symmetry

MDPI

*Article*

# A Band Selection Approach for Hyperspectral Image Based on a Modified Hybrid Rice Optimization Algorithm

**Zhiwei Ye [1,2,3,*], Wenhui Cai [1], Shiqin Liu [1], Kainan Liu [4], Mingwei Wang [1] and Wen Zhou [1]**

[1] School of Computer Science, Hubei University of Technology, Wuhan 430068, China; wenhuicai89@163.com (W.C.); liushiqin647@hbut.edu.cn (S.L.); wmwscola@hbut.edu.cn (M.W.); zw_mmwh@whu.edu.cn (W.Z.)
[2] Fujian Provincial Key Laboratory of Data Intensive Computing, Quanzhou 362000, China
[3] Key Laboratory of Intelligent Computing and Information Processing, Quanzhou 362000, China
[4] High Performance Computing Academician Workstation of Sanya University, Sanya 572022, China; kainanliu@sanyau.edu.cn
* Correspondence: weizhiye121@163.com

**Abstract:** Hyperspectral image (HSI) analysis has become one of the most active topics in the field of remote sensing, which could provide powerful assistance for sensing a larger-scale environment. Nevertheless, a large number of high-correlation and redundancy bands in HSI data provide a massive challenge for image recognition and classification. Hybrid Rice Optimization (HRO) is a novel meta-heuristic, and its population is approximately divided into three groups with an equal number of individuals according to self-equilibrium and symmetry, which has been successfully applied in band selection. However, there are some limitations of primary HRO with respect to the local search for better solutions and this may result in overlooking a promising solution. Therefore, a modified HRO (MHRO) based on an opposition-based-learning (OBL) strategy and differential evolution (DE) operators is proposed for band selection in this paper. Firstly, OBL is adopted in the initialization phase of MHRO to increase the diversity of the population. Then, the exploitation ability is enhanced by embedding DE operators into the search process at each iteration. Experimental results verify that the proposed method shows superiority in both the classification accuracy and selected number of bands compared to other algorithms involved in the paper.

**Keywords:** hyperspectral image; band selection; hybrid rice optimization algorithm; opposition-based learning; differential evolution



**Citation:** Ye, Z.; Cai, W.; Liu, S.; Liu, K.; Wang, M.; Zhou, W. A Band Selection Approach for Hyperspectral Image Based on a Modified Hybrid Rice Optimization Algorithm. *Symmetry* **2022**, *14*, 1293. https://doi.org/10.3390/sym14071293

Academic Editor: Mihai Postolache

Received: 31 May 2022
Accepted: 17 June 2022
Published: 22 June 2022



## 1. Introduction

Recently, hyperspectral remote sensing has been broadly and successfully applied in urban planning [1], precision agriculture [2], environmental monitoring [3] and other fields with the constantly increased spectral resolution of sensors. Hyperspectral remote sensing combines spectral features with spatial images that can accurately identify and detect ground objects, which provides strong technical support for ground feature extraction [4]. However, hyperspectral image (HSI) obtained in hundreds of narrow and contiguous bands from visible to infrared areas of the electromagnetic spectrum are characterized by high-dimensional space and a large number of spectral bands [5], which makes the processing and analysis of HSI a challenging task. Therefore, dimensionality reduction becomes a crucial task for hyperspectral data analysis [6].

Feature extraction and feature selection are two typical dimension-reduction methods. The original hyperspectral datasets are transformed into a low-dimensional and less-redundant feature space by feature extraction and common techniques such as independent component analysis (ICA) [7], principal component analysis (PCA) [8], and local linear embedding (LLE) [9]. Although these methods can extract valuable features from HSI datasets, they often lose physical information of the original data during the process of

data compression [10]. In contrast, feature selection can select the feature subset with the most information while preserving the physical meaning of the original data, which is an important and popular method for reducing dimensions [11]. In the traditional filter methods, the feature subset is built independently of the classifier or classification algorithm and can be evaluated based on different measures such as distance measures, correlation measures and information measures [12], while wrapper methods use the classifier model to estimate feature subsets. Although filter methods are computationally simple and fast [13], they are generally less accurate than wrapper methods because they are not guided by classifiers [14]. In general, feature-selection methods can be divided into supervised and unsupervised according to the availability of sample tags [15]. Unsupervised methods can select a subset of bands without class labels, but they tend to be unstable and biased due to the lack of prior information [16]. In comparison, supervised methods tend to obtain better feature-selection results with the assistance of class labels.

Supervised feature-selection methods include three search strategies: exhaustive search, sequential search, and random search [10]. Exhaustive search requires enumerating all possible combinations of features [17], which results in unacceptable time complexity for HSI. Sequential search contains sequential forward search (SFS), sequential backward search (SBS), and sequential floating forward search (SFFS) [18]. These methods require much computation while tending to get stuck to the local optima, and it is difficult to perform well for the existence of bands with strong correlation in HSI [19]. By contrast, random search introduces randomness into the search process to distance from the local optima and deliver promising results with higher efficiency. Recently, a number of nature-inspired stochastic search algorithms have been extensively utilized for feature selection based on their strong search ability in the large-scale space [20]. These include genetic algorithm (GA) [21], differential evolution (DE) algorithm [22], particle swarm optimization (PSO) [23], gray wolf optimizer (GWO) [24], cuckoo search (CS) algorithm [25], artificial bee colony (ABC) algorithm [26] and whale optimization algorithm (WOA) [27], which may have superior performance in dealing with feature-selection problems.

For HSI band selection, Nagasubramanian et al. [28] used GA to select the optimal subset of bands and support vector machine (SVM) to classify the infected and healthy samples. Additionally, the classification accuracy was replaced by F1-Score to alleviate the skewness caused by unbalanced datasets. The results showed that the bands chosen by this approach were more informative compared to RGB images. Xie et al. [29] proposed a band selection method based on ABC algorithm and enhanced subspace decomposition to apply in HSI classification. Subspace decomposition was realized by computing the relevance between adjacent bands, and ABC algorithm was guided by enhanced subspace decomposition and maximum entropy to optimize the combination of selected bands, which provided high classification accuracy compared with six related techniques. Wang et al. [30] proposed a wrapper feature-selection approach based on improved ant lion optimizer (ALO) and wavelet SVM to reduce the dimension of HSI. Lévy flight was used to help ALO jump out of local optimum and the wavelet SVM was introduced to improve the stability of classification result. The results showed that the proposed method can provide satisfactory classification accuracy in fewer frequency bands. Subsequently, Wang et al. [31] designed a new band selection method using chaos operation to set corresponding indices for the top three gray wolves in GWO to improve the optimization ability of GWO, and experimental results demonstrated that a suitable band subset can be obtained and superior classification accuracy can be achieved by this approach. Kavitha and Jenifa [32] used Discrete Wavelet transform with eight taps and four taps for extracting the important features and applied PSO algorithm for searching the optimal band subsets and utilized SVM as a classifier to classify HSI effectively. Medjahed et al. [33] introduced a novel band selection framework based on binary CS algorithm. The experiment compared the optimization ability of CS under two different objective functions and proved that it could obtain more excellent results than relevant approaches by adopting a few instances for training. Su et al. [34] proposed a modified firefly algorithm (FA) to deal with the band selection problem by

optimizing the minimum values of objective function, which outperformed superior results than SFS and PSO. In essence, band selection is a NP hard problem, as if the number of the bands increases, the above algorithms may suffer from premature convergence and even optimization stagnation.

Hybrid rice optimization (HRO) [35] is a newly proposed nature-inspired algorithm and has been successfully applied to image processing and knapsack problem because of its simple structure and strong optimization ability. For example, Liu et al. [36] presented an image segmentation method that used HRO to find the fittest multi-level thresholds by using Renyi's entropy as the fitness function, and experiments proved that HRO prevailed over the other six commonly used evolutionary algorithms on most metrics. Su et al. [37] designed two different hybrid models for the complex large-scale 0–1 knapsack problem by using novel combinations of improved HRO and binary ant colony algorithm, which achieved better performance on different size datasets. In addition, Ye et al. [38] regarded the band selection problem as a combinatorial optimization problem and employed binary HRO to select the optimal band set for HSI, which obtained good results in classification precision and execution efficiency. Although HRO algorithm has contributed to acquiring satisfactory results, primary HRO performs the exploitation of the current best solution during each search process inadequately.

Recently, DE algorithm has been successfully combined with other swarm intelligence algorithms for solving diverse optimization problems. Tubishat et al. [39] employed evolutionary operators from DE algorithm to help each whale seek better positions and improve the local search capability of WOA for feature selection in sentiment analysis. Jadon et al. [40] proposed a hybrid DE algorithm with ABC algorithm to enhance the convergence and the balance between exploration and exploitation. Houssein et al. [41] hybridized the adaptive guided DE algorithm with slime mold algorithm for combinatorial optimization problems, which verified that evolutionary operators could boost the local search capability of swarm agents. Hence, a modified HRO (MHRO) based on opposition-based learning (OBL) strategy and differential evolution (DE) operators is proposed to overcome the disadvantages of standard HRO in the paper. The main contributions of this paper are concluded as follows:

(1) OBL strategy is introduced to enhance the diversity of the initial population and accelerate the convergence of MHRO;
(2) DE operators are embedded into the search process of MHRO to enhance the local exploitation ability;
(3) The MHRO algorithm is applied in band selection and its performance is demonstrated on standard HSI datasets.

The remainder of the paper is organized as follows: Section 2 briefly gives a fundamental overview of the related technique and standard HRO algorithm. The methodology and the specific workflow of the proposed band selection approach are introduced in Section 3. Section 4 presents the experimental results and comparative studies. At last, conclusions and future work are summarized in the final section.

## 2. Background

### *2.1. Overview of Hybrid Rice Optimization Algorithm*

HRO algorithm is a meta-heuristic algorithm that simulates the breeding process of three-line hybrid rice. At each iteration, the rice seed population is sorted by the fitness from superior to inferior and divided into three sub-populations. According to self-equilibrium and symmetry, each sub-population is designed as an equal number of individuals. The individuals in the top third of the fitness ranking are selected into the maintainer line, the bottom third as the sterile line, and the remaining belong to the restorer line. The algorithm consists of three stages: hybridization, selfing, and renewal.

#### 2.1.1. Hybridization

Hybridization is performed to renew the rice seed genes in the sterile line. Two kinds of rice seeds presented to reconstruct one new individual are randomly chosen in the maintainer line and sterile line, respectively. If the new rice seed is superior to the current one, the current rice seed will be replaced by the new one. The new gene by hybridizing is shown in Equation (1).

$$X_{new(i)}^{k} = \frac{r_1 X_{s,r}^{k} + r_2 X_{m,r}^{k}}{r_1 + r_2} \tag{1}$$

where $X_{new(i)}^{k}$ represents the new $k$-th gene of the $i$-th rice seed in sterile line, $X_{s,r}^{k}$ is the $k$-th gene of a individual randomly selected from the sterile line, $X_{m,r}^{k}$ is the $k$-th gene of a individual randomly selected from the maintainer line, $r_1$ and $r_2$ are random numbers between $[-1, 1]$.

#### 2.1.2. Selfing

Selfing is the behavior that optimizes the gene sequence of rice seeds in the restorer line, which makes rice seeds gradually approach the best one, and the updated Equation is shown in (2).

$$X_{new(i)} = rand(0,1) \cdot (X_{best} - X_{j,r}) + X_i \tag{2}$$

where $X_{new(i)}$ is the new individual produced by selfing of the $i$-th restorer, $X_{best}$ represents the current optimal solution and $X_{j,r}$ is the $j$-th individual randomly selected from the restorer line ($i \neq j$). If the new individual is superior to the old individual, the old individual is replaced by the new and the current self-crossing number ($t_i$) is set to 0, otherwise the $t_i = t_i + 1$.

#### 2.1.3. Renewal

This stage is a reset operation for rice seeds in the restorer line that has not been updated for $t_{\max}$ consecutive times (i.e., reaching the maximum selfing time), and the renewal strategy is shown in Equation (3).

$$X_{new(i)} = X_i + rand(0,1) \cdot (R_{\max} - R_{\min}) + R_{\min} \tag{3}$$

where $X_{new(i)}$ is the new individual produced by renewal of the $i$-th restorer, $R_{\max}$ is the upper bound of the search space and $R_{\min}$ is the lower bound.

In summary, the flow of HRO is described in Algorithm 1.

**Algorithm 1.** Pseudo-Code of HRO

**Input:** the predefined parameters of HRO
**Output:** the global best solution and its fitness function value
**Initialize** the rice seed population randomly
**Initialize** $t_i = 0, k = 0$
**While** ($k$ < maximum number of iterations)
  Calculate the fitness function for each rice seed
  Divide the rice seeds into three lines
  **for** each rice seed in the sterile line
    Randomly select corresponding rice seeds in the sterile line and in the maintainer line
    The new gene is obtained by Equation (1)
    **if** the new rice seed is better
      Update the current rice seed
    **end if**
  **end for**
  **for** each rice seed in the restorer line
    **if** $t_i < t_{\max}$
      The new rice seed is obtained by Equation (2)

```
18:        if the new rice seed is better
19:          Update the rice seed
20:          t_i = 0
21:        else
22:          t_i = t_i + 1
23:        end if
24:      else
25:        The rice seed is renewed by Equation (3)
26:      end if
27:    end for
26:    k = k + 1
27: end while
```

### *2.2. The Binary Coding*

In general, data can be divided into two different types, that is, continuous or discrete. The basic HRO is presented for optimization problems with continuous search space. However, band selection for HSI is regarded as a discrete optimization problem, which is difficult to be solved by adopting the standard HRO. For binary coding, each individual in HRO is represented by a binary string where each element is only limited to 0 or 1. In order to solve the band selection problem, the continuous value of each candidate solution in the rice seeds population must be mapped to a probability value taking 0 or 1. Therefore, a sigmoid function is used to achieve data transform in the paper and is given as Equations (4) and (5).

$$S(x) = \frac{1}{1 + e^{-x}} \tag{4}$$

$$X_i^k = \begin{cases} 1, \ S(X_i^k) > 0.5 \\ 0, \ else \end{cases} \tag{5}$$

where $x$ is a real number, $X_i^k$ represents the $k$-th gene of the $i$-th new rice seed.

### *2.3. The Opposition Based Learning*

Generally, a good initial position of the population individual can accelerate the convergence speed of the algorithm. If the initial guess tends to be far from the position of the unknown optimal solution, the algorithm converges more slowly. OBL strategy can consider the current solution and its opposite solution to improve the diversity of the population. Then, the OBL method selected the fittest solutions from all initial solutions as the initial population, which can effectively broaden the search space of the algorithm.

**Definition 1.** *Let $x$ be a real number between $lb$ and $ub$, the opposite number $\widetilde{x}$ of $x$ is calculated as Equation (6).*

$$\widetilde{x} = lb + ub - x \tag{6}$$

where $lb$ and $ub$ are the lower and upper bounds of the search space, respectively. Similarly, the opposite number can also be used in multidimensional space.

**Definition 2.** *Let $x = \{x_1, x_2, \ldots, x_D\}$ be a point in D-dimensional space, where $x_i \in [lb_i, ub_i]$. The opposite point $\widetilde{x} = \{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_D\}$ can be defined in Equation (7).*

$$\widetilde{x}_i = lb_i + ub_i - x_i \tag{7}$$

**Definition 3.** *Let $x = \{x_1, x_2, \ldots, x_D\}$ be a bit string in D-dimensional space, where $x_i$ represents 0 or 1. The incomplete opposite point $\widetilde{x} = \{\widetilde{x}_1, \widetilde{x}_2, \ldots, \widetilde{x}_D\}$ can be expressed as Equation (8).*

$$\widetilde{x}_i = \begin{cases} 1 - x_i, \ rand(0,1) > r \\ x_i, \ else \end{cases} \tag{8}$$

where $r$ represents the proportion of taking the opposite value, and $r \in [0,1]$. It takes $r$= 0.5 in the paper.

### *2.4. Differential Evolution*

DE algorithm, which includes mutation, crossover and selection operators, is a simple and potential method to solve optimization problems.

#### 2.4.1. Mutation

The aim of this step is to form a new vector by randomly selecting three different target vectors in the population. In each iteration, a mutant vector $V_i$ ($i = 1, \ldots, N$, $N$ is the population size) is generated using Equation (9).

$$V_i = X_{r1} + F \cdot (X_{r2} - X_{r3}) \tag{9}$$

where $X_{r1}$, $X_{r2}$ and $X_{r3}$ are random solution vectors selected from the population and $F$ is the mutation factor.

#### 2.4.2. Crossover

After mutation operation, the mutant vector $V_i$ will crossover with its corresponding target vector $X_i$. The crossover process is defined as Equation (10).

$$U_i^j = \begin{cases} V_i^j, \ if \ rand(0,1) \le CR \ or \ j = j_{rand} \\ X_i^j, else \end{cases} \tag{10}$$

where $j = 1, \ldots, D$ and $D$ represents the dimension of the problem. $CR$ is the crossover rate and $j_{rand}$ is a randomly chosen integer within $[1, D]$.

#### 2.4.3. Selection

The selection process is to evaluate the fitness function of the target vector $X_i$ and the trial vector $U_i$ obtained after crossover operator, and the better vector will remain in the next generation. The selection strategy is given as Equation (11).

$$X_i = \begin{cases} U_i, if \ f(U_i) < f(X_i) \\ X_i, else \end{cases} \tag{11}$$

where $f(U_i)$ and $f(V_i)$ are the fitness value of vectors $U_i$ and $V_i$, respectively.

## 3. The Proposed Band Selection Method

To overcome the disadvantages of the primary HRO algorithm, two strategies are used to enhance the performance of HRO for handling the band selection problem. The main steps of the proposed technique are described in the following subsections.

### *3.1. The Coding Scheme*

The key factor to handling the band selection issue is to make an appropriate mapping between the problem solution and algorithm coding. For band selection of HSI, each band has two candidate states of being selected or not being selected, which is suitable to be represented by binary coding. In HRO, each gene bit is represented by "1" or "0", where "1" means that the corresponding band is selected and will be utilized for training, and "0" represents that the corresponding band is not chosen. Supposing that HSI contains ten bands, the binary coding of MHRO is "1001100101". That is, the 1st, 4th, 5th, 8th and 10th bands will be selected to complete the subsequent classification task.

### *3.2. The Objective Function*

Further, the proposed band selection method is developed to minimize the fitness function or the objective function by adopting MHRO algorithm. The main purpose of this method is to select the bands with the most informative subset from the original bands, so as to maximize the classification accuracy. Accordingly, SVM is adapted to conduct the classification on the HSI datasets, and the classification accuracy is selected as part of the objective function. In band selection technique, classification accuracy is an important measure metric, but how to reduce the number of redundant bands is also one of the most crucial goals. Therefore, the objective function as shown in Equation (12) is utilized in the paper.

$$Fitness = \alpha \cdot (1 - \mathrm{OA}) + (1 - \alpha) \cdot \frac{n_s}{n_c} \tag{12}$$

where *Fitness* denotes the fitness value, OA represents the overall classification accuracy and its concept is described in Appendix A.1. Note that $n_c$ and $n_s$ are the entire and the selected number of bands. $\alpha$ is a weight factor that balances classification accuracy and selected number of bands. It takes $\alpha = 0.99$ in the paper.

### *3.3. The Implementation of the Proposed MHRO*

The proposed band selection method is easy to implement, and its idea is to choose the optimal band subset with satisfactory classification results. Two improvements contained in the proposed algorithm MHRO are presented as shown in Figure 1. The first improvement is to adopt OBL in the initialization stage, whose aim is to improve the population diversity. The second improvement is the combination of DE operators and binary HRO algorithm, which improves the local search ability of the algorithm. The main procedure of these strategies utilized in MHRO is described as follows:

**OBL**: In the stage of population initialization, the position of each rice seed is randomly generated in the specified space. Then, a new population is formed by generating the corresponding opposite individual for each rice seed in the initial population by using OBL mechanism. Next, the individuals in the initial and new populations are sorted according to their fitness value, and the top $N$ individuals are selected to enter the final population. The main steps to initialize the population by OBL are as follows:

(1) Initialize the location of each rice seed randomly, Let $X_i = \{x_{i1}, x_{i2}, \ldots, x_{ij}, \ldots, x_{iD}\}$ be the $i$-th rice seed in the initial population $X$, where $i = 1, 2, \ldots, N$ and $j = 1, 2, \ldots, D$. $N$ denotes the population size and $D$ represents the dimension of the problem;
(2) A new population $OX$ was obtained by using the Equation (8) for each rice seed in the population $X$;
(3) The $N$ fittest individuals are chosen from the set $\{\mathrm{X} \cup \mathrm{OX}\}$ to constitute the new initial population of the MHRO algorithm.

**DE operators**: In HRO, only individuals in sterile and restorer lines are updated, while maintenance lines are ignored, which reduces the search performance of the algorithm on high-dimensional band selection. Therefore, DE evolution operators are applied to the genetic sequences of each rice seed in the maintainer line to find better rice seeds by using Equations (9)–(11). In order to degrade the possibility of falling into the local optimum, the mutation factor $F$ in Equation (9) is set as a random number between 0 and 1, where $X_{r1}$, $X_{r2}$ and $X_{r3}$ are randomly selected individuals in the maintainer line. If the fitness value of the newly generated trial solution is better than the current individual, the current individual will be replaced. Otherwise, it is not replaced.

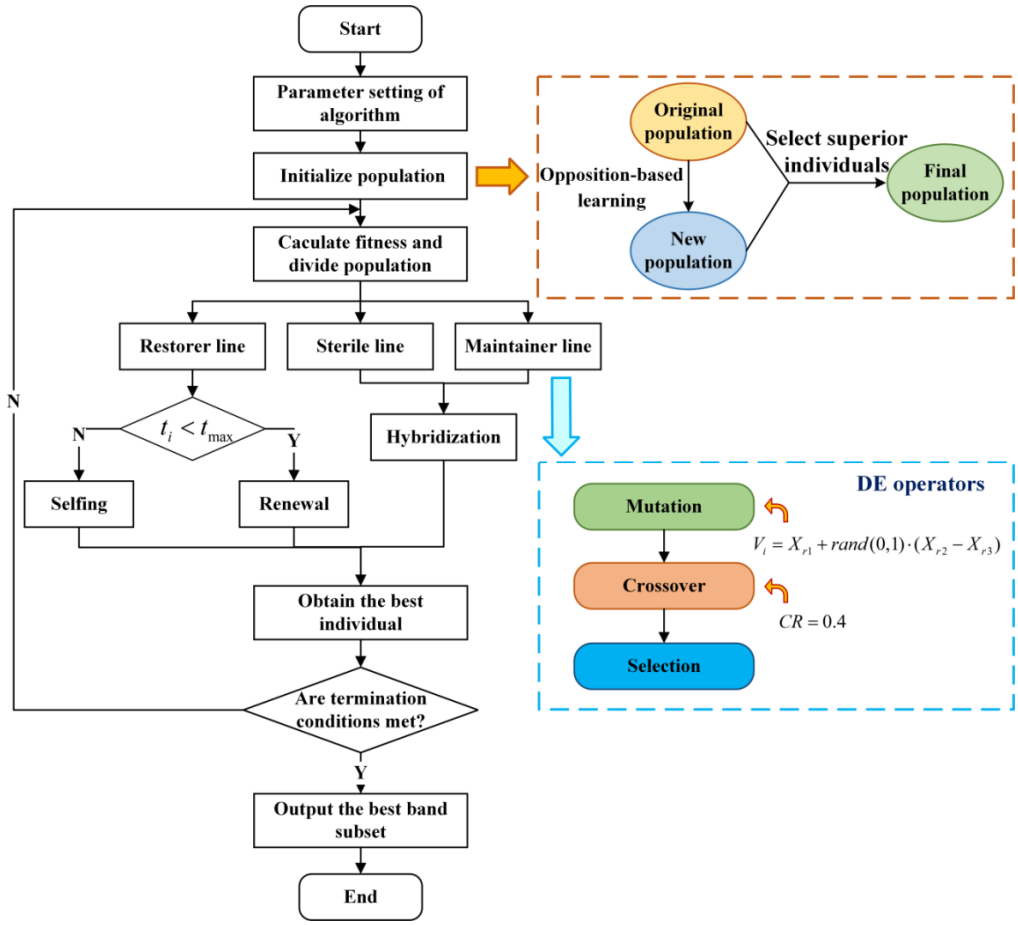



**Figure 1.** The overall flowchart of MHRO.

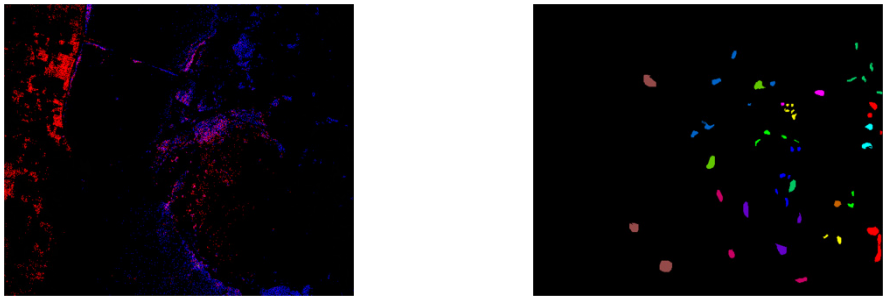
## 4. Experimental Results and Discussions

In this section, two sets of experiments are carried out to compare the proposed approach MHRO with multiple swarm intelligence algorithms and filter-based feature-selection methods. All the algorithms in these experiments are implemented by the language of python 3.9 and run on a PC with Intel(R) Core(TM) i7-10700 @ 2.9 GHz CPU and 16 GB memory under Windows 10 operating system.

### *4.1. Datasets Description*

Five public HSI datasets are employed to evaluate the performance of the proposed band selection method, including the Kennedy Space Center (KSC), Botswana, Indian Pines, Salinas and Pavia University datasets.

**KSC**: KSC was captured by AVIRIS sensor over the Kennedy Space Center in Florida. The size of the KSC image is 512 × 614 pixels from 0.4 to 2.5 µm. The data contain 224 bands and have a spatial resolution of 18 m. The noisy bands of low signal-to-noise or water absorption are dropped, and the remaining 176 bands are utilized for analysis. Figure 2 shows the false-color image of the KSC dataset and the corresponding ground truth image. Details of the category information are listed in Table 1.



(**a**) (**b**)

**Figure 2.** (**a**) KSC HSI. (**b**) Ground truth.

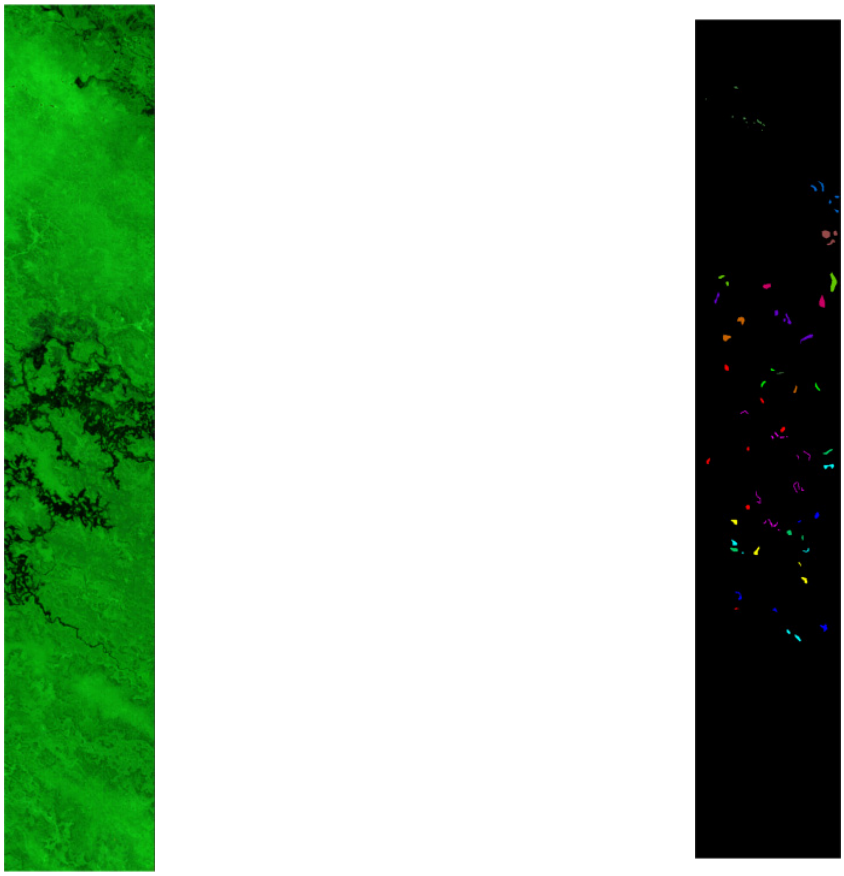
**Table 1.** Detailed category information in KSC dataset.

| Class Number | Class Name | Number of Samples |
|---|---|---|
| 1 | Scrub | 761 |
| 2 | Willow swamp | 243 |
| 3 | Cabbage palm hammock | 256 |
| 4 | Cabbage palm/Oak hammock | 252 |
| 5 | Slash pine | 161 |
| 6 | Oak/Broadleaf hammock | 229 |
| 7 | Hardwood swamp | 105 |
| 8 | Graminoid marsh | 431 |
| 9 | Spartina marsh | 520 |
| 10 | Cattail marsh | 404 |
| 11 | Salt marsh | 419 |
| 12 | Mud flats | 503 |
| 13 | Water | 927 |
| | Total | 5211 |

**Botswana**: The Botswana images, built over Okavango Delta, Botswana on May 31, 2001 by the NASA EO-1 satellite, are of size 1476 × 256. EO-1 acquires data in 242 bands from the 0.4 µm to 2.5 µm portion of the spectrum in 10 nm windows. Noise bands are discarded, and the remaining 145 bands are chosen as candidate features. The false-color image of the Botswana dataset and the corresponding ground truth image are illustrated in Figure 3. The specific category information is listed in Table 2.

**Table 2.** Detailed category information in Botswana dataset.

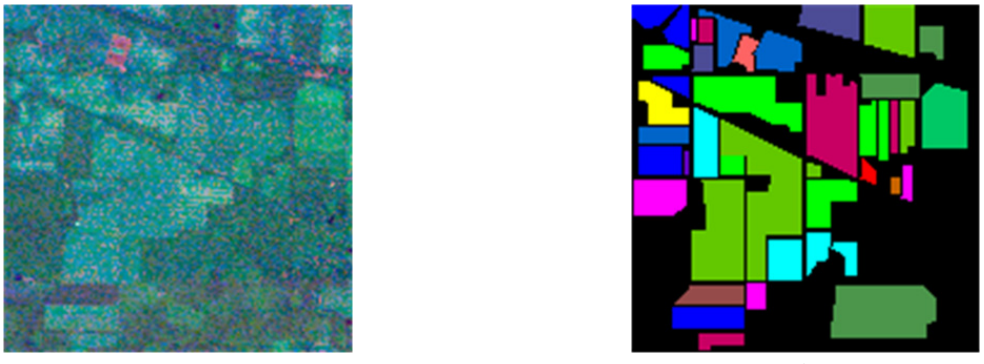
| Class Number | Class Name | Number of Samples |
|---|---|---|
| 1 | Water | 270 |
| 2 | Hippo grass | 101 |
| 3 | Floodplain grasses 1 | 251 |
| 4 | Floodplain grasses 2 | 215 |
| 5 | Reeds | 269 |
| 6 | Riparian | 269 |
| 7 | Firescar | 259 |
| 8 | Island interior | 203 |
| 9 | Acacia woodlands | 314 |
| 10 | Acacia shrublands | 248 |
| 11 | Acacia grasslands | 305 |
| 12 | Short mopane | 181 |
| 13 | Mixed mopane | 268 |
| 14 | Exposed soils | 95 |
| | Total | 3248 |



(**a**) (**b**)

**Figure 3.** (**a**) Botswana HSI. (**b**) Ground truth.

**Indian Pines:** The Indian Pines dataset was gathered by AVIRIS sensor in northwestern Indiana. The acquired image consists of 145 × 145 pixels and 224 original spectral bands with a spectrum range from 0.4 µm to 2.5 µm. The total number of bands is reduced to 200 by removing bands containing the area of water absorption. Figure 4 shows the false-color image of the Indian Pines dataset and the corresponding ground truth image. The specific category information is given in Table 3.



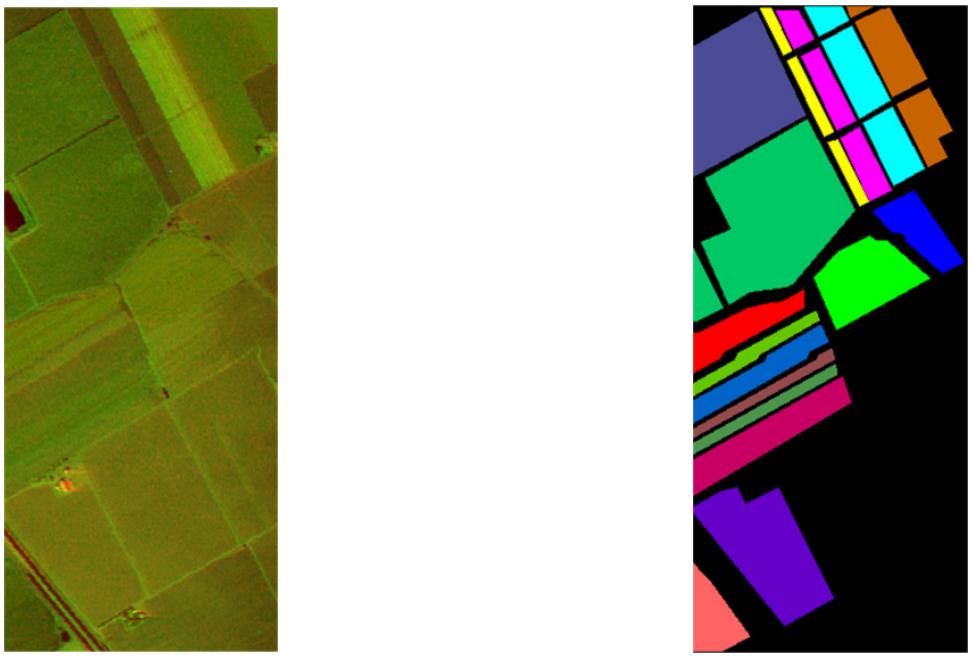
(**a**) (**b**)

**Figure 4.** (**a**) Indian Pines HSI. (**b**) Ground truth.

**Table 3.** Detailed category information in Indian Pines dataset.

| Class Number | Class Name | Number of Samples |
|---|---|---|
| 1 | Alfalfa | 46 |
| 2 | Corn-notill | 1428 |
| 3 | Corn-mintill | 830 |
| 4 | Corn | 237 |
| 5 | Grass-pasture | 483 |
| 6 | Grass-trees | 730 |
| 7 | Grass-pasture-mowed | 28 |
| 8 | Hay-windowed | 478 |
| 9 | Oats | 20 |
| 10 | Soybean-notill | 972 |
| 11 | Soybean-mintill | 2455 |
| 12 | Soybean-clean | 593 |
| 13 | Wheats | 205 |
| 14 | Woods | 1265 |
| 15 | Building-Grass-Trees-Drivers | 386 |
| 16 | Stone-Steel-Towers | 93 |
| | Total | 10,249 |

**Salinas**: The Salinas dataset shown in Figure 5 was obtained by an AVIRIS sensor on Salinas Valley. The HSI is formed by 512 × 217 pixels and 224 bands in the spectrum range 0.4–2.5 µm. Further, 204 bands in this scene are retained by discarding 20 water absorption bands. The detailed category information of the Salinas dataset is given in Table 4.



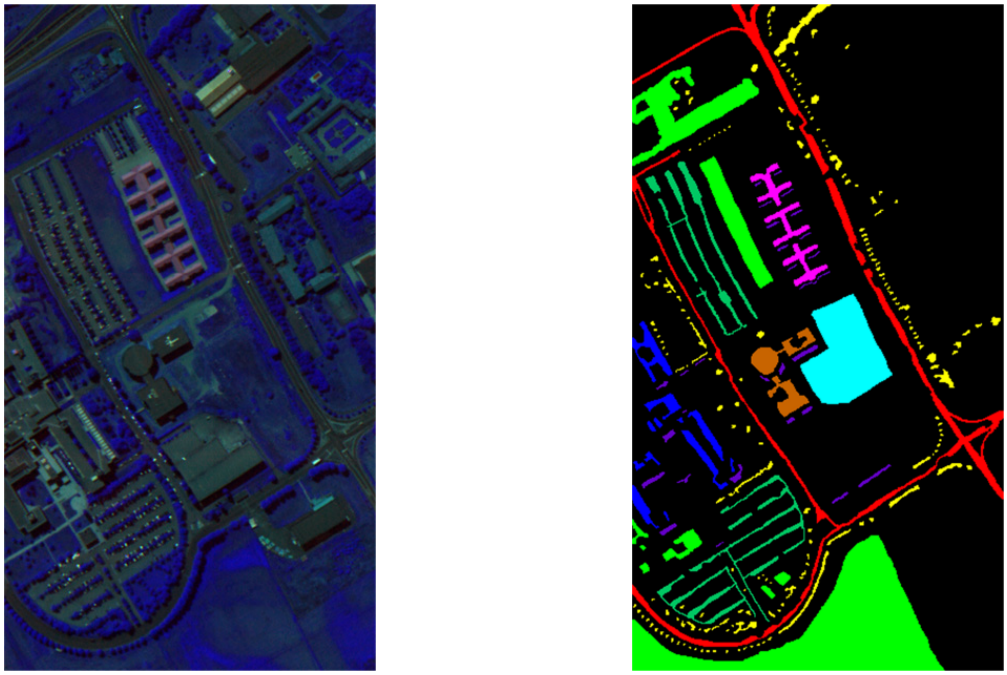
(**a**) (**b**)

**Figure 5.** (**a**) Salinas HSI. (**b**) Ground truth.

**Table 4.** Detailed category information in Salinas dataset.

| Class Number | Class Name | Number of Samples |
|---|---|---|
| 1 | Brocoli_green_weeds_1 | 2009 |
| 2 | Brocoli_green_weeds_2 | 3726 |
| 3 | Fallow | 1976 |
| 4 | Fallow_rough_plow | 1394 |
| 5 | Fallow_smooth | 2678 |
| 6 | Stubble | 3959 |
| 7 | Celery | 3579 |
| 8 | Grapes_untrained | 11,271 |
| 9 | Soil_vinyard_develop | 6203 |
| 10 | Corn_senesced_green_weeds | 3278 |
| 11 | Lettuce_romaine_4wk | 1068 |
| 12 | Lettuce_romaine_5wk | 1927 |
| 13 | Lettuce_romaine_6wk | 916 |
| 14 | Lettuce_romaine_7wk | 1070 |
| 15 | Vinyard_untrained | 7268 |
| 16 | Vinyard_vertical_trellis | 1807 |
| | Total | 54,129 |

**Pavia University**: The last dataset was collected from Pavia University in 2002. Pavia University is a 610 × 340 pixels image, and the number of spectral bands is 103. This image was taken by ROSIS on the wavelength range of 0.43 μm to 0.86 μm. The false-color image of Pavia University dataset and the corresponding ground truth image are illustrated in Figure 6. The concrete category information is listed in Table 5.



(**a**) (**b**)

**Figure 6.** (**a**) Pavia University HSI. (**b**) Ground truth.

**Table 5.** Detailed category information in Pavia University dataset.

| Class Number | Class Name | Number of Samples |
|---|---|---|
| 1 | Asphalt | 6631 |
| 2 | Meadows | 18,649 |
| 3 | Gravel | 2099 |
| 4 | Trees | 3064 |
| 5 | Painted metal sheets | 1345 |
| 6 | Bare soil | 5029 |
| 7 | Bitumen | 1330 |
| 8 | Self-Blocking Bricks | 3682 |
| 9 | Shadows | 947 |
| | Total | 42,776 |

### *4.2. Parameter Settings*

Appropriate parameter settings can improve the optimization ability of the algorithm. In the following experiments, the proposed band selection technique is compared with GA [28], PSO [32], CS [33], FA [34] and HRO [38]. The corresponding parameter settings of each algorithm are listed in Table 6. To make a fair comparison, all of the algorithms are adopted by binary coding and the corresponding band subset is used as input to SVM for classification. Each algorithm has an initial population size of 20 and a maximum number of iterations of 30. For all HSI datasets, 20% of samples are randomly selected as training data, and the remaining 80% are chosen as testing data. All algorithms are independently run 10 times for a case, and the average results are recorded.

**Table 6.** Parameters setting of each algorithm.

| Algorithm | Parameters | Value |
|---|---|---|
| GA | Crossover rate $C_R$ | 0.8 |
| | Mutation rate $C_M$ | 0.01 |
| PSO | Acceleration coefficients $c_1$, $c_2$ | 2 |
| | Minimum inertia weight $\omega_{min}$ | 0.2 |
| | Maximum inertia weight $\omega_{max}$ | 0.9 |
| CS | Detection probability $p_a$ | 0.25 |
| | Levy flight parameter $\beta$ | 1.5 |
| FA | Absorption coefficient $\gamma$ | 1 |
| | Initial attraction $\beta_0$ | 1 |
| | Randomization parameter $\alpha$ | 0.5 |
| HRO, MHRO | Maximum selfing time $t_{max}$ | 10 |

### *4.3. Experiments for Different Optimization Algorithms*

In this section, five benchmark datasets were used to test the performance of band selection based on MHRO. The overall classification accuracy (OA), kappa coefficient (The concept of kappa coefficient is described in Appendix A.2), number of selected bands and fitness function of each algorithm are utilized as evaluation indicators, as shown in Tables 7 and 8. It could be seen from Table 7 that the optimization ability of the proposed method is apparently superior to GA, PSO, CS, FA and HRO algorithms in terms of OA and kappa coefficients. It indicates DE operators employed by the individuals in the maintainer line have more probability to enhance the exploitation ability of MHRO in local search. For the Indian Pines dataset, OA of MHRO is 6.6% higher than GA, 4.89% higher than CS, and 6.17% higher than FA. In addition, the kappa coefficient of MHRO in Botswana, Salinas and Pavia University datasets are all over 0.94. This shows that the classification results are basically consistent with the real category labels.

**Table 7.** OA and Kappa coefficient of six algorithms.

| Dataset | Metrics | GA | PSO | CS | FA | HRO | MHRO |
|---|---|---|---|---|---|---|---|
| KSC | OA (%) | 92.84 ± 0.14 | 93.24 ± 0.1 | 92.92 ± 0.07 | 92.83 ± 0.06 | 93.34 ± 0.19 | **93.60 ± 0.10** |
| | Kappa | 0.9201 ± 0.0016 | 0.9246 ± 0.0012 | 0.9211 ± 0.0007 | 0.9200 ± 0.0007 | 0.9257 ± 0.0021 | **0.9287 ± 0.0011** |
| Botswana | OA (%) | 94.69 ± 0.14 | 95.49 ± 0.23 | 94.79 ± 0.07 | 94.74 ± 0.09 | 95.63 ± 0.22 | **95.96 ± 0.10** |
| | Kappa | 0.9425 ± 0.0015 | 0.9511 ± 0.0024 | 0.9436 ± 0.0007 | 0.9431 ± 0.0010 | 0.9526 ± 0.0024 | **0.9562 ± 0.0010** |
| Indian Pines | OA (%) | 82.65 ± 0.85 | 88.75 ± 0.50 | 84.36 ± 0.62 | 83.08 ± 0.42 | 88.43 ± 0.27 | **89.25 ± 0.32** |
| | Kappa | 0.8018 ± 0.0098 | 0.8718 ± 0.0057 | 0.8215 ± 0.0072 | 0.8068 ± 0.0049 | 0.8682 ± 0.0031 | **0.8776 ± 0.0036** |
| Salinas | OA (%) | 94.39 ± 0.03 | 94.64 ± 0.07 | 94.49 ± 0.05 | 94.49 ± 0.03 | 94.59 ± 0.06 | **94.73 ± 0.04** |
| | Kappa | 0.9375 ± 0.0004 | 0.9403 ± 0.0008 | 0.9386 ± 0.0006 | 0.9386 ± 0.0003 | 0.9397 ± 0.0007 | **0.9413 ± 0.0005** |
| Pavia University | OA (%) | 95.19 ± 0.09 | 95.46 ± 0.04 | 95.30 ± 0.05 | 95.34 ± 0.07 | 95.41 ± 0.04 | **95.53 ± 0.06** |
| | Kappa | 0.9360 ± 0.0012 | 0.9396 ± 0.0005 | 0.9376 ± 0.0007 | 0.9380 ± 0.0010 | 0.9390 ± 0.0006 | **0.9406 ± 0.0008** |

**Table 8.** The number of selected bands and fitness value of six algorithms.

| Dataset | Metrics | GA | PSO | CS | FA | HRO | MHRO |
|---|---|---|---|---|---|---|---|
| KSC | Num | 81.3 | 65.5 | 82 | 89.7 | 41 | **37.5** |
| | Fitness | 0.0756 ± 0.0015 | 0.0707 ± 0.0008 | 0.0747 ± 0.0006 | 0.0761 ± 0.0007 | 0.0683 ± 0.0020 | **0.0655 ± 0.0009** |
| Botswana | Num | 70.5 | 56.2 | 72.4 | 77.6 | 37.9 | **35.3** |
| | Fitness | 0.0574 ± 0.0016 | 0.0486 ± 0.0025 | 0.0566 ± 0.0008 | 0.0574 ± 0.0006 | 0.0459 ± 0.0020 | **0.0425 ± 0.0010** |
| Indian Pines | Num | 96.5 | 69.4 | 89.1 | 106.4 | **41.1** | **44.1** |
| | Fitness | 0.1766 ± 0.0085 | 0.1149 ± 0.0051 | 0.1593 ± 0.0063 | 0.1728 ± 0.0042 | 0.1166 ± 0.0027 | **0.1086 ± 0.0030** |
| Salinas | Num | 99.6 | 91.8 | 102.8 | 109.5 | **82.6** | **85.8** |
| | Fitness | 0.0604 ± 0.0003 | 0.0575 ± 0.0009 | 0.0596 ± 0.0005 | 0.0599 ± 0.0002 | 0.0576 ± 0.0004 | **0.0564 ± 0.0004** |
| Pavia University | Num | 51.1 | 46.2 | 52.9 | 56.6 | 42.7 | **41.9** |
| | Fitness | 0.0526 ± 0.0008 | 0.0495 ± 0.0004 | 0.0516 ± 0.0004 | 0.0517 ± 0.0003 | 0.0496 ± 0.0003 | **0.0484 ± 0.0004** |

According to Table 8, it is observed that the number of selected bands using HRO and MHRO is significantly less than those of GA, PSO, CS and FFA. The average number of selected bands with FA is 88 in the five datasets, nearly 1.8 times that of MHRO. About 79% of the high correlation and redundancy bands from the KSC dataset are removed by MHRO, with an average of only 37 bands with satisfactory classification accuracy remaining. Moreover, it is noticed that MHRO has slightly more band subsets than HRO for Indian Pines and Salinas datasets, which is caused by the fact that MHRO prioritizes the superior precision between high classification accuracy and a smaller number of bands at each iteration. With regard to the fitness, the proposed method has a better fitness value compared with the other five algorithms, and its corresponding standard deviation does not exceed 0.003 in any dataset, which verifies that MHRO has a slight fluctuation in HSI datasets. More importantly, the standard deviation of fitness value with MHRO is only 0.0004 for Salinas and Pavia University datasets, which is stable for independent operations. Figure 7 depicts the variation of average fitness with the number of iterations for all the algorithms used on five datasets.

As it is shown in Figure 7, the initial fitness value of MHRO is lower than HRO in all datasets, which proved that the OBL strategy used in the initial stage could enhance population diversity and provide more high-quality solutions. With the increase of iteration times, the iteration curves of GA, CS and FA gradually tend to be stable, while MHRO keeps a downward trend. This shows that DE operators can help to improve the exploration and exploitation ability, and also implies that MHRO has a powerful potential to find better solutions. As a result, MHRO has excellent optimization capability on HSI datasets and can obtain an optimal band subset with satisfactory classification accuracy.

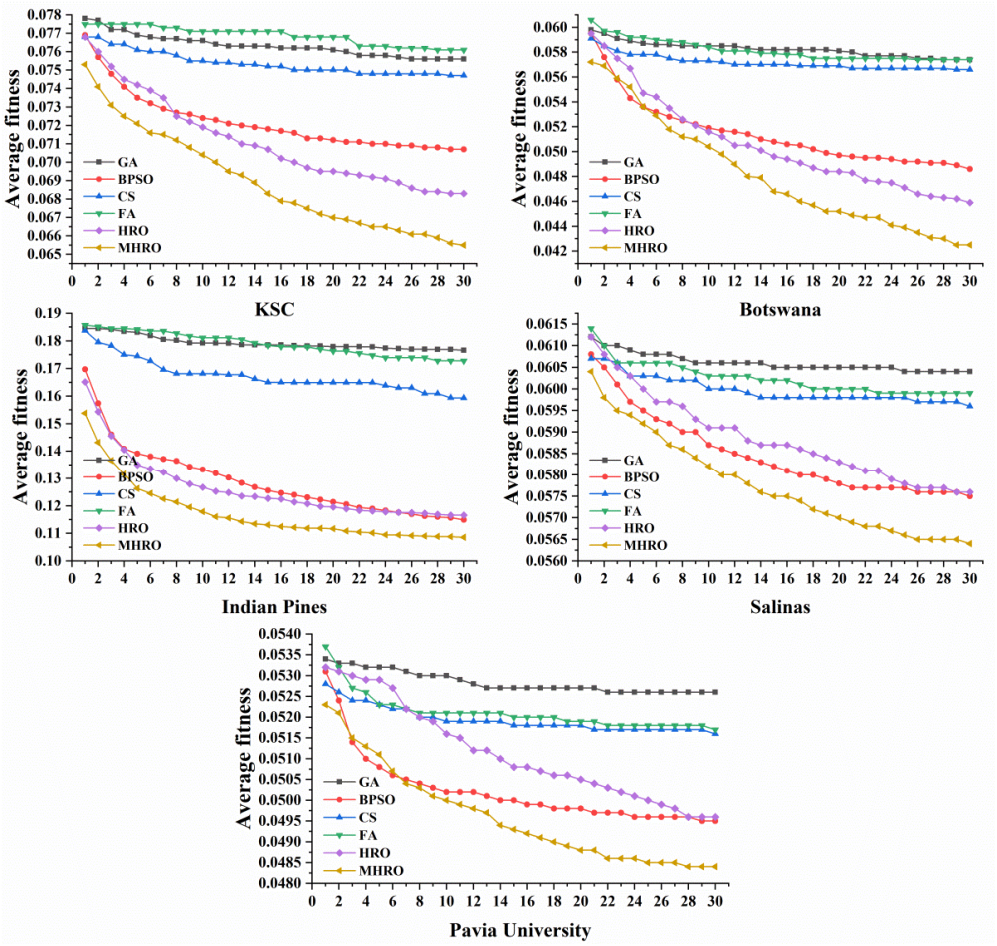



**Figure 7.** The changing process of average fitness on five datasets.

### *4.4. Experiments for Other Related Techniques*

To further verify the reliability of the proposed algorithm, the MHRO is compared with two common feature-selection methods, including Joint Mutual Information (JMI) and its improvement, named Joint Mutual Information Maximization (JMIM) [42] on five datasets. The experiments are implemented with bands varying from 10% to 30% of the total number of bands in each dataset. The results of classification accuracy for each class, OA and kappa coefficients are recorded in Tables 9–13.

**Table 9.** The classification results for KSC dataset via different number of bands.

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | JMI | JMIM | MHRO | JMI | JMIM | MHRO | JMI | JMIM | MHRO |
| 1 | 92.5750 | **93.4319** | 92.5350 | **94.1935** | 93.9145 | 93.9490 | **94.4805** | 93.3116 | 93.5024 |
| 2 | **92.8144** | 92.6554 | 90.0000 | **93.7500** | 93.2203 | 91.5344 | 93.7107 | **93.8889** | 91.9786 |
| 3 | 65.2985 | **91.2821** | 87.0466 | 77.8723 | 87.4419 | **89.9471** | 76.6520 | 87.7934 | **87.8307** |
| 4 | 58.9286 | 61.4504 | **73.0233** | 70.2564 | 63.6000 | **73.3032** | 66.5116 | 63.0522 | **74.0741** |
| 5 | 76.1364 | 78.0220 | **79.1304** | **79.6117** | 81.3187 | **82.7957** | 85.1852 | **81.4433** | 80.7692 |
| 6 | 76.5432 | 60.3659 | **78.9474** | 79.1667 | 68.1818 | 75.7225 | 75.2688 | 67.5497 | **80.2395** |
| 7 | 64.6552 | 71.6981 | **81.7073** | 65.2893 | 72.6415 | **88.8889** | 64.1667 | 75.2475 | **85.3659** |
| 8 | 84.6939 | 89.0141 | **93.6599** | 90.0990 | 88.5559 | **90.7609** | 89.0625 | 88.1081 | **91.1602** |
| 9 | 77.3469 | 93.9252 | **94.0909** | 83.5470 | **95.2038** | 94.3052 | 87.1681 | **96.5602** | 94.5205 |
| 10 | 89.1374 | **100.00** | 97.0297 | 95.8333 | **100.00** | 96.0784 | 95.2381 | **100.00** | 96.4286 |

**Table 9.** *Cont.*

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** |
| 11 | 89.6739 | 99.6997 | **99.7050** | 91.1932 | 99.3976 | **99.4030** | 91.1932 | 99.1018 | **99.7015** |
| 12 | 98.0609 | 97.6982 | **98.5000** | 96.6752 | **98.9848** | 98.7406 | 97.4026 | **98.4925** | 98.0000 |
| 13 | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** | **100.00** |
| OA (%) | 86.6635 | 91.3408 | **93.1638** | 89.9736 | 92.0844 | **93.3557** | 89.9496 | 92.1084 | **93.3797** |
| Kappa | 0.8514 | 0.9036 | **0.9238** | 0.8883 | 0.9119 | **0.9259** | 0.8881 | 0.9121 | **0.9262** |

**Table 10.** The classification results for Botswana dataset via different number of bands.

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** |
| 1 | **100.00** | 99.5392 | **100.00** | 99.5392 | 99.5392 | **100.00** | 99.5392 | 99.5392 | **100.00** |
| 2 | 98.7500 | **98.7654** | 97.4359 | **100.00** | **100.00** | 98.7342 | **100.00** | **100.00** | 96.2963 |
| 3 | 98.0296 | 95.6522 | **99.0050** | 98.0100 | 97.5610 | **99.5050** | 98.0488 | 97.5728 | **99.0148** |
| 4 | 92.7374 | 94.9721 | **95.0276** | 92.2222 | **97.0930** | 96.6292 | 91.7582 | 94.3503 | **96.6480** |
| 5 | 85.2941 | **89.9471** | 88.7324 | 83.9623 | **90.0000** | 89.5735 | 88.4422 | 88.0597 | **90.3846** |
| 6 | 75.7576 | 76.9953 | **83.1050** | 80.9278 | 79.8122 | **88.6792** | 80.2885 | 80.4762 | **87.9070** |
| 7 | **100.00** | 99.5098 | 99.5074 | **100.00** | **100.00** | **100.00** | 99.5050 | 99.5000 | **100.00** |
| 8 | 94.8718 | **97.5155** | 96.3636 | 96.2025 | **99.3548** | 94.7368 | 98.7261 | **98.7421** | 96.4286 |
| 9 | 81.2977 | 86.3469 | **94.7154** | 86.5672 | 87.0370 | **96.4844** | 87.5472 | 88.3895 | **96.0784** |
| 10 | 83.5616 | 83.1776 | **91.3265** | 86.3014 | 87.0192 | **92.0213** | 88.6792 | 88.2075 | **94.2708** |
| 11 | 95.3191 | 94.7826 | **96.1373** | **97.0085** | 94.1423 | 94.1423 | 95.7627 | 96.1207 | **96.6102** |
| 12 | 93.8776 | **94.5946** | **91.2162** | **95.2381** | 93.3333 | 94.2857 | 91.4474 | 93.8776 | **94.3262** |
| 13 | 88.8889 | 95.2607 | **97.0588** | 93.5484 | 93.0556 | **96.1905** | 92.9245 | 92.2018 | **97.1154** |
| 14 | 97.3333 | **100.00** | 97.5000 | 98.6301 | **100.00** | 98.7342 | 98.6301 | 98.6486 | **98.7342** |
| OA (%) | 90.9196 | 92.4586 | **94.5748** | 92.6510 | 93.2282 | **95.4598** | 92.9588 | 93.2282 | **95.8830** |
| Kappa | 0.9016 | 0.9183 | **0.9412** | 0.9204 | 0.9266 | **0.9508** | 0.9237 | 0.9266 | **0.9554** |

**Table 11.** The classification results for Indian Pines dataset via different number of bands.

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** |
| 1 | 55.5556 | 91.6667 | **100.00** | 76.1905 | **100.00** | **100.00** | 72.7273 | 92.8571 | **100.00** |
| 2 | 43.4932 | 80.5106 | **89.8605** | 58.5537 | 84.8908 | **89.2364** | 58.6438 | 85.1363 | **89.6140** |
| 3 | 26.0976 | 79.1209 | **86.7621** | 51.7355 | 84.2975 | **85.4430** | 57.3574 | 85.3135 | **85.8034** |
| 4 | 36.3636 | 76.3006 | **76.6169** | 56.0510 | **75.2747** | 73.7864 | 49.2754 | 75.7396 | **77.2021** |
| 5 | 68.4807 | 90.3226 | **94.9735** | 86.1461 | 89.5141 | **93.6869** | 89.2857 | 88.5496 | **92.1182** |
| 6 | 79.3939 | 87.2699 | **92.6350** | 83.4609 | 88.4013 | **94.6667** | 84.1079 | 88.1620 | **94.6932** |
| 7 | 0.0000 | 95.2381 | **95.4545** | 64.5161 | **90.9091** | 88.0000 | 73.0769 | 90.9091 | **100.00** |
| 8 | 79.7357 | 95.6853 | **97.2152** | 94.1919 | 96.6667 | **98.7113** | 89.3671 | 97.1503 | **98.7147** |
| 9 | 0.0000 | 62.5000 | **90.0000** | 0.0000 | 66.6667 | **87.5000** | 20.0000 | 71.4286 | **86.6667** |
| 10 | 57.6493 | **82.2102** | 78.0652 | 65.2226 | **85.1316** | 79.3143 | 74.5739 | **84.9490** | 80.9187 |
| 11 | 56.7493 | 83.2692 | **86.5589** | 73.9841 | 84.9903 | **88.6812** | 75.5808 | 86.6535 | **89.0834** |
| 12 | 34.8485 | 79.4769 | **86.0324** | 52.9968 | 84.2536 | **87.8351** | 59.7531 | 84.4622 | **87.5510** |
| 13 | 78.5714 | **96.2025** | 94.8276 | 96.7105 | **98.0892** | 96.5116 | **99.3506** | 98.0892 | 98.8095 |
| 14 | 89.2120 | 92.4027 | **92.5785** | 93.6770 | 93.6416 | **93.9189** | 93.2740 | **94.5259** | 93.1232 |
| 15 | 60.8696 | **77.9661** | 73.4615 | 71.1297 | 78.5088 | **79.9228** | 69.7479 | 76.3780 | **78.4000** |
| 16 | 98.4615 | 98.5507 | **98.6111** | **98.4848** | 97.1014 | 98.3607 | 98.0392 | **98.5075** | 98.3607 |
| OA (%) | 61.0366 | 84.7927 | **87.8780** | 73.1463 | 87.0244 | **88.9024** | 74.5366 | 87.5244 | **89.2317** |
| Kappa | 0.5475 | 0.8260 | **0.8618** | 0.6926 | 0.8517 | **0.8737** | 0.7084 | 0.8575 | **0.8774** |

**Table 12.** The classification results for Salinas dataset via different number of bands.

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** |
| 1 | 99.8101 | 99.8101 | **100.00** | 99.7498 | 99.9373 | **100.00** | 99.9374 | **100.00** | **100.00** |
| 2 | 99.0017 | 99.0017 | **99.8321** | 99.5983 | 99.5654 | **99.7985** | 99.6317 | **99.8993** | 99.7650 |
| 3 | 87.9785 | 87.9785 | **96.1353** | 92.8571 | 93.1847 | **99.3769** | 95.1703 | 97.5686 | **99.5633** |
| 4 | 98.4889 | 98.4889 | **99.3789** | 98.6655 | 98.7533 | **99.2902** | 99.1063 | 98.9295 | **99.3783** |
| 5 | 88.9628 | 88.9628 | **99.5342** | 95.4922 | 95.3606 | **99.5818** | 98.4870 | 98.9686 | **99.6281** |
| 6 | 99.6211 | 99.6211 | **99.9373** | 99.9050 | 99.8735 | **99.9373** | 99.7472 | 99.7788 | **99.9373** |
| 7 | 99.3338 | 99.3338 | **99.9644** | 99.8252 | 99.7556 | **99.9288** | **99.9650** | 99.8602 | 99.9644 |
| 8 | 72.4820 | 72.4820 | **80.7262** | 77.6299 | 78.1479 | **83.2298** | 79.0656 | 81.4202 | **84.3931** |
| 9 | 93.7717 | 93.7717 | **99.4389** | 96.2108 | 96.3827 | **99.5787** | 97.1542 | 98.1235 | **99.5986** |
| 10 | 92.6040 | 92.6040 | **97.7493** | 95.2437 | 95.1362 | **98.3116** | 95.6164 | 95.9938 | **98.3109** |
| 11 | 79.8193 | 79.8193 | **98.9260** | 86.9677 | 87.5635 | **99.6433** | 88.6499 | 87.1111 | **99.4055** |
| 12 | 95.8543 | 95.8543 | **98.7630** | 96.9046 | 96.9562 | **99.0209** | 96.8434 | 97.4603 | **99.2801** |
| 13 | 95.7333 | 95.7333 | **99.3151** | 96.6622 | 98.5034 | **99.8626** | 97.8408 | 98.3762 | **99.8623** |
| 14 | 96.7153 | 96.7153 | **98.5849** | 97.8417 | 98.0000 | **98.8249** | 97.6387 | 98.2332 | **98.5899** |
| 15 | 66.6223 | 66.6223 | **84.2755** | 75.7640 | 75.5780 | **84.2284** | 77.6699 | 80.8717 | **84.7408** |
| 16 | 98.2244 | 98.2244 | **99.5830** | 98.8842 | 98.8137 | **99.2388** | 99.1028 | **99.1701** | **99.3070** |
| OA (%) | 86.9573 | 86.9573 | **93.2662** | 90.3589 | 90.5159 | **94.0467** | 91.3195 | 92.4834 | **94.3816** |
| Kappa | 0.8543 | 0.8543 | **0.9249** | 0.8924 | 0.8942 | **0.9336** | 0.9032 | 0.9162 | **0.9374** |

**Table 13.** The classification results for Pavia University dataset via different number of bands.

| Class Number | 10% | | | 20% | | | 30% | | |
|---|---|---|---|---|---|---|---|---|---|
| | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** | **JMI** | **JMIM** | **MHRO** |
| 1 | 86.7749 | 91.5299 | **94.2054** | 89.3364 | 93.0755 | **95.4459** | 91.6985 | 94.4086 | **95.5572** |
| 2 | 81.0733 | 90.0861 | **95.0046** | 83.7405 | 93.8672 | **96.2513** | 87.2954 | 95.9982 | **96.6656** |
| 3 | 69.8551 | 78.1069 | **87.9648** | 71.1213 | 82.0709 | **89.7384** | 76.1051 | 83.6127 | **89.7283** |
| 4 | 92.3801 | **96.3346** | 95.7699 | 92.0854 | 96.5762 | **97.3862** | 93.7824 | 97.6637 | **97.7580** |
| 5 | 99.7180 | **99.9065** | 99.4550 | 99.4403 | **99.9064** | 99.3642 | 99.5323 | **100.00** | 99.6357 |
| 6 | 89.9167 | 92.0507 | **93.0775** | 89.3754 | 91.2009 | **95.0250** | 88.6102 | 94.7454 | **95.7552** |
| 7 | 73.9530 | 79.6693 | **86.4476** | 77.1178 | 83.2370 | **89.2323** | 84.4828 | 85.8252 | **88.9437** |
| 8 | 79.6457 | 80.0737 | **84.4270** | 80.5669 | 82.9960 | **84.8779** | 82.3899 | **86.5560** | 86.2015 |
| 9 | 99.8681 | **99.8682** | **99.8682** | 99.8681 | 99.8681 | **99.8682** | **99.8681** | 99.8681 | **99.8681** |
| OA (%) | 83.1887 | 89.6701 | **93.4426** | 85.2138 | 92.1452 | **94.6992** | 88.1155 | 94.2082 | **95.1316** |
| Kappa | 0.7677 | 0.8607 | **0.9126** | 0.7976 | 0.8951 | **0.9295** | 0.8390 | 0.9229 | **0.9353** |

Table 9 shows OA and kappa values generated for the KSC dataset under different band subsets. It is clear that the proposed MHRO has the highest OA compared with other filter feature techniques. In particular, OA of MHRO is 6.5% higher than JMI and 1.8% higher than JMIM via 10% entire number of bands. In addition, only OA and Kappa coefficient of JMI are lower than 90% with different band subsets, and MHRO both exceed 92%. In brief, it is verified that the proposed band selection approach has good practicability for KSC dataset.

Table 10 reports OA and kappa values generated for the Botswana dataset under different band subsets. It is induced that OA corresponding to MHRO is satisfactory, which is more than 95% via exceeding 20% the number of bands. More importantly, the classification accuracy by MHRO is superior to JMI and JMIM in 12 categories with 30% of the total number of bands. Moreover, kappa coefficients of MHRO are approximately 0.02–0.04, 0.02–0.03 and 0.028–0.032 higher than other approaches in the total number of bands 10%, 20% and 30%, respectively. In sum, it is an effective band selection technique for the Botswana dataset.

OA and kappa values generated for the Indian Pines dataset under different band subsets are given in Table 11. It is observed that OA of the proposed MHRO significantly outperformed JMI, and the difference is over 14% even 26%. Further, the accuracy achieved by using MHRO reaches 100% on class number 1 and just 55.55% by using JMI via 10% of all number of bands. It is worth noting that the category named Oats is difficult to classify correctly by using JMI and JMIM, and the number of samples correctly distinguished by MHRO exceeded 86%. Therefore, it is a more promising technique than filtering feature-selection methods on the Indian Pines dataset.

Table 12 reports OA and kappa values generated for the Salinas dataset under different band subsets. It is obviously revealed that the proposed band selection approach can still obtain the best classification accuracy when the selected number of bands does not exceed 20%, and OA is lower than 87% by using JMI and JMIM techniques via 10% total number of bands. More than 98% of the samples could be correctly distinguished in the 13 categories by MHRO and the corresponding kappa coefficient is more than 0.92, which infers that the precisions are basically consistent with the real category labels. In short, the proposed MHRO has a strong optimization ability on the Salinas dataset.

According to Table 13, it has been concluded that the classification accuracy rate obtained by the proposed MHRO performs better than JMI and JMIM in most categories. For 10% total number of bands, OA is lower than 84% by using JMI and lower than 90% by JMIM. Kappa value by using MHRO has reached 0.91, 0.92 and 0.93 in the different number of bands, respectively, which are 0.9–15 higher than JMI. In conclusion, the proposed MHRO is a robust and feasible feature-selection approach for the Pavia University dataset.

## 5. Discussion

As shown in Tables 7 and 8, the experimental results of MHRO on all datasets performed better than other swarm intelligence algorithms. Except for Indian Pines and Salinas, HRO has slightly smaller band subsets than MHRO. It can be seen from Figure 7 that the fitness function values of GA, CS and FA are only slight fluctuations throughout the iteration, indicating that they easily fall into local optimum at the early stage of iteration. In contrast, the fitness value of MHRO is always the lowest and keeps a decreasing trend, which implies that MHRO has strong optimization ability and is superior to algorithms GA [28], PSO [32], CS [33], FA [34] and HRO [38].

According to Tables 9–13, MHRO achieves higher accuracy than filter techniques JMI and JMIM [42] under different band subsets. This can be explained by the fact that the filter methods use mutual information to select feature subsets and they are independent of the classifier, while the wrapper approach MHRO proposed in the paper calculates the fitness function based on the accuracy of the classifier and the number of selected bands.

## 6. Conclusions

Band selection is a crucial phase to remove high-correlation bands and improve the classification accuracy for HSI. In the paper, a band selection approach based on MHRO is proposed and the basic idea is to obtain the fittest band combination. Experimental results are compared with commonly used feature-selection approaches optimized by GA, PSO, CS, FA and standard HRO on five datasets. In general, it is concluded that the proposed MHRO has excellent optimization capability and is able to achieve the highest classification accuracy with fewer bands. Moreover, OA and kappa coefficient are obviously higher than other related feature-selection techniques JMI and JMIM, which proved that the precisions are basically consistent with the real category labels. As a result, the proposed band selection technique has good robustness and practicability for HSI datasets. Future work will investigate other swarm intelligence algorithms and combine multiple optimization strategies and spatial information to further improve the performance of band selection. In addition, it is worthwhile to formulate different assessment criteria to solve the multi-objective optimization problem for feature selection on large-scale datasets.

**Author Contributions:** Conceptualization, Z.Y. and M.W.; methodology, W.C. and S.L.; software, W.C. and S.L.; validation, Z.Y. and K.L.; formal analysis, W.C.; investigation, W.C. and S.L.; resources, Z.Y.; data curation, W.C., K.L. and M.W.; writing—original draft preparation, Z.Y. and W.C.; writing—review and editing, Z.Y. and W.Z.; visualization, W.C.; supervision, Z.Y. and K.L.; project administration, Z.Y.; funding acquisition, Z.Y. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research was funded by the National Natural Science Foundation of China under Grant No. 61502155, 61772180, funded by Fujian Provincial Key Laboratory of Data Intensive Computing and Key Laboratory of Intelligent Computing and Information Processing, Fujian No. BD201801.

**Institutional Review Board Statement:** Not applicable.

**Informed Consent Statement:** Not applicable.

**Data Availability Statement:** The public data in section "Datasets description" are available at https://www.ehu.eus/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes (accessed on 25 March 2022).

**Conflicts of Interest:** The authors declare no conflict of interest.

## Appendix A

Overall classification accuracy (OA) and Kappa coefficient are commonly used in HSI classification, and they will be described in detail as follows.

### *Appendix A.1. Overall Classification Accuracy*

OA represents the ratio between the number of correctly classified samples by the classifier or classification algorithm and the total number of samples, and the mathematical expression is shown as Equation (A1).

$$\mathrm{OA} = \frac{\sum_{i=1}^{N_c} C_{ii}}{\sum_{j=1}^{N_c} \sum_{i=1}^{N_c} C_{ij}} \tag{A1}$$

where $N_c$ is the number of classes, $C_{ii}$ represents the number of samples correctly classified to class $i$, and $C_{ij}$ denotes the number of samples of $i$-th class assigned to $j$-th category.

### *Appendix A.2. Kappa Coefficient*

Kappa coefficient is a statistical indicator to measure the agreement between the final classification results and the ground-truth map, and its value is in the range of $[-1, 1]$. If the value of kappa coefficient is closer to 1, it indicates that the classification result is better. Kappa coefficient is given as Equation (A2).

$$\mathrm{Kappa} = \frac{N_s \sum_{i=1}^{N_c} C_{ii} - \sum_{i=1}^{N_c} C_{i+} C_{+i}}{{N_s}^2 - \sum_{i=1}^{N_c} C_{i+} C_{+i}} \tag{A2}$$

where $N_s$ is the number of samples, $C_{i+}$ represents the number of samples in class $i$ and $C_{+i}$ denotes the total number of samples of non-category $i$ predicted to be category $i$.

## References

1. Weber, C.; Aguejdad, R.; Briottet, X.; Avala, J.; Fabre, S.; Demuynck, J.; Zenou, E.; Deville, Y.; Karoui, M.S.; Benhalouche, F.Z.; et al. Hyperspectral Imagery for Environmental Urban Planning. In Proceedings of the IGARSS 2018-2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 1628–1631.
2. Sethy, P.K.; Pandey, C.; Sahu, Y.K.; Behera, S.K. Hyperspectral imagery applications for precision agriculture-a systemic survey. *Multimed. Tools Appl.* **2022**, *81*, 3005–3038. [CrossRef]
3. Yi, L.; Chen, J.M.; Zhang, G.; Xu, X.; Guo, W. Seamless Mosaicking of UAV-Based Push-Broom Hyperspectral Images for Environment Monitoring. *Remote Sens.* **2021**, *13*, 4720. [CrossRef]
4. Wang, Z.; Tian, S. Ground object information extraction from hyperspectral remote sensing images using deep learning algorithm. *Microprocess. Microsyst.* **2021**, *87*, 104394. [CrossRef]

5. Xie, W.; Li, Y.; Lei, J.; Yang, J.; Li, J.; Jia, X.; Li, Z. Unsupervised spectral mapping and feature selection for hyperspectral anomaly detection. *Neural Netw.* **2020**, *132*, 144–154. [CrossRef]
6. Sawant, S.; Manoharan, P. Hyperspectral band selection based on metaheuristic optimization approach. *Infrared Phys. Technol.* **2020**, *107*, 103295. [CrossRef]
7. Li, R.; Zhang, H.; Chen, Z.; Yu, N.; Kong, W.; Li, T.; Wang, E.; Wu, X.; Liu, Y. Denoising method of ground-penetrating radar signal based on independent component analysis with multifractal spectrum. *Measurement* **2022**, *192*, 110886. [CrossRef]
8. Xie, S. Feature extraction of auto insurance size of loss data using functional principal component analysis. *Expert Syst. Appl.* **2022**, *198*, 116780. [CrossRef]
9. Liu, Q.; He, H.; Liu, Y.; Qu, X. Local linear embedding algorithm of mutual neighborhood based on multi-information fusion metric. *Measurement* **2021**, *186*, 110239. [CrossRef]
10. Ding, X.; Li, H.; Yang, J.; Dale, P.; Chen, X.; Jiang, C.; Zhang, S. An improved ant colony algorithm for optimized band selection of hyperspectral remotely sensed imagery. *IEEE Access* **2020**, *8*, 25789–25799. [CrossRef]
11. Zhang, A.; Ma, P.; Liu, S.; Sun, G.; Huang, H.; Zabalza, J.; Wang, Z.; Lin, C. Hyperspectral band selection using crossover-based gravitational search algorithm. *IET Image Processing* **2019**, *13*, 280–286. [CrossRef]
12. Ambusaidi, M.A.; He, X.; Nanda, P.; Tan, Z. Building an intrusion detection system using a filter-based feature selection algorithm. *IEEE Trans. Comput.* **2016**, *65*, 2986–2998. [CrossRef]
13. Wah, Y.B.; Ibrahim, N.; Hamid, H.A.; Abdul-Rahman, S.; Fong, S. Feature Selection Methods: Case of Filter and Wrapper Approaches for Maximising Classification Accuracy. *Pertanika J. Sci. Technol.* **2018**, *26*, 329–340.
14. Ghosh, M.; Guha, R.; Sarkar, R.; Abraham, A. A wrapper-filter feature selection technique based on ant colony optimization. *Neural Comput. Appl.* **2020**, *32*, 7839–7857. [CrossRef]
15. Wang, J.; Tang, C.; Li, Z.; Liu, X.; Zhang, W.; Zhu, E.; Wang, L. Hyperspectral band selection via region-aware latent features fusion based clustering. *Inf. Fusion* **2022**, *79*, 162–173. [CrossRef]
16. Shi, J.; Zhang, X.; Liu, X.; Lei, Y.; Jeon, G. Multicriteria semi-supervised hyperspectral band selection based on evolutionary multitask optimization. *Knowl. -Based Syst.* **2022**, *240*, 107934. [CrossRef]
17. Bhadra, T.; Bandyopadhyay, S. Supervised feature selection using integration of densest subgraph finding with floating forward–backward search. *Inf. Sci.* **2021**, *566*, 1–18. [CrossRef]
18. Li, A.D.; Xue, B.; Zhang, M. Improved binary particle swarm optimization for feature selection with new initialization and search space reduction strategies. *Appl. Soft Comput.* **2021**, *106*, 107302. [CrossRef]
19. Ghosh, A.; Datta, A.; Ghosh, S. Self-adaptive differential evolution for feature selection in hyperspectral image data. *Appl. Soft Comput.* **2013**, *13*, 1969–1977. [CrossRef]
20. Turky, A.; Sabar, N.R.; Dunstall, S.; Song, A. Hyper-heuristic local search for combinatorial optimisation problems. *Knowl. -Based Syst.* **2020**, *205*, 106264. [CrossRef]
21. Faris, H.; Ala'M, A.Z.; Heidari, A.A.; Aljarah, I.; Mafarja, M.; Hassonah, M.A.; Fujita, H. An intelligent system for spam detection and identification of the most relevant features based on evolutionary random weight networks. *Inf. Fusion* **2019**, *48*, 67–83. [CrossRef]
22. Hancer, E. Differential evolution for feature selection: A fuzzy wrapper–filter approach. *Soft Comput.* **2019**, *23*, 5233–5248. [CrossRef]
23. Amoozegar, M.; Minaei-Bidgoli, B. Optimizing multi-objective PSO based feature selection method using a feature elitism mechanism. *Expert Syst. Appl.* **2018**, *113*, 499–514. [CrossRef]
24. Al-Tashi, Q.; Md, R.H.; Abdulkadir, S.J.; Mirjalili, S.; Alhussian, H. A review of grey wolf optimizer-based feature selection methods for classification. *Evol. Mach. Learn. Tech.* **2020**, 273–286. [CrossRef]
25. Aziz, M.A.E.; Hassanien, A.E. Modified cuckoo search algorithm with rough sets for feature selection. *Neural Comput. Appl.* **2018**, *29*, 925–934. [CrossRef]
26. Hancer, E.; Xue, B.; Zhang, M.; Karaboga, D.; Akay, B. Pareto front feature selection based on artificial bee colony optimization. *Inf. Sci.* **2018**, *422*, 462–479. [CrossRef]
27. Mafarja, M.; Mirjalili, S. Whale optimization approaches for wrapper feature selection. *Appl. Soft Comput.* **2018**, *62*, 441–453. [CrossRef]
28. Nagasubramanian, K.; Jones, S.; Sarkar, S.; Singh, A.K.; Singh, A.; Ganapathysubramanian, B. Hyperspectral band selection using genetic algorithm and support vector machines for early identification of charcoal rot disease in soybean stems. *Plant Methods* **2018**, *14*, 86. [CrossRef]
29. Xie, F.; Li, F.; Lei, C.; Yang, J.; Zhang, Y. Unsupervised band selection based on artificial bee colony algorithm for hyperspectral image classification. *Appl. Soft Comput.* **2019**, *75*, 428–440. [CrossRef]
30. Wang, M.; Wu, C.; Wang, L.; Xiang, D.; Huang, X. A feature selection approach for hyperspectral image based on modified ant lion optimizer. *Knowl. -Based Syst.* **2019**, *168*, 39–48. [CrossRef]
31. Wang, M.; Liu, W.; Chen, M.; Huang, X.; Han, W. A band selection approach based on a modified gray wolf optimizer and weight updating of bands for hyperspectral image. *Appl. Soft Comput.* **2021**, *112*, 107805. [CrossRef]
32. Kavitha, K.; Jenifa, W. Feature Selection Method for Classifying Hyper Spectral Image Based on Particle Swarm Optimization. In Proceedings of the 2018 International Conference on Communication and Signal Processing (ICCSP), Chennai, India, 3–5 April 2018; pp. 119–123.

33. Medjahed, S.A.; Saadi, T.A.; Benyettou, A.; Ouali, M. Binary cuckoo search algorithm for band selection in hyperspectral image classification. *IAENG Int. J. Comput. Sci.* **2015**, *42*, 183–191.
34. Su, H.; Yong, B.; Du, Q. Hyperspectral band selection using improved firefly algorithm. *IEEE Geosci. Remote Sens. Lett.* **2015**, *13*, 68–72. [CrossRef]
35. Ye, Z.; Ma, L.; Chen, H. A hybrid rice optimization algorithm. In Proceedings of the 2016 11th International Conference on Computer Science & Education (ICCSE); Institute of Electrical and Electronics Engineers (IEEE), Nagoya, Japan, 23–25 August 2016; pp. 169–174.
36. Liu, W.; Huang, Y.; Ye, Z.; Cai, W.; Yang, S.; Cheng, X.; Frank, I. Renyi's entropy based multilevel thresholding using a novel meta-heuristics algorithm. *Appl. Sci.* **2020**, *10*, 3225. [CrossRef]
37. Shu, Z.; Ye, Z.; Zong, X.; Liu, S.; Zhang, D.; Wang, C.; Wang, M. A modified hybrid rice optimization algorithm for solving 0-1 knapsack problem. *Appl. Intell.* **2021**, *52*, 5751–5769. [CrossRef]
38. Ye, Z.; Liu, S.; Zong, X.; Shu, Z.; Xia, X. A Band Selection Method for Hyperspectral Image Based on Binary Coded Hybrid Rice Optimization Algorithm. In Proceedings of the 2021 11th IEEE International Conference on Intelligent Data Acquisition and Advanced Computing Systems: Technology and Applications (IDAACS), Cracow, Poland, 22–25 September 2021; Volume 1, pp. 596–600.
39. Tubishat, M.; Abushariah, M.A.M.; Idris, N.; Aljarah, I. Improved whale optimization algorithm for feature selection in Arabic sentiment analysis. *Appl. Intell.* **2019**, *49*, 1688–1707. [CrossRef]
40. Jadon, S.S.; Tiwari, R.; Sharma, H.; Bansal, J.C. Hybrid artificial bee colony algorithm with differential evolution. *Appl. Soft Comput.* **2017**, *58*, 11–24. [CrossRef]
41. Houssein, E.H.; Mahdy, M.A.; Blondin, M.J.; Shebl, D.; Mohamed, W.M. Hybrid slime mould algorithm with adaptive guided differential evolution algorithm for combinatorial and global optimization problems. *Expert Syst. Appl.* **2021**, *174*, 114689. [CrossRef]
42. Bennasar, M.; Hicks, Y.; Setchi, R. Feature selection using joint mutual information maximisation. *Expert Syst. Appl.* **2015**, *42*, 8520–8532. [CrossRef]