TABLE 7. The overall accuracy (%), per-class accuracy (%), Kappa coefficients (K), and the significance testing results (p-value) between the proposed method and other algorithms in terms of classification accuracy with the Botswana university dataset ($m = 40$); Class names are in Table 3.

Class No.	ACA-BS	GA-BS	GAPSO-BS	PSO-BS	IMACA-BS	SFFS-BS	Full-bands	ALO-BS
1	98.51±0.11	98.48±0.11	98.51±0.00	98.51±0.00	**98.52±0.00**	98.42	94.81	98.14±0.00
2	92.57±1.20	92.77±1.32	92.77±0.66	92.77±0.66	91.09±1.16	**93.07**	83.16	91.35±1.30
3	96.29±0.51	96.37±1.41	**96.81±0.53**	96.05±0.43	96.41±0.56	94.02	86.45	96.06±1.00
4	93.58±0.49	95.44±1.62	95.20±1.29	95.62±1.61	94.42±0.40	**97.21**	88.37	93.48±1.91
5	71.26±1.63	71.71±2.44	71.19±1.32	72.34±2.14	**76.58±1.52**	73.23	69.88	70.92±2.40
6	83.68±1.39	83.94±1.88	84.05±1.12	84.31±2.56	**85.50±1.65**	84.39	82.7	82.91±3.47
7	95.59±0.19	96.10±0.64	96.06±0.35	95.56±0.35	95.37±1.32	**96.53**	88.41	95.04±0.68
8	92.77±1.10	94.53±1.12	94.67±1.22	94.63±0.67	**96.06±1.21**	93.60	90.14	92.69±1.43
9	84.89±0.65	85.28±1.51	85.31±1.16	85.71±1.78	**88.22±1.78**	85.35	81.21	83.73±1.15
10	93.87±0.71	92.25±1.78	**95.83±1.40**	94.07±2.08	95.16±2.36	92.74	90.72	92.72±3.09
11	90.19±1.32	90.03±1.51	91.72±1.21	90.23±1.49	**94.10±1.25**	92.79	84.26	90.20±1.93
12	92.92±0.57	**93.59±0.91**	92.21±1.20	92.54±0.59	91.71±0.88	91.71	86.74	92.89±1.21
13	88.47±1.42	89.29±1.40	**90.73±1.42**	90.66±1.07	89.18±1.33	88.06	86.19	87.00±2.74
14	84.10±1.90	85.15±2.50	85.52±2.38	85.78±2.00	84.21±2.14	**90.53**	78.94	84.60±2.77
OA	87.50±0.55	89.47±0.41	89.80±0.49	89.68±0.48	**91.32±0.47**	90.45	85.26	89.19±0.49
K	0.86±0.01	0.88±0.01	0.89±0.01	0.88±0.01	**0.90±0.01**	0.84	0.83	0.88±0.005
p-value	0.006±0.011	0.001±0.002	0.017±0.014	0.017±0.026	--	0.0134	0.0376	0.0154±0.003

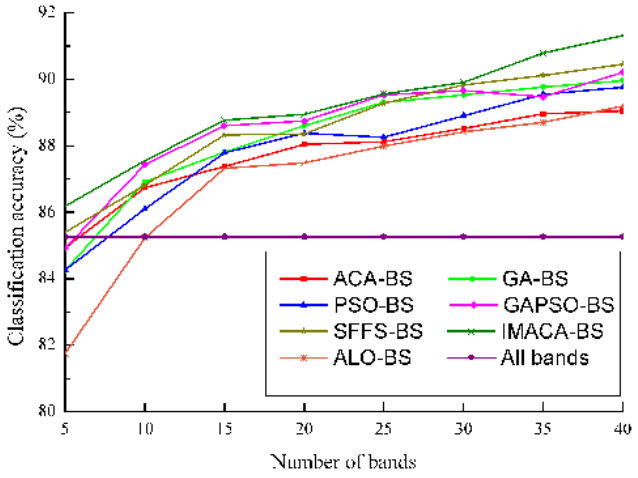


FIGURE 6. The OA of full-bands and the average OAS achieved by the band subsets with different band sizes (5, 10, 15, 20, 25, 30, 35 and 40) selected from the botswana dataset by the IMACA-BS, ACA-BS, GA-BS, PSO-BS, GAPSO-BS, ALO-BS, and SFFS-BS, respectively.

classification (class5) which was often misclassified as Floodplain grasses2 (class 4) and Riparian (class 6). With the bands selected by the IMACA-BS, the average classification accuracy of Reeds1 was improved to 76.58%, about a 3-5% increase compared with those of benchmarks.

Similar to previous experiments, the significance of the difference between the proposed IMACA-BS and other benchmarks was tested using the McNemar test at a significance level of 0.05. The testing results including average p-values and the standard deviations are presented in Table 7. It can be seen from the table that all of the average p-values were lower than the significance level. The classification results achieved by the IMACA-BS were, thus, statistically significantly better than those by other algorithms.

V. DISCUSSION

Band selection is essential for the effective utilization of heavily redundant bands in hyperspectral images (HSI). However, band selection from HSI data sets is not a trivial task, considering the huge solution space. The proposed IMACA-BS method has proved to be superior to the meta-heuristic ACA-BS, GA-BS, PSO-BS, GAPSO-BS, and the traditional SFFS over all three experiments. In the GA-BS, the cross and mutation operations are randomly operated on candidate solutions, which leads to very slow convergence [34]. In contrast, the PSO-BS converges rapidly due to the use of the global optimal solution and personal optima in the process of population evolution [35]. However, such an evolving strategy may make the solutions resemble each other. The particle swarm, thus, cannot explore new promising areas, and suffers from premature convergence [35]. Similarly, the traditional SFFS also tends to become trapped on local optima because of a lack of global observation on selected feature subsets. Our experiments illustrate (Tables 5, 6 and 7) that it is difficult to achieve promising band selection results with these band selection methods.

Though the ACA possesses advantages over the above-mentioned algorithms in terms of searching mechanism (e.g., strong positive feedback), the ACA suffers from slow convergence and local optima problems (resulting from loss of population diversity). Improvements are, thus, necessary to make the ACA suitable to solve the NP-hard HSI datasets band selection issue. In this research, a novel improved ant colony algorithm (IMACA) was proposed for band selection of HSI data. In comparison with the standard ACA, the IMACA overcomes the two shortcomings (i.e. slow convergence and local optima problems) of the standard ACA in solving the complicated HSI band selection problem. For the former, a pre-filter is adopted to improve the heuristic desirability of the ant colony system by reducing the randomness of the search ants at the

beginning of the search process. Besides, the similarity (i.e. the degree of redundancy) of the selected bands is measured and serves as the heuristic function in the IMACA-BS, which significantly reduces data redundancy. Such strategies markedly improve the searching efficiency of the ant colony, and thus significantly accelerate the convergence of the IMACA-BS. For the latter, a pseudo-random rule was applied to the IMACA-BS, which enables the paths with limited pheromone concentration to be selected as candidate solutions with a certain probability. At the same time, an adaptive information update strategy is employed to update the pheromone dynamically, which prevents pheromone on the potential optimal edges being volatilized. As demonstrated in our previous work, the computational complexity of the ACA-BS $O(P_s \times m \times n \times T \times C_0)$ is lower than that of other benchmarks (i.e. GA-BS, GAPSO-BS, PSO-BS, and SFFS-BS) [36]. In comparison to the ACA-BS, the IMACA-BS further decreases the computational complexity since the number of candidate bands (n) is lowered by using a pre-filter. Therefore, the proposed method is effective and efficient for practical usage.

The above-mentioned advantages ensure the proposed IMACA-BS will always select better band subsets compared to ACA-BS as well as other benchmarks, including GA-BS, GAPSO-BS, PSO-BS, and SFFS-BS. As demonstrated in the three experiments, the IMACA-BS consistently acquired the band subsets with the highest capacity to separate land cover classes in most of cases, regardless of the size of band subsets. With the selected bands, high classification accuracies were achieved for most land cover types, especially for those complicated and confusing land cover classes with high between-class spectral heterogeneity and similarity, e.g., Corn and Corn-mintill in Indian Pines Experiment, Meadows and Bare soil in Pavia University Experiment, and Reeds1 and Riparian in Botswana Experiment. Our experimental results demonstrate that the proposed IMACA-BS is capable of getting rid of suboptimal solutions and eventually approaching or reaching the global optimal solution.

The proposed method involves three parameters, each of which exerts a huge impact on the band selection results. Herein, a ''grid search'' approach was employed to achieve the optimal parameter combinations. Such a strategy enables the IMACA-BS to achieve encouraging results, but it is tedious and time-consuming. In contrast, the signal to noise ratio method (S/N) that can indicate how a solution changes among experiments [37], might be an alternative to efficiently search for the optimal parameters [38]–[40], which deserves further investigations. Although the proposed method considers the degree of redundancy between candidate band and selected band subsets to accelerate convergence, this might affect the final classification accuracy in some cases. Future studies may further improve the IMACA-BS by directly measuring the contribution of a candidate band to classification accuracy in the band selection process. Finally, it is usually difficult to determine the optimal number of selected bands for band selection algorithms. Fortunately, deep learning-based methods have become the new hot topic in a wide range of research fields (e.g. change detection [41] and transportation mode classification [42]), owing to their capacity to solve complex issues in an end-to-end fashion. Deep learning-based band selection methods, thus, deserve further investigation.

VI. CONCLUSION

In this study, an improved ant colony algorithm (IMACA-BS) for band selection from a hyperspectral image (HSI) was proposed, which enables the ant colony to avoid being trapped on local optima by accelerating the convergence and maintaining the population diversity. The effectiveness of the proposed IMACA-BS was tested with three public HSI data sets (Indian Pines, Pavia University and Botswana datasets), and compared with benchmark comparators using the overall classification accuracy of the SVM. The experimental results demonstrated that the proposed IMACA-BS method consistently produced the most accurate classification results and significantly outperformed the ant colony algorithm (ACA-BS), the genetic algorithm (GA-BS), the particle swarm optimization (PSO-BS), the hybridization of GA and PSO (GAPSO-BS), and the traditional sequential forward floating selection (SFFS-BS). As well, the proposed method achieved the greatest classification accuracies for most of the classes in each experiment, especially for those complex classes with similar spectral characteristics. The proposed method is, therefore, suggested to be an alternative for band selection from HSI data sets.

REFERENCES

[1] C. Huo, R. Zhang, and D. Yin, ''Compression technique for compressed sensing hyperspectral images,'' *Int. J. Remote Sens.*, vol. 33, no. 5, pp. 1586–1604, Mar. 2012.

[2] G. Hughes, ''On the mean accuracy of statistical pattern recognizers,'' *IEEE Trans. Inf. Theory*, vol. 14, no. 1, pp. 55–63, Jan. 1968.

[3] C. Sukawattanavijit, J. Chen, and H. Zhang, ''GA-SVM algorithm for improving land-cover classification using SAR and optical remote sensing data,'' *IEEE Geosci. Remote Sens. Lett.*, vol. 14, no. 3, pp. 284–288, Mar. 2017.

[4] L. Feng, A.-H. Tan, M.-H. Lim, and S. W. Jiang, ''Band selection for hyperspectral images using probabilistic memetic algorithm,'' *Soft Comput.*, vol. 20, no. 12, pp. 4685–4693, Dec. 2016.

[5] J. Feng, L. Jiao, F. Liu, T. Sun, and X. Zhang, ''Unsupervised feature selection based on maximum information and minimum redundancy for hyperspectral images,'' *Pattern Recognit.*, vol. 51, pp. 295–309, Mar. 2016.

[6] J. C. Ang, A. Mirzal, H. Haron, and H. N. A. Hamed, ''Supervised, unsupervised, and semi-supervised feature selection: A review on gene selection,'' *IEEE/ACM Trans. Comput. Biol. Bioinf.*, vol. 13, no. 5, pp. 971–989, Sep. 2016.

[7] K. Sechidis and G. Brown, ''Simple strategies for semi-supervised feature selection,'' *Mach. Learn.*, vol. 107, no. 2, pp. 357–395, Feb. 2018.

[8] Q. Du and H. Yang, ''Similarity-based unsupervised band selection for hyperspectral image analysis,'' *IEEE Geosci. Remote Sens. Lett.*, vol. 5, no. 4, pp. 564–568, Oct. 2008.

[9] V. Kumar and S. Minz, ''Feature selection: A literature review,'' *Smart Comput. Rev.*, vol. 4, no. 3, pp. 211–229, 2014.

[10] P. M. Narendra and K. Fukunaga, ''A branch and bound algorithm for feature subset selection,'' *IEEE Trans. Comput.*, vol. C-26, no. 9, pp. 917–922, Sep. 1977.

[11] H. Almuallim and T. G. Dietterich, ''Learning with many irrelevant features,'' in *Proc. 9th Nat. Conf. AI*, Anaheim, CA, USA, Jul. 1991, pp. 547–552.

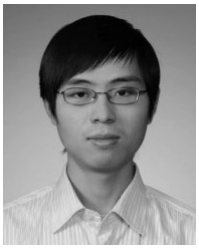
[12] J. Nascimento and J. Dias, "Does independent component analysis play a role in unmixing hyperspectral data?" *IEEE Trans. Geosci. Remote Sens.*, vol. 43, no. 1, pp. 175–187, Jan. 2005.
[13] A. Al-Ani, "Feature subset selection using ant colony optimization," *Int. J. Comput. Intell.*, vol. 2, no. 1, pp. 53–58, 2006.
[14] X. Wang, J. Yang, X. Teng, W. Xia, and R. Jensen, "Feature selection based on rough sets and particle swarm optimization," *Pattern Recognit. Lett.*, vol. 28, no. 4, pp. 459–471, Mar. 2007.
[15] S. Yu, S. De Backer, and P. Scheunders, "Genetic feature selection combined with composite fuzzy nearest neighbor classifiers for hyperspectral satellite imagery," *Pattern Recognit. Lett.*, vol. 23, nos. 1–3, pp. 183–190, Jan. 2002.
[16] Y. Zhong and L. Zhang, "A fast clonal selection algorithm for feature selection in hyperspectral imagery," *Geo-Spatial Inf. Sci.*, vol. 12, no. 3, pp. 172–181, Jan. 2009.
[17] H. Yang, Q. Du, and G. Chen, "Particle swarm optimization-based hyperspectral dimensionality reduction for urban land cover classification," *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 5, no. 2, pp. 544–554, Apr. 2012.
[18] S. Mirjalili, "The Ant Lion optimizer," *Adv. Eng. Softw.*, vol. 83, pp. 80–98, May 2015.
[19] J. Wang, J. Liu, and Y. Zhong, "A novel ant colony algorithm for assembly sequence planning," *Int. J. Adv. Manuf. Technol.*, vol. 25, nos. 11–12, pp. 1137–1143, Jun. 2005.
[20] B. Chen, L. Chen, and Y. Chen, "Efficient ant colony optimization for image feature selection," *Signal Process.*, vol. 93, no. 6, pp. 1566–1576, Jun. 2013.
[21] 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.
[22] H. R. Kanan and K. Faez, "An improved feature selection method based on ant colony optimization (ACO) evaluated on face recognition system," *Appl. Math. Comput.*, vol. 205, no. 2, pp. 716–725, Nov. 2008.
[23] F. Samadzadegan and T. Partovi, "Feature selection based on ant colony algorithm for hyperspectral remote sensing images," in *Proc. 2nd Workshop IEEE Hyperspectral Image Signal Process., Evol. Remote Sens. (WHISPERS)*, Reykjavik, Iceland, Jun. 2012.
[24] Y. Chen, D. Miao, and R. Wang, "A rough set approach to feature selection based on ant colony optimization," *Pattern Recognit. Lett.*, vol. 31, no. 3, pp. 226–233, 2010.
[25] 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, pp. 39–48, Mar. 2019.
[26] L. Bruzzone and S. B. Serpico, "A technique for feature selection in multiclass problems," *Int. J. Remote Sens.*, vol. 21, no. 3, pp. 549–563, 2000.
[27] P. Ghamisi, M. S. Couceiro, and J. A. Benediktsson, "A novel feature selection approach based on FODPSO and SVM," *IEEE Trans. Geosci. Remote Sens.*, vol. 53, no. 5, pp. 2935–2947, May 2015.
[28] M. Dorigo, G. D. Caro, and L. M. Gambardella, "Ant algorithms for discrete optimization," *Artif. Life*, vol. 5, no. 2, pp. 137–172, Apr. 1999.
[29] H. W. Quan and G. J. Wen, "Subspace genetic algorithm for TSP," *Math. Theory Appl.*, vol. 22, no. 1, pp. 36–39, 2002.
[30] M. H. Aghdam, N. Ghasem-Aghaee, and M. E. Basiri, "Text feature selection using ant colony optimization," *Expert Syst. Appl.*, vol. 36, no. 3, pp. 6843–6853, 2009.
[31] R. K. Sivagaminathan and S. Ramakrishnan, "A hybrid approach for feature subset selection using neural networks and ant colony optimization," *Expert Syst. Appl.*, vol. 33, no. 1, pp. 49–60, Jul. 2007.
[32] M. Dorigo and C. Blum, "Ant colony optimization theory: A survey," *Theor. Comput. Sci.*, vol. 344, nos. 2–3, pp. 243–278, Nov. 2005.
[33] M. Dorigo, E. Bonabeau, and G. Theraulaz, "Ant algorithms and stigmergy," *Future Gener. Comput. Syst.*, vol. 16, no. 8, pp. 851–871, Jun. 2000.
[34] P. Ghamisi and J. A. Benediktsson, "Feature selection based on hybridization of genetic algorithm and particle swarm optimization," *IEEE Geosci. Remote Sens. Lett.*, vol. 12, no. 2, pp. 309–313, Feb. 2015.
[35] H. Li, S. Zhang, and C. Zhang, "A novel unsupervised Levy flight particle swarm optimization (ULPSO) method for multispectral remote-sensing image classification," *Int. J. Remote Sens.*, vol. 38, no. 23, pp. 6970–6992, 2017.
[36] X. Ding, S. Zhang, and H. Li, "A restrictive polymorphic ant colony algorithm for the optimal band selection of hyperspectral remote sensing images," *Int. J. Remote Sens.*, vol. 41, no. 3, pp. 1093–1117, 2020.
[37] E. B. Tirkolaee, M. Alinaghian, and A. A. R. Hosseinabadi, "An improved ant colony optimization for the multi-trip capacitated arc routing problem," *Comput. Elect. Eng.*, vol. 77, pp. 457–470, Jul. 2019.
[38] A. Goli, E. B. Tirkolaee, and B. Malmir, "A multi-objective invasive weed optimization algorithm for robust aggregate production planning under uncertain seasonal demand," *Computing*, vol. 101, no. 6, pp. 499–529, 2019.
[39] T. E. Babaee, A. Goli, and M. Pahlevan, "A robust bi-objective multi-trip periodic capacitated arc routing problem for urban waste collection using a multi-objective invasive weed optimization," *Waste Manage. Res.*, vol. 37, no. 11, pp. 1089–1101, 2019.
[40] E. B. Tirkolaee, I. Mahdavi, M. M. S. Esfahani, and G.-W. Weber, "A hybrid augmented ant colony optimization for the multi-trip capacitated arc routing problem under fuzzy demands for urban solid waste management," *Waste Manag. Res.*, vol. 38, no. 2, pp. 156–172, Feb. 2020, doi: 10.1177/0734242x19865782.
[41] N. Lv, C. Chen, T. Qiu, and A. K. Sangaiah, "Deep learning and superpixel feature extraction based on contractive autoencoder for change detection in SAR images," *IEEE Trans. Ind. Informat.*, vol. 14, no. 12, pp. 5530–5538, Dec. 2018.
[42] R. Zhang, P. Xie, C. Wang, G. Liu, and S. Wan, "Classifying transportation mode and speed from trajectory data via deep multi-scale learning," *Comput. Netw.*, vol. 162, Oct. 2019, Art. no. 106861.

XIAOHUI DING received the B.Sc. degree in cartography and geography information system from Jiangxi Normal University, Nanchang, China, in 2013, and the M.S. and Ph.D. degrees in cartography and geography information system from the Northeast Institute of Geography and Agroecology, CAS, Changchun, China, in 2016 and 2019, respectively. Since 2019, he is a Postdoctoral Fellow in the Guangzhou Institute of Geography, Guangzhou, China. His major research interests include remote sensing image analysis, artificial intelligence, and GIS.

HUAPENG LI received the B.Sc. degree in GIS from the Northeast Forestry University, Harbin, China, in 2007, and the Ph.D. degree from the Northeast Institute of Geography and Agroecology, CAS, Changchun, China, in 2012. He is currently an Assistant Professor with the Northeast Institute of Geography and Agroecology, CAS. He has published over ten articles in peer-reviewed international scientific journals. His major research interests include geospatial data mining, artificial intelligence, machine learning (especially deep learning), and remotely sensed image analysis.

JI YANG received the B.Sc. and the M.S. degrees in photogrammetry and remote sensing from Wuhan University, Wuhan, China, in 2004 and 2006, respectively, and the Ph.D. degree from the University of Chinese Academy of Sciences, Beijing, China, in 2018. Since 2018, he took the role of the Director of the Guangzhou Institute of Geography. His major research interests include UAV remote sensing and spatio-temporal big data analysis.

PATRICIA DALE received the B.A. (Hons) degree in geography from Southampton University, U.K., in 1964, the M.Soc.Sc. degree in geography and the L.L.B. (laws) from The University of Queensland, Australia, in 1979, in 1993, respectively, and the Ph.D. in ecology degree from the Griffith University, Brisbane, and 1997. Since 1980, she has been researching land cover ecology at Griffith University, using a variety of remote sensing techniques to assess land cover characteristics and to monitor change. She is currently an Emeritus Professor with Griffith University with continued interest in remote sensing of environment. She has supervised postgraduate student's remote sensing research.



XIANGCONG CHEN received the B.Sc. and M.S. degrees in geodesy and survey engineering from Wuhan University, Wuhan, China, in 2006 and 2008, respectively. Since 2008, he took the role of the Director of the Harbin Institute of Geotechnical Investigation and Surveying, Harbin, China. His research interest includes remote sensing and application of geography information systems.

CHUNLEI JIANG received the B.Sc. degree in physics from Jilin Normal University, Siping, China, in 2003, and the M.S. degree in biological information from Sichuan University, Chengdu, China, in 2010, and the Ph.D. degree in cartography and geography information system from the Northeast Institute of Geography and Agroecology, CAS, in 2016. Since 2016, he has been a Research Assistant Professor with Jilin Jianzhu University, Changchun, China. His research interests include remote sensing and 3D geography information systems.

SHUQING ZHANG received the Ph.D. degree from the Northeast Institute of Geography and Agroecology, CAS, Changchun, China, in 2001. He is currently a Professor and the Head of GIS Research Group, Northeast Institute of Geography and Agroecology, CAS, with a specific interest in GIS (3D/4D) modeling, landscape change analysis, and machine learning. He has authored over 30 peer-reviewed articles in international scientific journals.