Spatial-Spectral Hyperspectral Images Classification Based on Krill Herd Band Selection and Edge-Preserving Transform Domain Recursive Filter (KHBS-ETDRF)

Mohammed Abdulmajeed Moharram
VIT-AP University
Divya Meena Sundaram (divyameena.sundaram@gmail.com)
VIT-AP University

Research Article

Keywords: Heuristic Search, Krill Herd, Band Selection, Edge-Preserving Filter, Transform Domain

Posted Date: April 14th, 2022

DOI: https://doi.org/10.21203/rs.3.rs-1539336/v1

Spatial-Spectral Hyperspectral Images Classification Based on Krill Herd Band Selection and Edge-Preserving Transform Domain Recursive Filter (KHBS-ETDRF)

Mohammed Abdulmajeed Moharram, Divya Meena Sundaram*

Research Scholar, School of Computer Science and Engineering, VIT-AP University, Amaravati, India

*Assistant Professor Sr. Grade 1, School of Computer Science and Engineering, VIT-AP University, India

*divyameena.sundaram@gmail.com

ABSTRACT

Hyperspectral images (HSI) have recently been exploited in several aspects as they contain many contiguous narrow spectral informative-rich bands. The curse of dimensionality in hyperspectral images is an essential challenge as it possesses plenty of redundant bands that lead to the Hughes phenomenon. However, many feature selection or band selection techniques have been performed for dimensionality reduction of HSI. In this manuscript, firstly, a novel approach for spectral bands selection process is presented for hyperspectral images dimensionality reduction using Krill Herd (KH) Algorithm. However, KH is a heuristic search method that seeks to reach the optimum global solution within the search space and evade falling into the local optima. KH relies on simulating the herding behavior of krill in the sea in order to determine the most informative and relevant bands. Secondly, an Edge-preserving filter (EPF) was utilized to extract the spatial characteristics while reducing noise and obtaining a suitable smoothing that improves the performance of the classification process. Finally, the support vector machine (SVM) classifier at pixel level was performed for the classification of HSI. Moreover, the proposed work was compared to the Harmony Search (HS), Genetic Algorithm (GA), Bat Algorithm (BA), Particle Swarm Optimization (PSO), and Firefly Algorithm (FA). In addition, the classification results for overall accuracy (OA) on four popular publicly datasets were 96.54%, 98.93%, 99.78%, and 98.66% for the Indian Pines scene, the Pavia University scene, the Salinas scene, and the Botswana scene, respectively.

KEYWORDS: Heuristic Search, Krill Herd, Band Selection, Edge-Preserving Filter, Transform Domain.

I INTRODUCTION

HSI has attracted a lot of interest recently because it contains the most informative bands such as spectral discriminative information at various wavelengths. Moreover, hyperspectral imaging has been used in many diverse domains such as analysis of biological and chemical properties of plants [1], and early detection of diseases [2]. Despite the remarkable progress in hyperspectral imaging applications, the problem of dealing with high-dimensional data still constitutes a major dilemma due to the robust correlations between them, which lead to many computational complexities [3]. Hyperspectral images comprise redundant information which is described by a three-dimensional hyperspectral cube representing two dimensions of spatial information and a third dimension representing spectral information [4]. HSI comprises hundreds of adjacent bands with discriminative

spectral information. However, several studies have been conducted to decrease the number of bands while ensuring bands of high discriminative significance, which in turn remove redundant information, reduce processing time for remote sensing images, and maintain high classification accuracy [5].

Two common approaches for dimensionality reduction of HSI are feature extraction as well as feature selection. Where, the feature extraction technique using linear and nonlinear transformations is considered inefficient to reduce redundant information and also unsuitable for obtaining low-dimensional of HSI because it does not conserve the physical interpretation of the actual image. On the other hand, there is a significant correlation between adjacent spectral bands representing only 10 nm, so the enormous dimensions of HSI should be reduced using feature selection or band selection techniques to obtain low-dimensional HSI with reasonable processing time and desired classification accuracy [6]. Band selection methods are classified depending on the availability of labeled data into two major types supervised [7] and unsupervised [8] approaches. Whereas, the supervised methods assume the existence of labeled information per class for the training phase, and usually the classification performance of band selection using these methods is better compared to band selection unsupervised methods. Conversely, the unsupervised methods do not assume the existence of any information about the class label and rely on the study of statistical standards such as mutual information, correlation analysis, or various clustering techniques [9].

The band selection of HSI is achieved by determining a subset of the whole bands which are containing the most important information and then building an HSI classification model based on the selected informative-rich bands [10]. Several swarm intelligence approaches have recently been performed in the range of HSI feature selection which provided satisfactory classification results by selecting an appropriate subset of entire features such as the Artificial Bee Colony (ABC) [11], Differential Evolution (DE) [12], PSO [13], Whale Optimization Algorithm (WOA) [14], GA [15], and Cuckoo Search Algorithm (CSA) [16]. Moreover, dimensionality reduction techniques have been conducted on hyperspectral images to reduce processing time and computational complexities, such as selecting a subset of the bands based on anomaly detection of hyperspectral images which depends on selecting multiple bands simultaneously, unlike traditional band selection methods that select one multiple bands at a time [17], two methods were used for dimensionality reduction of HSI as a pre-processing step of classification process, principal component analysis as well as negative matrix factorization [18], a hybrid method consisting of principal component analysis that identifies new features as a pre-processing step and then normalized Mutual Information (nMI) measure effectively applied as feature selection [19], a locally Fisher's discriminant analysis for hyperspectral image dimensionality reduction was also introduced, which ensures that the multimodal structure is preserved [20], and a novel approach for dimensionality reduction of HSI based on locality adaptive discriminant analysis that focuses on the close points of spectral and spatial range in addition to being able to adaptively use the local manifold structure of the data in contrast to conventional linear discriminant analysis [21].

Many recent studies rely on the involvement of both spectral and spatial characteristics in the classification process in order to obtain more useful information and improve classification performance [22]. In general, spectral-spatial classifiers are split into two types. The first one represents the spectral-spatial characteristics within the classification process, where the spatial contextual features are included in the pixel-wise classification process. The second one represents the process of including spatial features after classification,

where the spatial contextual features are included in the pixel-wise post-classification process [23]. Moreover, the popular pixel-wise classification methods lack spatial information, which in turn reduces classification performance. Therefore, to overcome the previous dilemma, spatial characteristics should be integrated in the classification processes. As in [24] a watershed segmentation technique was introduced in order to integrate spatial information structures within hyperspectral image classification processes. A hyperspectral image segmentation and classification approach is presented in order to obtain a spatial-spectral classification map depending on a minimum spanning forest by the markers selected from the most reliable classified pixels [25]. Also, for the spectral-spatial characteristics classification of HSI, a multiple-classifier procedure is introduced in which many classifiers are used independently in order to maintain the pixel as a trait of the spatial area with the related class label and finally the minimum spanning forest (MSF) is created to obtain the spectral-spatial classification map [26].

Moreover, many studies have been conducted to exploit spatial information in the feature extraction stage. As in [27], extended morphological transformations that simultaneously exploit spatial and spectral characteristics have been used. Also, the spatial characteristics of high-resolution images are incorporated into a hyperspectral image classification process depending on an independent component analysis (ICA) and extended attribute profile (EAP) [28]. Further, high-resolution spatial features based on morphological area filtering were exploited by spatial information modeling as a set of pixels connected to an identical gray value, as for the spectral information that represents the original pixel value, finally kernel methods are used in the process of classifying spatial-spectral information [29]. An approach has been developed for extracting spatial and spectral characteristics based on a tensor discriminative locality alignment for dimensionality reduction HSI as well as a tensor scheme to define the spatial-spectral feature of pixel [30]. However, all the previous studies mentioned previously had a significant role in improving the classification accuracy of HSI effectively by incorporating the spatial-spectral discriminant characteristics during the feature extraction phase.

The major contributions of this manuscript are outlined as follows:

(1) A new approach is introduced for hyperspectral images dimensionality reduction by selecting the relevant and discriminatory bands using the Krill Herd algorithm.

(2) The Krill Herd Algorithm brings together global and local strategies that act together in parallel. This in turn improves the algorithm's overall performance for selecting the most informative-rich bands.

(3) The Krill Herd algorithm contains crossover and mutation operators that exchange the rich information between krill individuals which in turn increases the random diversity within the search space, gives a global search optimization approach as well as avoids falling into the local optima.

(4) The Krill Herd Algorithm showed a high performance with a few relevant selected bands compared with the competing algorithms, as it showed the superiority of the proposed work on five common optimization algorithms: HS, BA, PSO, FA, and GA.

(5) The spectral information extracted from the Krill Herd algorithm is combined with the spatial information extracted by an edge-preserving filter in the classification process which in turn improves classification accuracy.

(6) The proposed work showed high accuracy taking into account the limited training samples, which represent 10 % of the labeled samples in our proposed work which was performed effectively on four standard common hyperspectral imaging datasets.

The manuscript was arranged as follows: the introduction was presented in section I, pursued by the Krill Herd algorithm in section II. The Edge-preserving filter was introduced in section III, followed by the proposed methodology in section IV, then the experimental results in section V, and ultimately the conclusion in section VI.

II KRILL HERD ALGORITHM

The Krill Herd (KH) algorithm is represented by performing the movement process, which includes three basic effects that assist in reaching the optimal solution within the search space. Moreover, there are also genetic factors consisting of crossover and mutation that will enhance the performance of the KH algorithm [31].

2.1 Movement Process of The KH Algorithm:

It represents a biologically inspired approach, emulating the herding behavior of krill in the sea. The movement of krill depends on three main effects:

(1) the motion caused by other krill individuals (the mutual interaction between other individuals);
(2) foraging activity (the orientation towards the location of the food);
(3) and arbitrary diffusion (the random search process).

The movement of the krill can be described by generalizing the Lagrangian model to an n-dimensional space through equation 1, which includes the previous three main effects together:

$$\frac{dP_i}{dt} = I_i + O_i + R_i \quad (1)$$

Where; I_i is the motion caused by the mutual interaction between other individuals; O_i is the foraging activity, and R_i is the arbitrary diffusion of the individuals. Where, the motion caused by the mutual interaction between other individuals and the foraging motion comprises two local and global strategies that act in parallel, making it a robust algorithm. As for arbitrary diffusion, it performs a random search to speed up access to the optimal solution. The position of the krill individuals can be referred to by equation 2, which constitutes the movement process during the interval t to t+Φt.

$$P_i(t + \Phi t) = P_i(t) + \Phi t * \frac{dP_i}{dt} \quad (2)$$

Where; $P_i(t)$ represents the previous position of the krill, and DT is considered one of the most significant constants in the algorithm because it represents a scale component of the velocity vector, and it can be indicated by equation 3:

$$\Phi t = \frac{\sum_{i=1}^{NB} |U_i - L_i|}{NB * J_{max}} \quad (3)$$

Where; NB is the number of selected bands, L_i and U_i are the lower and upper limits of the krill individuals, which represent the values of /1/ and the overall number of bands, respectively, and J_{max} is the maximum iteration.

2.1.1 Motion Caused by Other Krill Individuals:

The Krill individuals seek to preserve the elevated intensity and move through their mutual influences between individuals. The direction of movement caused by other krill individuals, a_i, is computed by local influence, target influence, and repulsive influence which are obtained by local swarm intensity, target swarm intensity, and repulsive swarm intensity, respectively. This motion can be illustrated for the krill individual as in equations 4, 5, and 6:

$$I_i^{new} = n^{max} * a_i + w_{iter} * I_i^{old} \quad (4)$$

Where;

$$w_{iter} = 0.1 + 0.8 * \left(1 - \frac{J}{J_{max}}\right) \quad (5)$$

$$a_i = a_i^{local} + a_i^{target} \quad (6)$$

and n^{max} is maximum speed caused by other krill individuals, w_{iter} is the inertia weight of the motion caused by the mutual interaction between other individuals that lies within the range between [0, 1], I_i^{old} is the previous motion caused, a_i^{local} is the local influence supplied by neighbors, a_i^{target} is the impact of the objective direction supplied by the best krill, J is the present iteration and J_{max} is the maximum iteration.

2.1.2 Foraging Motion:

The strategy of foraging movement is denoted by two basic parameters, the first parameter represents the location of the food and the second parameter represents the past experience about the location of the food, where the acceleration coefficient was added e, which will increase the exploration within search space to reach the food location effectively. The formula for searching for food can be referred to as in equations 7, 8, and 9:

$$O_i^{new} = S_f * b_i + e * w_{iter} * O_i^{old} \quad (7)$$

Where;

$$e = (-1)^{iter} \quad (8)$$

$$b_i = b_i^{food} + b_i^{best} \quad (9)$$

and S_f is the speed of foraging, w_{iter} is the inertia weight of the foraging motion that lies in the range between [0, 1], O_i^{old} is previous foraging motion, b_i^{food} is food attractive coefficient and b_i^{best} is the influence of best fitness for krill.

2.1.3 Arbitrary Diffusion:

Arbitrary diffusion represents a random process of krill individuals. This movement can be defined by two basic parameters: the maximum diffusion velocity, and the arbitrary directional vector. This movement can be referred to by equation 10:

$$R_i = R^{max} * \left(1 - \frac{J}{J_{max}}\right) * \delta \quad (10)$$

Where; R^{max} is the maximum diffusion velocity, δ is the arbitrary directional vector that ranges between [-1, 1], J is the present iteration and J_{max} is the maximum iteration.

2.2 Genetic Factors:

The genetic reproduction mechanisms in the KH algorithm have an essential significance to improve the algorithm performance and it is inspired by the traditional differential evolution algorithm, which represents crossover and mutation operators [31].

2.2.1 Crossover:

The crossover operator is an effective global search approach inspired by the genetic algorithm (GA). It mainly depends on mixing the representations of the solutions with each other in order to converge towards the optimal solution within the search space. Figure 1 illustrates the mechanism used of the crossover operator in our proposed work, where a one-crossover point was employed, which depends on the exchange of information among krill individuals to increase diversity in the global search process and obtain informative-rich bands. Firstly, a random crossover point between values [1, 10] is selected, where /10/ represents the number of discriminant bands selected in our proposed work. Secondly, the parents are randomly selected from the krill individuals for crossover in order to exchange information between them. Moreover, it enhances the process of diversity and thus obtains informative-rich bands effectively. Where [A1 A2 ... A10] and [B1 B2 ... B10] are two representations of randomly selected solutions from the population of krill individuals for the first and second parents, respectively, which are the set of the bands.

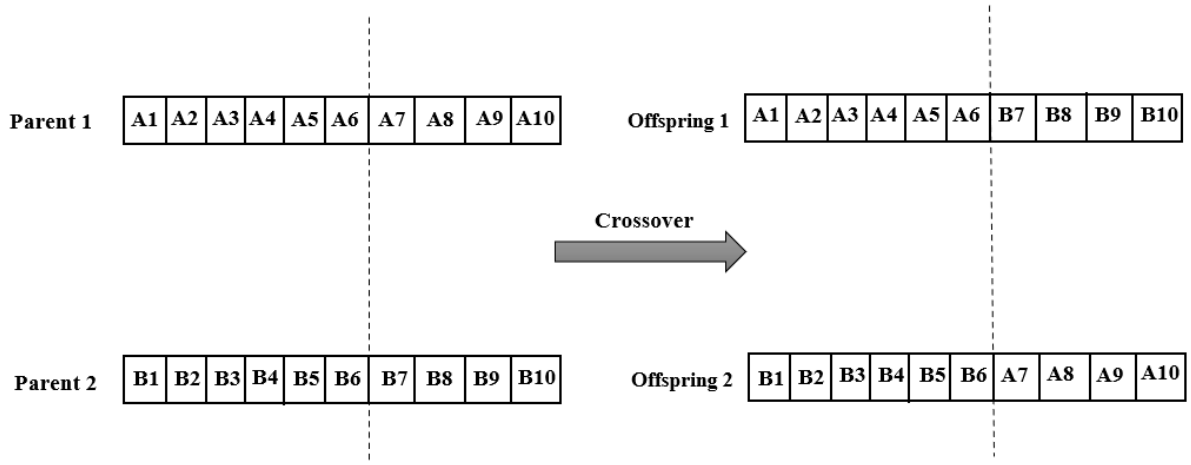


Figure 1: crossover Process

2.2.2 Mutation:

Mutation helps to increase the random diversity within the search space as well as escape from the local optima [32]. The mutation is also a significant factor in many evolutionary algorithms and is obtained by adjusting some parts of the solution. Figure 2 represents the mutation mechanism in our proposed work. Firstly, the location of the mutation is randomly selected between [1, 10], where /10/ represents the number of discriminant bands selected in our proposed work. Secondly, one solution from the krill individuals is selected randomly. Where the mutation process is performed by adjusting the mutation location at the randomly selected solution with a randomly selected value between the values 1 and the total number of bands. Where [A1 ... A6 A7 A8 A9 A10] is the randomly selected parent from the Krill individuals, and [A1 ... A6 A11 A8 A9 A10] is the offspring after the mutation process is performed.

Figure 2: Mutation Process

Algorithm 1 represents the main steps for selecting the most significant spectral bands by the Krill Herd (KH) algorithm.

Algorithm 1: Band Selection Based on Krill Herd (KH)

Input: Hyperspectral data, number of krill individuals, and positions of krill individuals.

Output: The optimum solution according to the best fitness function (subset of selected bands).

Begin:

1. Initializing parameters, maximum iteration, and determining the lower and upper limits of bands.
2. Generating an initial random population of Krill positions in the search space (initialize a subset of bands).
3. Computing the fitness function (overall accuracy).
4. Choosing the best position based on the best fitness function.

For I = 1: Maximum Iteration

5. Motion caused by other krill individuals using equation No. 4.
6. Foraging motion using equation No. 7.
7. Arbitrary diffusion using equation No. 10.
8. Calculating the new movement of the krill position using equation No. 2.
9. Performing genetic reproduction operators (crossover and mutation).
10. Updating the positions of krill individuals based on the fitness function.
11. Updating the best krill position based on the best fitness function.

End For

12. Printing the global best solution (subset of selected bands).

End

III EDGE-PRESERVING FILTER

Filters are one of the necessary operations in image processing that aim to obtain images without noise, smooth images, and preserve edges. Recently, EPFs have been used effectively in many applications, such as a bilateral filter that can smooth the colors in images and reduce delusion colors as well as preserve edges according to human perception [33], an edge-preserving model is introduced based on multi-scale image decomposition for multi-scale detail extraction using a weighted least squares optimization method [34], a guided filter that contains a rapid and non-approximate linear time algorithm which is derived from a linear model, has been shown to effectively improve detail, preserve and smooth edges [35], an approach based on merging several images to generate an information-rich image has been proposed. It consists of two layers, the base layer which contains immense scale divergences in density, and the detail layer comprise small scale details, in addition to guided filtering-based weighted average technique to fully exploit spatial features [36], and an approach has been proposed to classify HSI based on the amalgamation of the spatial and the spectral characteristics by three basic steps where the hyperspectral images are classified based on support vector machine classifier (SVM) and several probability maps are generated. In the next step, the EPF is applied to the resulting probability maps, and in the last step, the label of every pixel is decided based on the highest probability [37].

3.1 Transform Domain Recursive Filter (TDRF)

This section will explain an effective edge-preserving filter for incorporating the spatial information into the classification process of HSI which is called edge-preserving transform domain recursive filter [38]. An approach has been introduced to classify spatial-spectral hyperspectral images that relies on image merging and a TDRF to efficiently extract the characteristics. Where the HSIs are divided to multiple subsets of contiguous bands and then the contiguous spectral bands are combined for each subset by averaging and finally a TDRF is utilized to prepare the features for the classification phase [39]. Furthermore, the spectral-spatial characteristics of hyperspectral images are determined by applying an TDRF several times with different parameters and then the results of edge-preserving filters are stacked together and finally the spectral dimension of the stacked filters is decreased by the principal component analysis [40]. The transformed signal is processed through edge-preserving transform domain recursive filter as in equation 11:

$$J[z] = (1 - F^{L}) * E[z] + F^{L} * J[z-1] \tag{11}$$

Where; $J[z]$ is the filtered outcome, and the feedback coefficient F is calculated by equation 12:

$$F = \exp\left(\frac{-\sqrt{2}}{\sigma_H}\right);\ K \in [0,1] \tag{12}$$

Where; the standard deviation σ_{H_i} is calculated to find the feedback coefficient K, which represents the standard deviation of the kernel used in the iteration(i) as in equation 13:

$$\sigma_{H_i} = \sigma_h * \sqrt{3} * \frac{2^{Tot-i}}{\sqrt{4^{Tot}-1}} \tag{13}$$

Where; Tot is the overall number of iterations and σ_h is the standard deviation of the desired kernel. Moreover, L is estimated by the distance between the neighboring samples x_z and x_{z-1} in transform domain (Ω_ω) which represents the difference between the two values $G(x_z)$ and $G(x_{z-1})$ as in equation 14:

$$L = G(x_z) - G(x_{z-1}) \tag{14}$$

The domain transform of the input signal E(x) through the G(u) function which is used to compute the distance L is defined by equation 15:

$$G(u) = \int_0^u 1 + \frac{\delta_{spatial}}{\delta_{range}} |E'(x)| \, dx \quad (15)$$

Where the input signal E(x) is converted to the transform domain (Ω_ω); $u \in \Omega_\omega$. Also, $E'(x)$ represents the derivative of the input signal E(x) as well as the spatial and range parameters $\delta_{spatial}$ and δ_{range} of the edge-preserving filter, respectively.

IV PROPOSED METHODOLOGY

HSI has a tremendous number of adjacent spectral bands that give rise to the Hughes phenomenon [41]. However, there are many redundant bands that should be removed in order to minimize the processing time and the computational complexity and also improve classification performance. Schematic figure 3 demonstrates the main steps of the proposed work as it mainly encompasses three steps: the first step represents the process of hyperspectral images dimensionality reduction by effectively selecting the most informative-rich discriminant bands using a biologically inspired heuristic search algorithm called Krill Herd (KH). The second step is the feature extraction process, where spatial information is extracted by using an edge-preserving filter called transform domain recursive filter, which significantly reduces noise and remarkably improves the classification accuracy. The last step is to classify HSI using the SVM algorithm and the evaluation process.

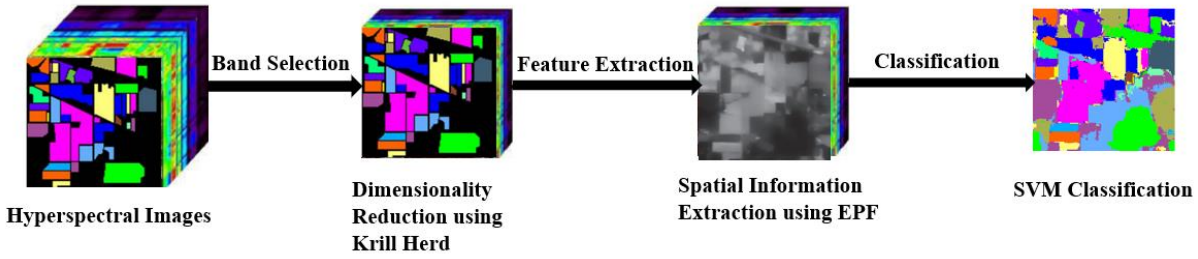


Figure 3: Schematic representation of spatial-spectral hyperspectral images classification using KHBS-ETDRF

Figure 4 also illustrates an overall flowchart of the proposed work. The major points will be explained in detail as follows:

Band Selection Using KH:

The hyperspectral images are prepared for the process of selecting the relevant and most informative bands using the Krill Herd (KH) algorithm, where the basic parameters of the algorithm are initialized as in table 1:

Table 1: Basic parameters of the KH

Parameter	Value
Number if Krills	5
Dimensions of the Krill Individuals (number of the selected bands)	10
Maximum Iteration	250
n^{max}	0.5
S_f	0.5
R^{max}	0.5

Firstly, the primitive population is initialized with unique random values which lie between the value /1/ and the total number of bands that describe the smallest number of bands and the largest number of bands, respectively. Table 2 determines the lower and upper limit of the population according to the four datasets used in this work.

Table 2: The lower and upper limit of the primitive population

Dataset	Lower Limit	Upper Limit
Indian Pines	1	200
Pavia University	1	103
Salinas	1	204
Botswana	1	145

Secondly, the fitness function is calculated for each solution (a subset of bands) in the population that represents the overall accuracy (OA) by passing the subset of bands (the ten most informative and relevant bands) out of the total bands as in equation 16:

$$\text{Overall Accuracy (OA)} = \frac{A}{B} \tag{16}$$

Where; A is the overall number of correctly classified pixels, and B is the overall number of pixels. After the initialization process, the KH algorithm works effectively through the search space to find the best solution (the ten most significant discriminatory bands) during /250/ iterations. Moreover, the basic motion of the krill algorithm is influenced by three main movements: motion caused by the mutual interaction between other individuals, foraging activity, and Arbitrary diffusion. Where the movement caused by the mutual interaction between other individuals and the foraging motion possess local and global search strategies that work in parallel to obtain the informative-rich bands. As for Arbitrary diffusion, it represents a random search process to effectively find the most relevant bands. In addition, the KH algorithm is characterized by genetic reproduction factors represented by crossover and mutation that enhance the performance of the KH. Moreover, crossover and mutation play a significant role in increasing the diversity within the search space and exchanging significant information between krill individuals, and preventing falling into the local optima.

Extraction of The Spatial Information:

The spatial information has outstanding significance in the classification process of HSI. In this paper, an edge-preserving filter called a transform domain recursive filter has been employed to remove noise, small details, and weak edges from HSI. Recursive filter aims to improve classification accuracy by preserving strong edges and effectively smoothing HSI. A transform domain recursive filter can be indicated by equation 17:

$$O = \text{Recursive filter}\left(IM,\ \delta_{spatial},\ \delta_{range}\right) \tag{17}$$

Where; IM is the input image, O is the resulting filtered image, and $\delta_{spatial}$ and δ_{range} are the spatial standard deviation and the range standard deviation of the edge-preserving filter, respectively. After the relevant and significant spectral bands are selected by the KH algorithm, the hyperspectral images are smoothed and strong edges are preserved through the transform domain recursive filter. Moreover, by incorporating spectral and spatial characteristics, the HSI classification process can be effectively improved.

SVM Classification:

After extracting the relevant and the most informative bands by the KH algorithm as well as extracting the spatial features through the transform domain recursive filter, the hyperspectral images are classified using the SVM classifier. Moreover, the most popular pixel classifier is the SVM classifier and has shown remarkable performance in classifying HSI and also in terms of high classification accuracy. Moreover, the SVM classifier is distinguished by its superior ability to handle the dataset dimensions effectively [42, 43].

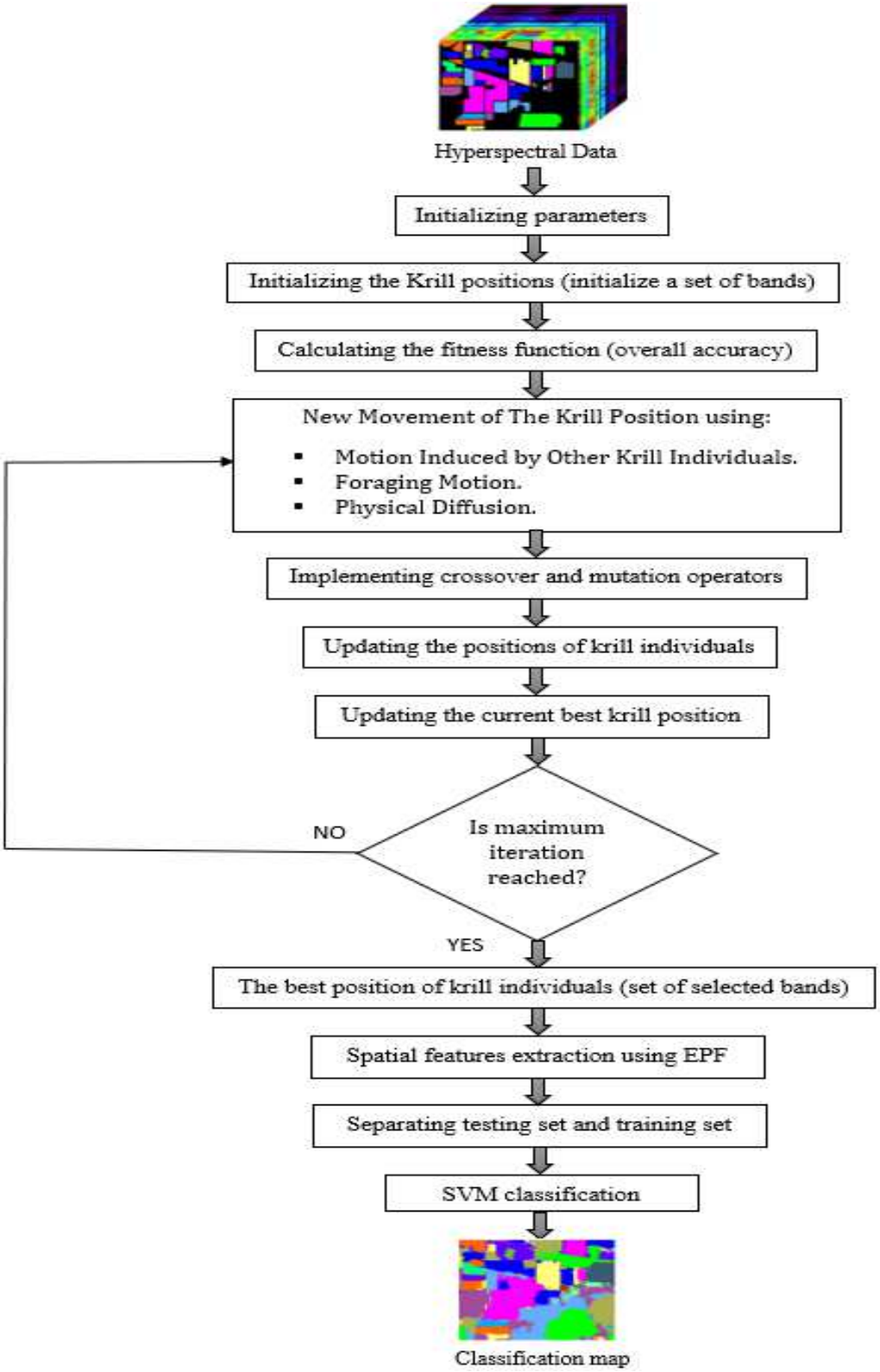


Figure 4: Overall flowchart of the proposed work

V EXPERIMENTAL RESULTS

The classification results of hyperspectral images obtained from four benchmark datasets (Indian Pines, Pavia University, Salinas, and Botswana) will be discussed for the proposed work in comparison with the five optimization algorithms (HS, BA, PSO, FA, and GA). In addition to the convergence analysis of the competing algorithms compared to the proposed work during the selection process of the most relevant discriminant bands.

5.1 Datasets Description:

5.1.1 Indian Pines Dataset

It was gathered from an agricultural area in northwestern Indiana in 1992, using AVIRIS sensor. The HSI of the Indian Pines scene contains 224 spectral bands and the image has 145*145 pixels. The spectral wavelength range between [0.4, 2.5] µm. Moreover, 24 bands representing water absorption spectral bands and bad bands are eliminated, thus the number of bands utilized in the investigations becomes 200 spectral bands. The ground truth of this scene comprises 16 classes as in table 3, which shows the division of data used for training and testing in this work.

Table 3: Training and testing data for Indian Pines scene

Class NO.	Training Sample	Testing Samples
1	5	41
2	143	1285
3	83	747
4	24	213
5	49	434
6	73	657
7	3	25
8	48	430
9	2	18
10	98	874
11	246	2209
12	60	533
13	21	184
14	127	1138
15	39	347
16	10	83
Total	1,031	9,218

5.1.2 Pavia University Dataset

It was collected from an urban region around the University of Pavia in 2002, using ROSIS sensor. The HSI of the University of Pavia scene contains 115 bands and the image has 610*340 pixels. The spectral wavelength range between [0.43, 0.86] µm. Moreover, 12 bands representing water absorption bands are eliminated, thus the number of bands used in the investigations becomes 103 bands. Where, the ground truth of this scene comprises 9 classes as in table 4, which shows the division of data used for training and testing in this work.

Table 4: Training and testing data for Pavia University scene

Class NO.	Training Sample	Testing Samples
1	664	5967
2	1865	16784
3	210	1889
4	307	2757
5	135	1210
6	503	4526
7	133	1197
8	369	3313
9	95	852
Total	4281	38495

5.1.3 Salinas Dataset

It was collected from an agricultural area along Salinas Valley, Southern California in 1998, using the AVIRIS sensor. The HSI of the Salinas scene contains 224 spectral band and the image has 512*217 pixels. The spectral wavelength range between [0.4, 2.5] µm. Moreover, 20 bands representing water absorption bands are eliminated, thus the number of bands used in the investigations becomes 204 spectral bands. Where, the ground truth of this scene comprises 16 classes as in table 5, which shows the division of data used for training and testing in this work.

Table 5: Training and testing data for Salinas scene

Class NO.	Training Sample	Testing Samples
1	201	1808
2	373	3353
3	198	1778
4	140	1254
5	268	2410
6	396	3563

7	358	3221
8	1128	10143
9	621	5582
10	328	2950
11	107	961
12	193	1734
13	92	824
14	107	963
15	727	6541
16	181	1626
Total	5418	48711

5.1.4 Botswana Dataset

It was gathered from the Okavango Delta in 2001, using the NASA EO 1 satellite as well as the Hyperion sensor. The HSI of the Botswana scene contains 242 spectral bands and the image has 1476*256 pixels. The spectral wavelength range between [0.4, 2.5] μm. Moreover, 97 bands representing water absorption bands are eliminated, thus the number of bands used in the investigations becomes 145 spectral bands. The ground truth of this scene comprises 14 classes as in table 6, which shows the division of data used for training and testing in this work.

Table 6: Training and testing data for Botswana scene

Class NO.	**Training Sample**	**Testing Samples**
1	27	243
2	11	90
3	26	225
4	22	193
5	27	242
6	27	242
7	26	233
8	21	182
9	32	282
10	25	223
11	31	274
12	19	162
13	27	241
14	10	85
Total	331	2917

5.2 Evaluation Criteria

The evaluation criteria were employed for the assessment process of the proposed work and the competing optimization algorithms, which are individual class accuracies (ICA), overall accuracy (OA), average accuracy (AA), and kappa coefficient (KC) [44].

5.3 Classification Results

It was obtained from the Indian Pines scene as in table 7, the Pavia University scene as in table 8, the Salinas scene as in table 9, and the Botswana scene as in table 10. Where each table shows ICA, OA, AA, and KC computed for the proposed work compared with five optimization algorithms. As shown in tables 7, 8, 9, and 10, we can find that the proposed work has the highest accuracy in classifying hyperspectral images effectively compared with all competing optimization algorithms due to the robustness of the proposed work in selecting the most informative and relevant discriminatory bands. However, there is only a slight misclassification in the proposed work for a small number of classes compared with all competing algorithms due to the intelligent behavior of the krill individuals in reaching the global solution within the search space and selecting the most informative discriminatory bands.

5.3.1 Analytical Results of Indian Pines Dataset

In table 7, the proposed method has shown an outstanding improvement in classification accuracy. For instance, the classification accuracy of the Grass-pasture- mowed class is 100% which has surpassed all competing algorithms, whereas the classification accuracy ranges between 47.92% and 95.65% for all competing algorithms. Moreover, there is a high improvement in the OA of the proposed work, which is 96.54%, significantly superior to all state-of-the-art algorithms, which ranged between 76.81% and 80.49%.

5.3.2 Analytical Results of Pavia University Dataset

In table 8, the proposed method significantly excelled in classifying the hyperspectral images with the highest classification accuracy. For instance, the classification accuracy for the Bare soil class is 99.98% which is the best classification accuracy compared with all competing algorithms, where the classification accuracy ranged between 87.78% and 92.60% for all competing algorithms. As for the overall accuracy of the proposed work, which is 98.93%, it clearly surpassed all competing algorithms, as it ranged between 89.93% and 93.72%.

5.3.3 Analytical Results of Salinas Dataset

Similarly in table 9, the proposed method showed a remarkable performance in classifying the hyperspectral images. For instance, the classification accuracy of the Broccoli green weeds 2 class, the Soil vineyard develop class, the Corn senesced green weeds class, the Lettuce romaine 4 wk class, the Lettuce romaine 6 wk class, and the Vineyard vertical trellis class are 100% in the proposed work which has outperformed classification accuracy compared with all competing algorithms. As for the competing algorithms, where the classification accuracy of the Broccoli green weeds 2 class ranged between 99.17% and 99.67%, the classification accuracy of the Soil vineyard develop class ranged between 98.61% and 99.38%, the classification accuracy of the Corn senesced green weeds class ranged between 95.33% and 97.66%, the classification accuracy of the Lettuce romaine 4 wk class ranged between 94.60% and 99.05%, the classification accuracy of the Lettuce romaine 6 wk class ranged

between 96.25% and 99.88%, and the classification accuracy of the Vineyard vertical trellis class ranged between 99.01% and 99.63%. Furthermore, the proposed work outperformed with a considerable improvement in terms of the overall accuracy which is 99.78% compared with all competing algorithms, which ranged between 91.86% and 93.23%.

5.3.4 Analytical Results of Botswana Dataset

Finally, in table 10, the proposed method significantly selected the most relevant bands which leads to classifying the hyperspectral images with high accuracy. For instance, the classification accuracy of the Short mopane class, the Acacia shrublands class, the Riparian class, the FloodPlain grasses 2 class, the FloodPlain grasses 1 class, and the Hippo grass class are 100% in the proposed work which has outperformed classification accuracy compared with all competing algorithms. As for the competing algorithms, where the classification accuracy of the Short mopane class ranged between 85.06% and 99.07%, the Acacia shrublands class ranged between 81.70% and 89.90%, the Riparian class ranged between 74.27% and 91.32%, the FloodPlain grasses 2 class ranged between 91.58% and 94.38%, the FloodPlain grasses 1 class ranged between 83.42% and 92.26%, and the Hippo grass class ranged between 85.65% and 93.13%. Moreover, the overall accuracy of the proposed work, which is 98.66 %, outperformed with a high improvement compared with all competing algorithms, which ranged between 87.21% and 91.91%.

Table 7: Classification results obtained from the Indian Pines dataset compared with the competing algorithms (in %)

Class NO.	Class Name	HS-SVM	BA-SVM	PSO-SVM	FA-SVM	GA-SVM	Proposed Method
1	Alfalfa	52.73	61.11	51.61	36.49	58.33	**97.62**
2	Corn-no till	76.95	71.20	72.98	71.40	71.65	**97.00**
3	Corn-min till	76.05	71.43	80.97	72.56	74.76	**95.74**
4	Corn	70.32	56.17	73.09	61.98	65.73	**95.07**
5	Grass-pasture	78.26	79.10	87.92	88.89	91.37	**96.74**
6	Grass-tree	88.57	85.69	87.46	82.79	88.26	**99.70**
7	Grass-pasture-mowed	88.00	47.92	91.67	95.65	64.29	**100**
8	Hay-windrowed	95.69	94.83	97.58	94.39	96.79	**99.08**
9	Oat	78.57	20.00	23.08	64.29	58.33	**94.74**
10	Soybean-no till	79.06	80.14	72.52	67.51	76.59	**94.35**
11	Soybean-min till	74.24	74.51	75.65	74.82	78.02	**95.03**
12	Soybean-clean	67.42	68.54	84.14	60.00	80.70	**93.44**
13	Wheat	85.51	83.33	87.50	98.40	87.44	**98.39**
14	Woods	92.76	90.81	90.29	90.50	90.06	**99.82**

15	Buildings-grass-trees-drives	73.13	66.67	73.75	68.95	73.30	**95.97**
16	Stone-steel- towers	**100**	**100**	98.68	95.24	**100**	98.80
OA		79.58	77.41	80.05	76.81	80.49	96.54
AA		79.83	71.97	78.06	76.49	78.48	96.97
KC		76.60	74.14	77.13	73.48	77.66	96.05

Table 8: Classification results obtained from the Pavia University dataset compared with the competing algorithms (in %)

Class NO.	Class Name	HS-SVM	BA-SVM	PSO-SVM	FA-SVM	GA-SVM	Proposed Method
1	Asphalt	94.09	90.78	92.10	93.01	93.96	**98.37**
2	Meadows	95.83	92.51	94.68	95.72	96.45	**99.88**
3	Gravel	81.82	72.03	80.28	82.40	82.42	**97.85**
4	Trees	95.53	92.92	93.74	95.43	94.34	**99.34**
5	Painted metal sheets	98.44	97.64	99.18	99.42	99.50	**99.92**
6	Bare soil	92.60	87.78	88.79	89.07	91.38	**99.98**
7	Bitumen	90.19	82.98	85.96	83.73	86.71	**90.83**
8	Self-blocking bricks	87.02	81.46	85.98	85.58	86.77	**96.77**
9	Shadows	99.88	**100**	99.88	99.88	**100**	99.02
OA		93.72	89.93	92.11	92.83	93.68	98.93
AA		92.82	88.68	91.18	91.58	92.39	97.99
KC		91.65	86.53	89.50	90.48	91.62	98.58

Table 9: Classification results obtained from the Salinas dataset compared with the competing algorithms (in %)

Class NO.	Class Name	HS-SVM	BA-SVM	PSO-SVM	FA-SVM	GA-SVM	Proposed Method
1	Broccoli green weeds 1	**100**	99.50	99.45	99.89	99.89	**100**
2	Broccoli green weeds 2	99.67	99.61	99.55	99.38	99.17	**100**
3	Fallow	97.29	97.83	98.07	96.12	97.52	**99.78**
4	Fallow rough plough	99.52	98.89	**99.60**	99.21	98.43	98.58
5	Fallow smooth	98.42	98.54	97.18	98.29	99.20	**99.71**
6	Stubble	99.97	99.97	99.92	**100**	99.97	**100**
7	Celery	99.84	99.66	99.23	99.66	**99.88**	99.63
8	Grapes untrained	83.05	78.24	80.92	83.63	82.60	**99.79**
9	Soil vineyard develop	99.38	99.25	98.61	99.09	99.28	**100**
10	Corn senesced green weeds	97.33	97.66	95.33	97.51	97.44	**100**
11	Lettuce romaine 4 wk	98.95	94.60	98.47	99.05	95.51	**100**
12	Lettuce romaine 5 wk	98.97	98.74	98.51	**99.65**	98.41	99.60
13	Lettuce romaine 6 wk	96.25	98.06	99.15	98.33	99.88	**100**
14	Lettuce romaine 7 wk	99.24	97.39	97.91	**99.36**	98.23	98.87
15	Vineyard untrained	80.22	78.85	79.23	76.81	80.36	**99.60**
16	Vineyard vertical trellis	99.32	99.01	99.25	99.63	99.26	**100**
OA		93.23	91.86	92.34	92.83	93.08	99.78
AA		96.71	95.99	96.27	96.60	96.56	99.72
KC		92.46	90.91	91.45	92.01	92.29	99.76

Table 10: Classification results obtained from the Botswana dataset compared with the competing algorithms (in %)

Class NO.	Class Name	HS-SVM	BA-SVM	PSO-SVM	FA-SVM	GA-SVM	Proposed Method
1	Exposes soils	**100**	99.59	**100**	**100**	99.59	**100**
2	Mixed mopane	93.62	**97.67**	92.71	93.75	95.65	86.54
3	Short mopane	99.07	95.20	87.25	85.06	94.04	**100**
4	Acacia grasslands	83.77	87.68	91.53	91.88	**95.16**	93.24
5	Acacia shrublands	87.32	81.70	82.20	89.90	88.64	**100**
6	Acacia woodlands	79.81	74.72	70.87	73.44	76.98	**96.79**
7	Island interior	92.44	97.02	98.26	97.37	99.56	**99.57**
8	Firescar	95.16	88.61	89.06	93.68	94.24	**99.45**
9	Riparian	83.60	77.71	81.34	74.27	91.32	**100**
10	Reeds	81.15	87.80	80.00	73.55	89.59	**99.55**
11	FloodPlain grasses 2	94.38	92.70	93.70	91.58	93.80	**100**
12	FloodPlain grasses 1	92.26	85.03	86.52	83.42	91.77	**100**
13	Hippo grass	93.13	85.65	87.62	90.74	86.97	**100**
14	Water	98.80	98.75	97.65	**100**	**100**	**100**
OA		90.26	88.48	87.76	87.21	91.91	98.66
AA		91.04	89.27	88.48	88.47	92.67	98.22
KC		89.45	87.52	86.75	86.15	91.23	98.55

The classification map of the proposed work and the five competing algorithms for the Indian Pines scene is shown in figure 5, the Pavia University scene is illustrated in figure 6, the Salinas scene is exhibited in figure 7, and the Botswana scene is depicted in figure 8. Where some misclassification can be observed in all competing algorithms due to the selection of irrelevant and insignificant bands. Compared with the proposed work, which is smoother and less noisy for all datasets due to the effectiveness of the KH algorithm in finding the most relevant and informative bands through local and global search strategies, in addition to using a transform domain recursive filter to extract the spatial characteristics effectively.

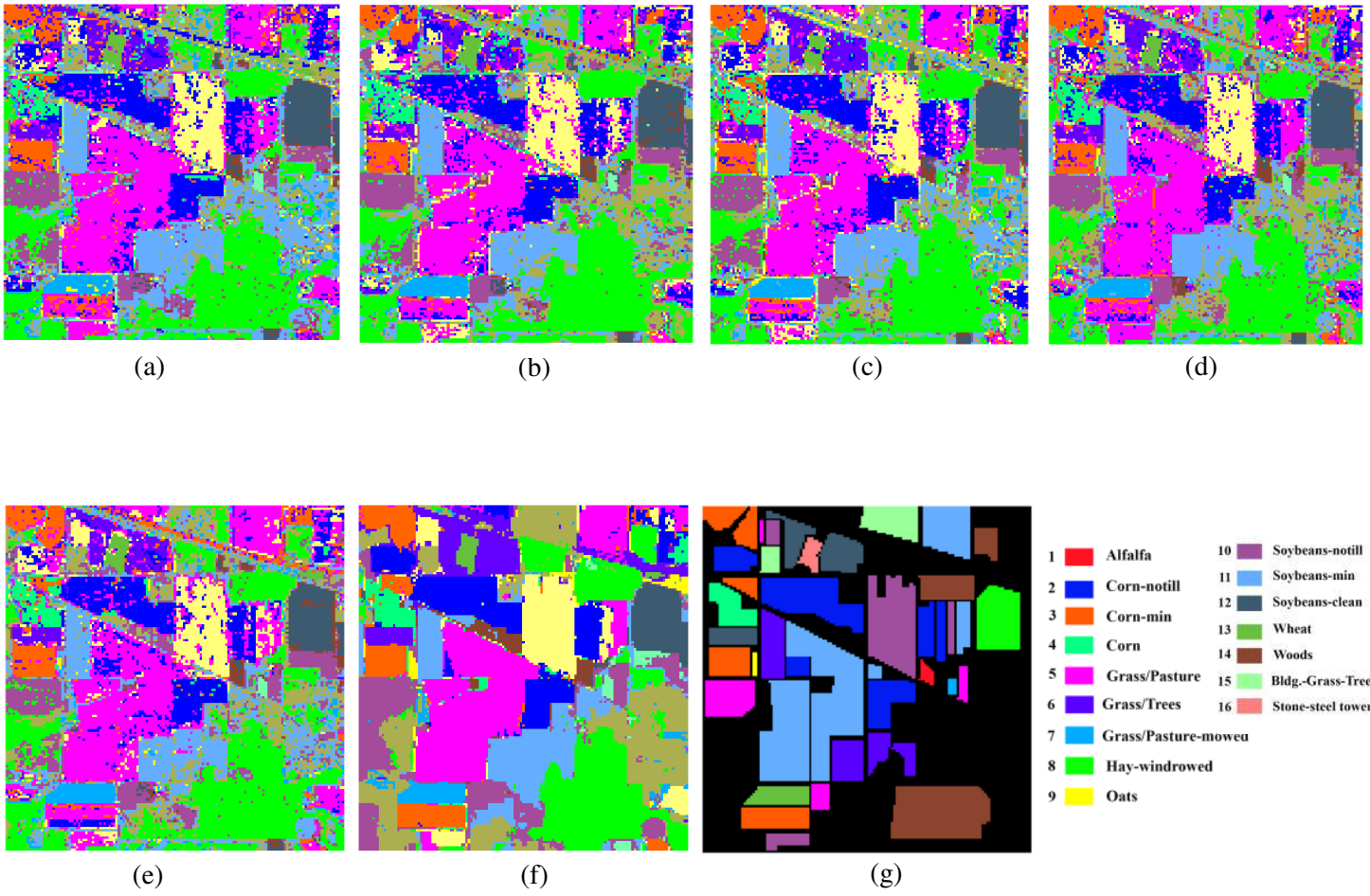
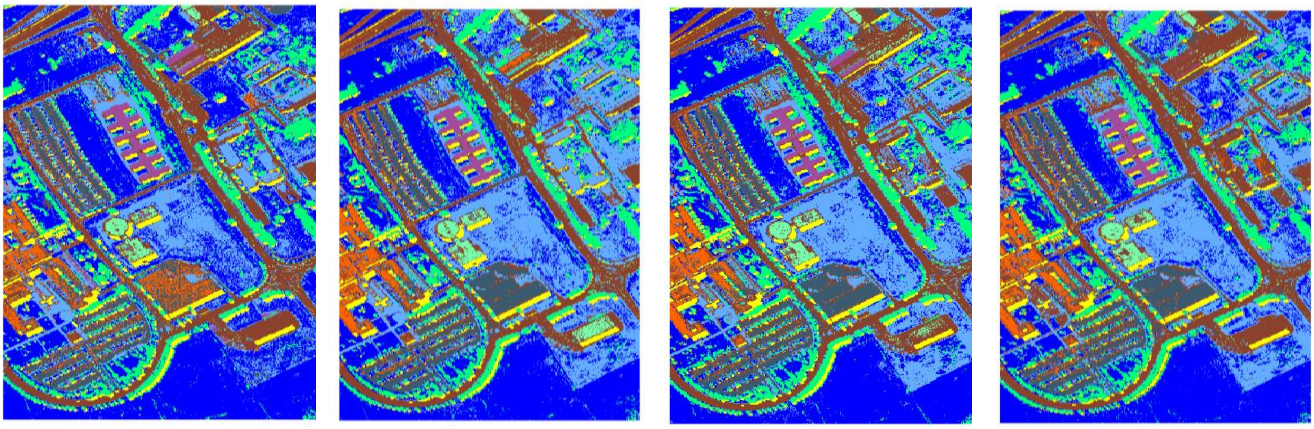
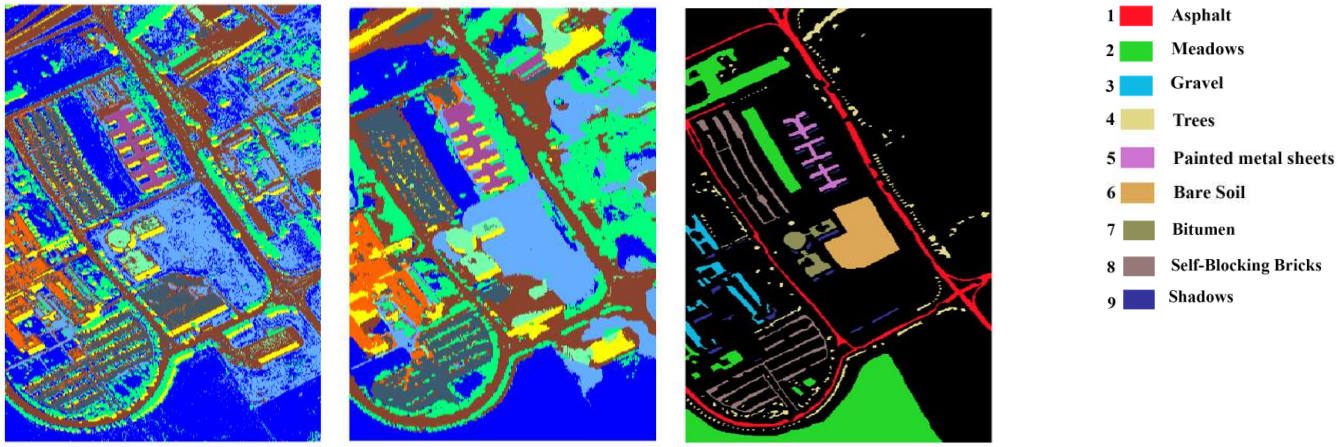


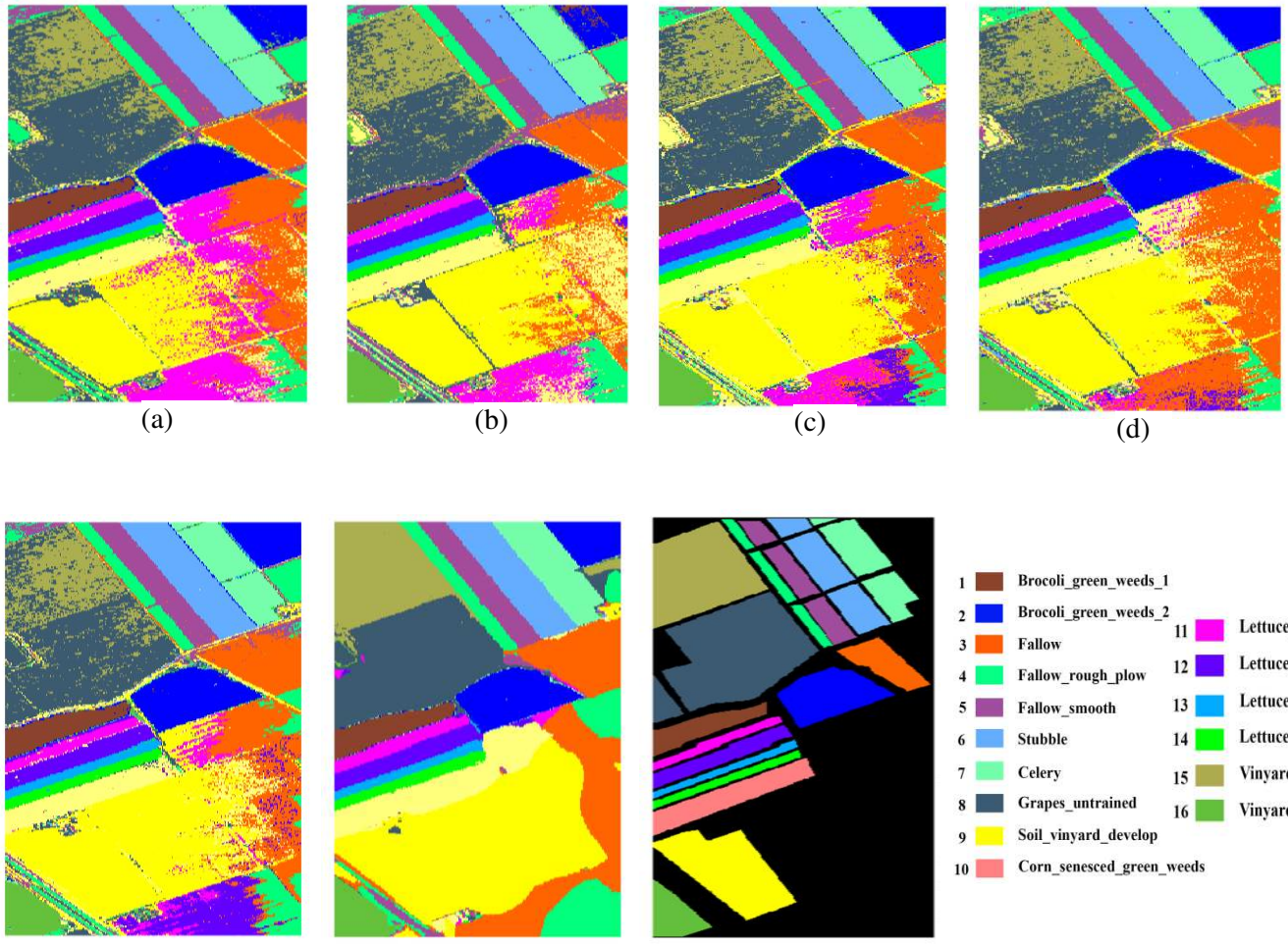
Figure 5: Classification maps on Indian pines dataset: (a) BA-SVM; (b) FA-SVM; (c) GA-SVM; (d) HS-SVM; (e) PSO-SVM; (f) Proposed Method; (g) Ground Truth.





(e) (f) (g)

Figure 6: Classification maps on Pavia University dataset: (a) BA-SVM; (b) FA-SVM; (c) GA-SVM; (d) HS-SVM; (e) PSO-SVM; (f) Proposed Method; (g) Ground Truth.



(a) (b) (c) (d)

(e) (f) (g)

Figure 7: Classification maps on Salinas dataset: (a) BA-SVM; (b) FA-SVM; (c) GA-SVM; (d) HS-SVM; (e) PSO-SVM; (f) Proposed Method; (g) Ground Truth.

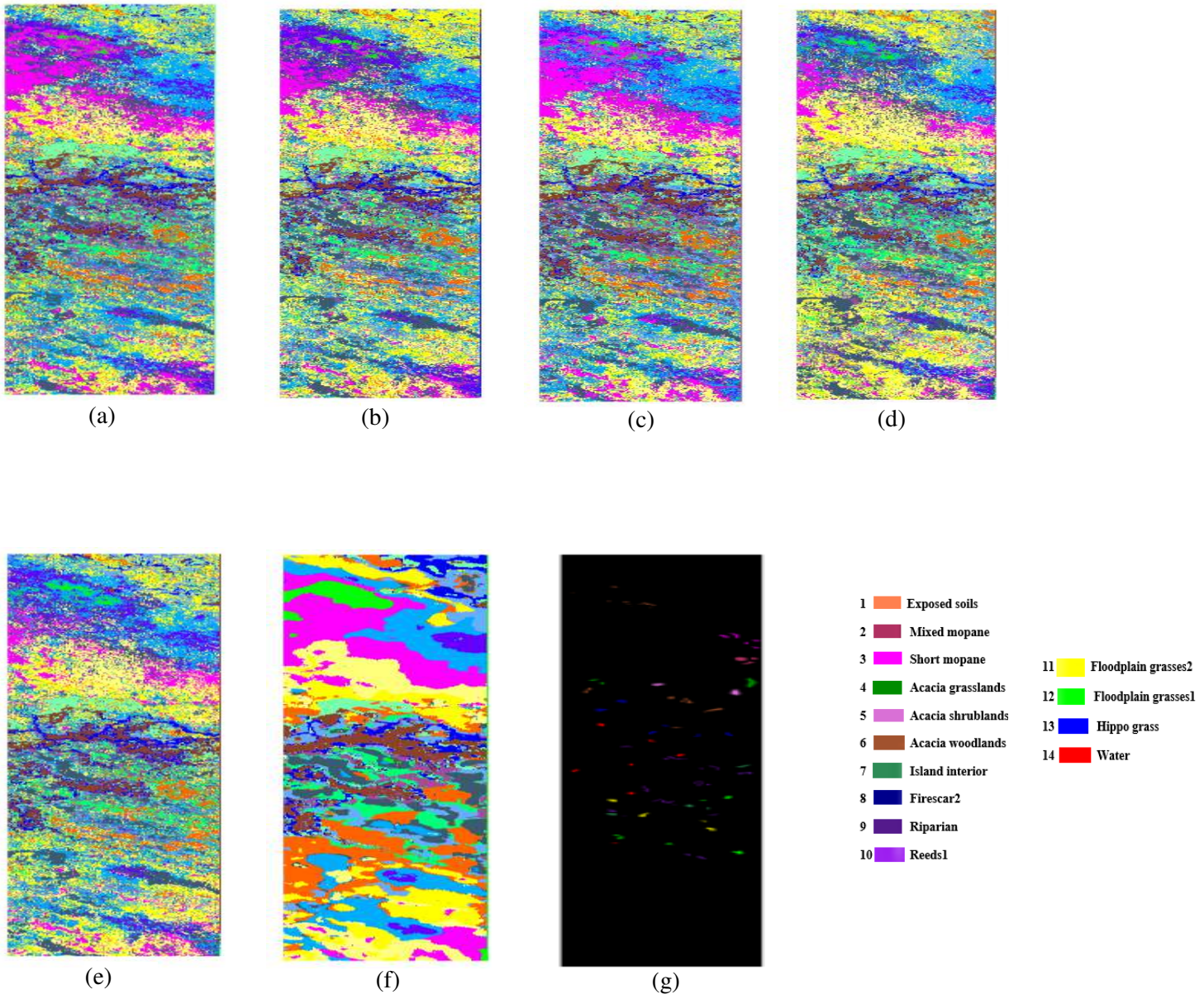


Figure 8: Classification maps on Botswana dataset: (a) BA-SVM; (b) FA-SVM; (c) GA-SVM; (d) HS-SVM; (e) PSO-SVM; (f) Proposed Method; (g) Ground Truth.

Tables 11, 12, 13, and 14 exhibit the selected spectral bands for the Indian Pines scene, the Pavia University scene, the Salinas scene, and the Botswana scene, respectively. Where these tables demonstrate the ten most informative and significant bands selected by the proposed work and the competing algorithms. Moreover, figures 9 (a), 9 (b), 9 (c), and 9 (d) indicate the distribution of the selected spectral bands obtained by the proposed work and the competing algorithms for the Indian Pines scene, the Pavia University scene, the Salinas scene, and the Botswana scene, respectively. Where the spectral bands were chosen for the Indian Pines scene in the range between 1 and 200, the Pavia University scene in the range between 1 and 103, the Salinas scene in the range between 1 and 204, and the Botswana scene in the range between 1 and 145.

Table 11: The ten most informative and significant bands selected by the proposed work and the competing algorithms for Indian Pines scene

Method	Selected Bands
HS	17 22 35 40 41 98 100 141 161 178
BA	56 27 13 172 137 79 162 70 148 178
PSO	60 43 64 123 34 178 165 20 49 53
FA	74 153 82 28 181 125 196 176 68 7
GA	7 18 165 129 164 69 21 59 32 39
Proposed Method	124 26 10 68 98 36 51 165 44 19

Table 12: The ten most informative and significant bands selected by the proposed work and the competing algorithms for Pavia University scene

Method	Selected Bands
HS	5 17 31 42 60 68 83 85 87 103
BA	8 25 30 49 79 6 85 24 69 12
PSO	27 59 87 68 37 47 93 83 19 70
FA	6 59 17 38 85 83 68 84 82 60
GA	15 99 9 28 70 38 83 65 19 86
Proposed Method	38 68 15 6 79 85 98 65 26 83

Table 13: The ten most informative and significant bands selected by the proposed work and the competing algorithms for Salinas scene

Method	Selected Bands
HS	19 38 50 52 82 96 136 140 156 169
BA	175 11 54 130 168 137 130 192 67 77
PSO	90 99 51 30 115 74 84 93 136 158
FA	4 78 141 18 32 51 102 14 162 40
GA	152 139 105 9 119 39 50 18 179 91
Proposed Method	139 52 30 111 6 88 181 38 19 12

Table 14: The ten most informative and significant bands selected by the proposed work and the competing algorithms for Botswana scene

Method	Selected Bands
HS	24 28 33 34 55 68 70 72 106 120
BA	97 42 79 45 112 49 94 85 131 27
PSO	38 54 116 62 53 61 16 65 105 27
FA	97 29 34 59 111 27 6 73 133 31
GA	31 6 122 101 27 69 98 138 27 51
Proposed Method	25 92 37 27 57 78 84 80 126 4

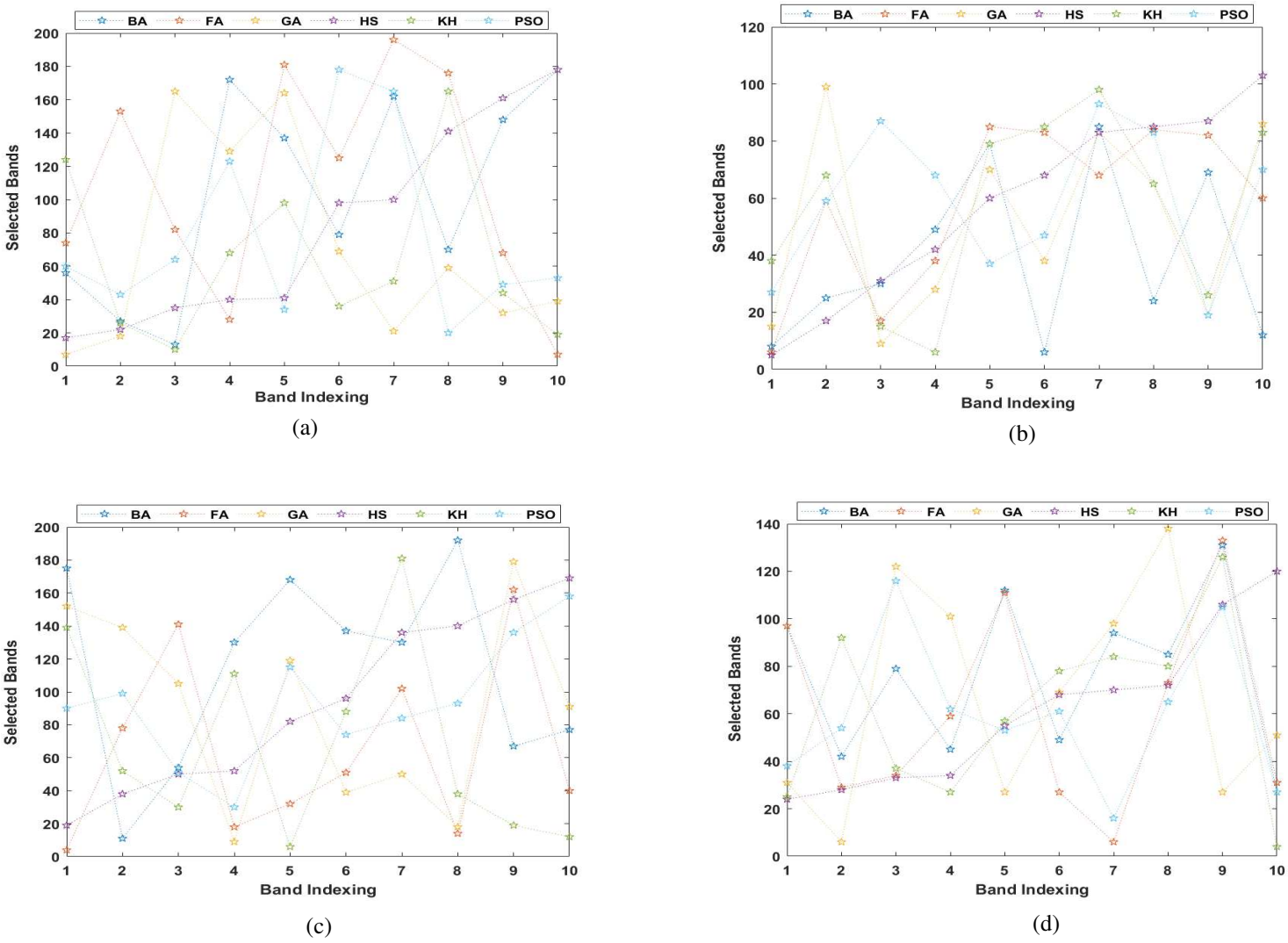


Figure 9: Distribution of the selected spectral bands obtained by the proposed work and the competing algorithms: (a) the Indian Pines scene; (b) the Pavia University scene; (c) the Salinas scene; (d) the Botswana scene.

5.3.5 Statistical Analysis of Fitness Function

Figures 10, 11, 12, and 13 illustrate the convergence behavior during the selection process of the most discriminative spectral bands for the Indian Pines scene, the Pavia University scene, the Salinas scene, and the Botswana scene, respectively. Where the convergence behavior is analyzed of the proposed work compared with the competing algorithms by plotting the number of iterations against the fitness function, which represents the OA of the hyperspectral image classification for the respective bands.

Figure 10 demonstrates the substantial superiority of the KH algorithm by reaching the best optimal solution after /250/ iteration with an overall accuracy of 74.22% compared with the competing algorithms due to discovering the most informative and relevant spectral bands. Compared with the competing algorithms, the BA algorithm has the worst overall accuracy, which is equal to 70.58%, and the remaining optimization algorithms FA, HS, PSO, and GA have an overall accuracy of 71.49%, 72.58%, 73.38%, and 73.80%, respectively.

Figure 11 exhibits the remarkable performance of the KH algorithm by obtaining the global solution as well as the optimal convergence compared with the competing algorithms, where the overall accuracy after /250/ iteration is 88.41% in addition to avoiding falling into the local optima during the searching process about the best bands. We also note the premature convergence of the BA algorithm and obtaining the worst performance with an overall accuracy of 84.92% due to being stuck into the local optima, and the other optimization algorithms PSO, FA, GA, and HS have an overall accuracy of 86.54%, 87.61%, 87.89%, and 87.95%, respectively.

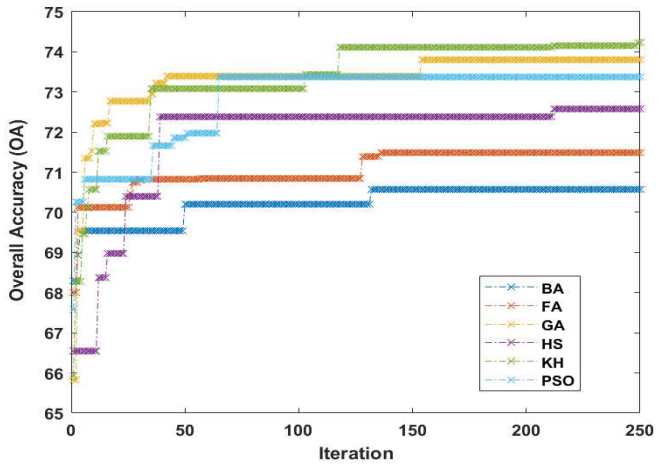


Figure 10: The best OA obtained in each iteration for the Indian Pines scene

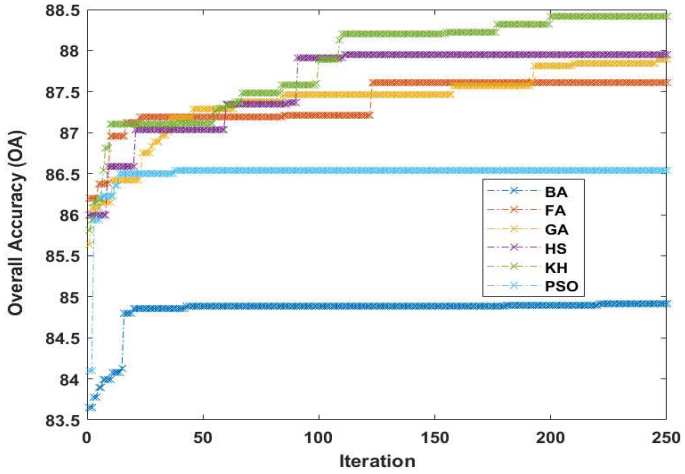


Figure 11: The best OA obtained in each iteration for the Pavia University scene

Figure 12 displays the convergence behavior of the KH algorithm which has reached the global optimal solution effectively compared with the competing algorithms with an overall accuracy after /250/ iteration of 91.47%, where the optimal solution represents the most informative spectral bands used in classifying hyperspectral images. Compared to competing algorithms, we find that the BA algorithm is being stuck into the local optima with an overall accuracy of 90.21%, and then the GA, PSO, HS, and FA algorithms come with an overall accuracy of 91.28%, 90.74%, 90.86%, and 90.98%, respectively.

Figure 13 shows a significant improvement in the convergence behavior of the KH algorithm towards the best bands used for classifying hyperspectral images with an OA after /250/ iteration of 91.94%. In contrast, the overall accuracy of the PSO algorithm is 88.86%, which represents the worst performance in the convergence process between the competing algorithms, while the overall accuracy of HS, BA, GA, and FA algorithms reached 90.16%. 90.71%, 91.53%, and 91.57%, respectively.

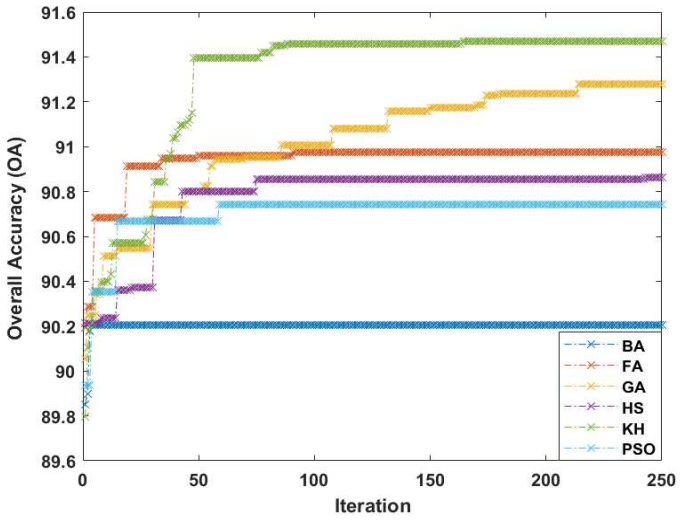


Figure 12: The best OA obtained in each iteration for the Salinas scene

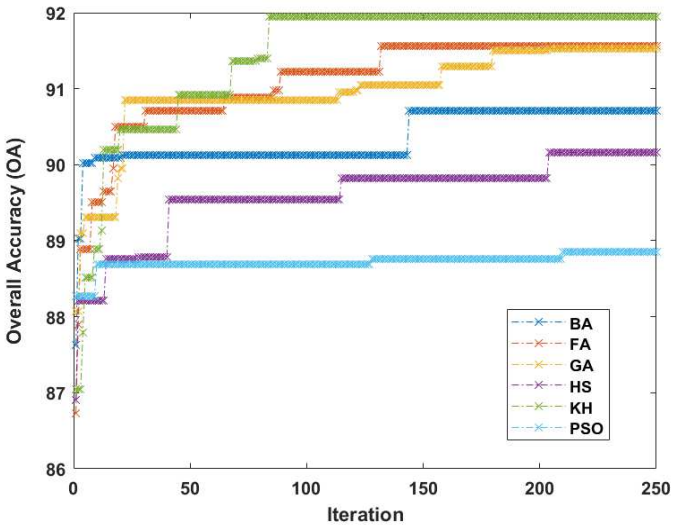


Figure 13: The best OA obtained in each iteration for the Botswana scene

Moreover, the considerable improvement in the convergence behavior of the proposed work is due to the benefit of local and global search strategies of the KH algorithm, in addition to the significant diversity in the search space towards the global optimal solution by using genetic reproduction operators. The analysis of the convergence behavior of the competing algorithms is due to the selection of irrelevant and less informative spectral bands compared with the proposed work that outperformed all the competing algorithms effectively and remarkably.

Figures 14, 15, 16, and 17 demonstrate the percentage values of OA, AA, and KC for the Indian Pines scene, the Pavia University scene, the Salinas scene, and the Botswana scene, respectively. The proposed work apparently overcomes all the competing algorithms and effectively classifies hyperspectral images.

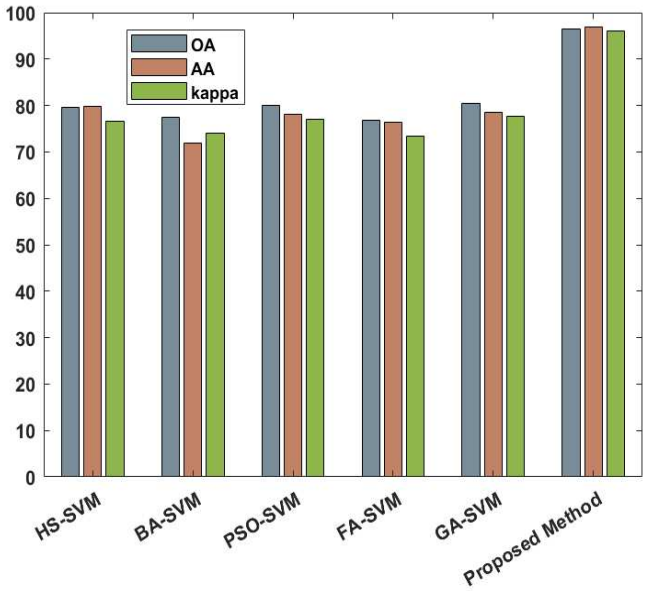


Figure 14: OA, AA, and KC values for the Indian Pines scene

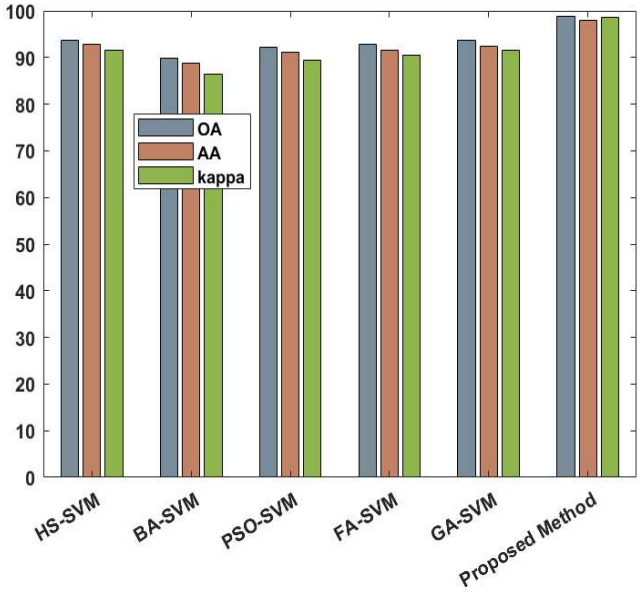


Figure 15: OA, AA, and KC values for the Pavia University scene

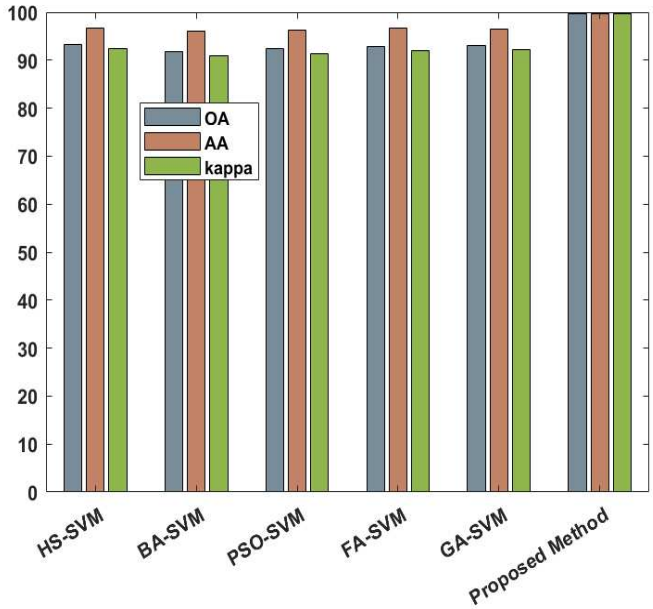


Figure 16: OA, AA, and KC values for the Salinas scene

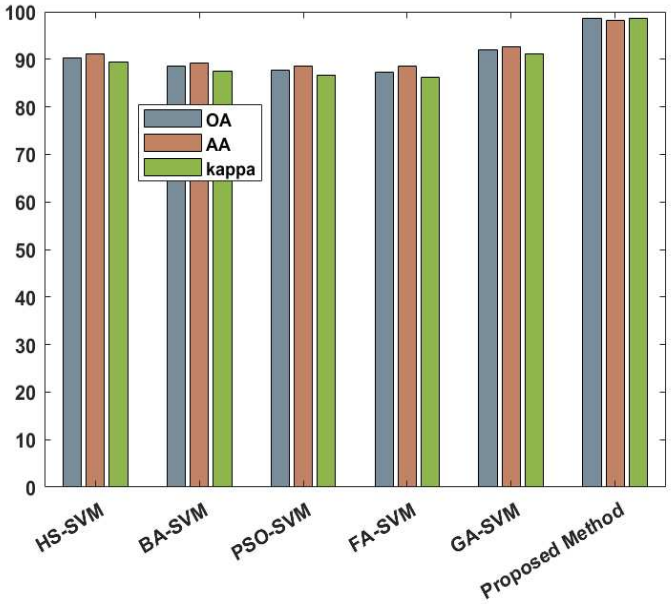


Figure 17: OA, AA, and KC values for the Botswana scene

VI CONCLUSION

The Hughes phenomenon is a considerable dilemma in the classification of HSI due to the redundant and insignificant spectral bands. In this manuscript, a novel approach is presented for HSIs dimensionality reduction by selecting the most discriminative spectral bands using the KH algorithm, which is characterized by local and global search strategies. The OA was used as a fitness function of the KH algorithm during the selection process

of the significant spectral bands. Moreover, the KH algorithm has genetic reproduction mechanisms (crossover and mutation) that provide a large diversity within the search space to effectively accelerate reaching the optimal global solution (the most informative spectral bands) as well as avoid falling into the local optima. In the next stage, a TDRF is used to extract the spatial characteristics, which significantly improves the accuracy of classifying hyperspectral images by smoothing the images and obtaining strong edges. Finally, the resulting image was classified using an SVM classifier, which is a robust classifier that is widely used to classify HSI at the pixel level. The proposed method demonstrated a substantial improvement in the classification of HSI compared to competing optimization algorithms (HS, BA, PSO, FA, and GA). The proposed method also revealed a significant superiority based on the limited training data of hyperspectral images on four benchmark datasets. In future work, new optimization techniques will be used to significantly reduce the dimensions of hyperspectral images. Furthermore, hybrid models will be employed to effectively select discriminative spectral bands as well as optimize the fitness function to enhance the classification accuracy of HSI. In addition to incorporating the spatial features in the hyperspectral image classification process by using appropriate filters to obtain better performance and high accuracy.

Ethics approval and consent to participate

Not applicable.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this manuscript.

Funding

No funding was received to assist with the preparation of this manuscript.

Authors' contributions

The manuscript was written by Mohammed Abdulmajeed Moharram and revised by Divya Meena Sundaram. All authors read and approved the final manuscript.

Data Availability Statement:

All hyperspectral image scenes utilized in this study are publicly available on the website: Hyperspectral Remote Sensing Scenes - Grupo de Inteligencia Computacional (GIC) (ehu.eus).

Acknowledgments

The authors thank the VIT-AP for providing a seed grant for carrying out this research work.

REFERENCES

[1] B. Luo, C. Yang, J. Chanussot, and L. Zhang, “Crop yield estimation based on unsupervised linear unmixing of multidate hyperspectral imagery,” IEEE Trans. Geosci. Remote Sens., vol. 51, no. 1, pp. 162–173, Jan. 2013.

[2] Terentev, Anton, et al. "Current State of Hyperspectral Remote Sensing for Early Plant Disease Detection: A Review." Sensors 22.3 (2022): 757.

[3] Wang, Qi, Fahong Zhang, and Xuelong Li. "Optimal clustering framework for hyperspectral band selection." IEEE Transactions on Geoscience and Remote Sensing 56.10 (2018): 5910-5922.

[4] Phaneendra Kumar, Boggavarapu LN, and Prabukumar Manoharan. "Whale optimization-based band selection technique for hyperspectral image classification." International Journal of Remote Sensing 42.13 (2021): 5105-5143.

[5] Li, Shijin, et al. "An effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine." Knowledge-Based Systems 24.1 (2011): 40-48.

[6] C. Liu, C. Zhao, Y. Zhang, A new method of hyperspectral remote sensing image dimensional reduction, J. Image Graph. 10 (2) (2005) 218–224.

[7] H. Yang, Q. Du, H. Su, and Y. Sheng, “An efficient method for supervised hyperspectral band selection,” IEEE Geosci. Remote Sens. Lett., vol. 8, no. 1, pp. 138–142, Jan. 2011.

[8] Sui, Chenhong, et al. "Unsupervised hyperspectral band selection with multigraph integrated embedding and robust self-contained regression." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-15.

[9] Patra, Swarnajyoti, Prahlad Modi, and Lorenzo Bruzzone. "Hyperspectral band selection based on rough set." IEEE Transactions on Geoscience and Remote Sensing 53.10 (2015): 5495-5503.

[10] Mou, Lichao, et al. "Deep reinforcement learning for band selection in hyperspectral image classification." IEEE Transactions on Geoscience and Remote Sensing 60 (2021): 1-14.

[11] E. Hancer, B. Xue, M. Zhang, D. Karaboga, B. Akay, Pareto Front feature selection based on artificial bee colony optimization, Inform. Sci. 422 (2018) 462–479.

[12] E. Hancer, Differential evolution for feature selection: a fuzzy wrapper-filter approach, Soft Comput. 10 (2018) 1–16.

[13] Paul, Arati, and Nabendu Chaki. "Band selection using spectral and spatial information in particle swarm optimization for hyperspectral image classification." Soft Computing (2022): 1-16.

[14] Phaneendra Kumar, Boggavarapu LN, and Prabukumar Manoharan. "Whale optimization-based band selection technique for hyperspectral image classification." International Journal of Remote Sensing 42.13 (2021): 5105-5143.

[15] Aghaee, Reza, Mehdi Momeni, and Payman Moallem. "Semi-Supervised Band Selection from Hyperspectral Images using Levy Flight-Based Genetic Algorithm." IEEE Geoscience and Remote Sensing Letters (2022).

[16] M.A.E. Aziz, A.E. Hassanien, Modified cuckoo search algorithm with rough sets for feature selection, Neural Comput. Appl. 29 (4) (2018) 925–934.

[17] Wang, Lin, et al. "Band subset selection for anomaly detection in hyperspectral imagery." IEEE Transactions on Geoscience and Remote Sensing 55.9 (2017): 4887-4898.

[18] Li, Tong, Junping Zhang, and Ye Zhang. "Classification of hyperspectral image based on deep belief networks." 2014 IEEE international conference on image processing (ICIP). IEEE, 2014.

[19] Hossain, Md Ali, Xiuping Jia, and Mark Pickering. "Subspace detection using a mutual information measure for hyperspectral image classification." IEEE Geoscience and Remote Sensing Letters 11.2 (2013): 424-428.

[20] Li, Wei, et al. "Locality-preserving dimensionality reduction and classification for hyperspectral image analysis." IEEE Transactions on Geoscience and Remote Sensing 50.4 (2011): 1185-1198.

[21] Wang, Qi, Zhaotie Meng, and Xuelong Li. "Locality adaptive discriminant analysis for spectral–spatial classification of hyperspectral images." IEEE Geoscience and Remote Sensing Letters 14.11 (2017): 2077-2081.

[22] Plaza, Antonio, et al. "Recent advances in techniques for hyperspectral image processing." Remote sensing of environment 113 (2009): S110-S122.

[23] Ran, Qiong, Wei Li, and Qian Du. "Spatial-spectral classification with local regional filter and Markov random field." 2015 7th Workshop on Hyperspectral Image and Signal Processing: Evolution in Remote Sensing (WHISPERS). IEEE, 2015.

[24] Tarabalka, Yuliya, Jocelyn Chanussot, and Jon Atli Benediktsson. "Segmentation and classification of hyperspectral images using watershed transformation." Pattern Recognition 43.7 (2010): 2367-2379.

[25] Tarabalka, Yuliya, Jocelyn Chanussot, and Jón Atli Benediktsson. "Segmentation and classification of hyperspectral images using minimum spanning forest grown from automatically selected markers." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 40.5 (2009): 1267-1279.

[26] Tarabalka, Yuliya, et al. "Multiple spectral–spatial classification approach for hyperspectral data." IEEE Transactions on Geoscience and Remote Sensing 48.11 (2010): 4122-4132.

[27] Plaza, Antonio, et al. "Dimensionality reduction and classification of hyperspectral image data using sequences of extended morphological transformations." IEEE Transactions on Geoscience and remote sensing 43.3 (2005): 466-479.

[28] Dalla Mura, Mauro, et al. "Classification of hyperspectral images by using extended morphological attribute profiles and independent component analysis." IEEE Geoscience and Remote Sensing Letters 8.3 (2010): 542-546.

[29] Fauvel, Mathieu, Jocelyn Chanussot, and Jon Atli Benediktsson. "A spatial–spectral kernel-based approach for the classification of remote-sensing images." Pattern Recognition 45.1 (2012): 381-392.

[30] Zhang, Liangpei, et al. "Tensor discriminative locality alignment for hyperspectral image spectral–spatial feature extraction." IEEE Transactions on Geoscience and Remote Sensing 51.1 (2012): 242-256.

[31] Gandomi, Amir Hossein, and Amir Hossein Alavi. "Krill herd: a new bio-inspired optimization algorithm." Communications in nonlinear science and numerical simulation 17.12 (2012): 4831-4845.

[32] Doerr, Benjamin. "Does comma selection help to cope with local optima?" Algorithmica (2022): 1-35.

[33] Tomasi, Carlo, and Roberto Manduchi. "Bilateral filtering for gray and color images." Sixth international conference on computer vision (IEEE Cat. No. 98CH36271). IEEE, 1998.

[34] Farbman, Zeev, et al. "Edge-preserving decompositions for multi-scale tone and detail manipulation." ACM Transactions on Graphics (TOG) 27.3 (2008): 1-10.

[35] He, Kaiming, Jian Sun, and Xiaoou Tang. "Guided image filtering." IEEE transactions on pattern analysis and machine intelligence 35.6 (2012): 1397-1409.

[36] Li, Shutao, Xudong Kang, and Jianwen Hu. "Image fusion with guided filtering." IEEE Transactions on Image processing 22.7 (2013): 2864-2875.

[37] Kang, Xudong, Shutao Li, and Jon Atli Benediktsson. "Spectral–spatial hyperspectral image classification with edge-preserving filtering." IEEE transactions on geoscience and remote sensing 52.5 (2013): 2666-2677.

[38] Gastal, Eduardo SL, and Manuel M. Oliveira. "Domain transform for edge-aware image and video processing." ACM SIGGRAPH 2011 papers. 2011. 1-12.

[39] Kang, Xudong, Shutao Li, and Jón Atli Benediktsson. "Feature extraction of hyperspectral images with image fusion and recursive filtering." IEEE Transactions on Geoscience and Remote Sensing 52.6 (2013): 3742-3752.

[40] Kang, Xudong, et al. "PCA-based edge-preserving features for hyperspectral image classification." IEEE Transactions on Geoscience and Remote Sensing 55.12 (2017): 7140-7151.

[41] Hassanzadeh, Shirin, Habibollah Danyali, and Mohammad Sadegh Helfroush. "Combined Spatial-Spectral Schroedinger Eigenmaps with Multiple Kernel Learning for Hyperspectral Image Classification Using a Low Number of Training Samples." Canadian Journal of Remote Sensing (2022): 1-13.

[42] Melgani, Farid, and Lorenzo Bruzzone. "Classification of hyperspectral remote sensing images with support vector machines." IEEE Transactions on geoscience and remote sensing 42.8 (2004): 1778-1790.

[43] Chen, Ying-Nong, et al. "Feature line embedding based on support vector machine for hyperspectral image classification." Remote Sensing 13.1 (2021): 130.

[44] Richards, John A., and J. A. Richards. Remote sensing digital image analysis. Vol. 3. Berlin: springer, 1999.