



A hybrid optimization approach for hyperspectral band selection based on wind driven optimization and modified cuckoo search optimization

Shrutika Sawant¹ · Prabukumar Manoharan¹ 

Received: 13 July 2019 / Revised: 11 July 2020 / Accepted: 25 August 2020 /
Published online: 9 September 2020
© Springer Science+Business Media, LLC, part of Springer Nature 2020

Abstract

Selection of useful bands plays a very important role in hyperspectral image classification. In the past decade, metaheuristic algorithms have been used as promising methods for solving this problem. However, many metaheuristic algorithms may provide unsatisfactory performance due to their slow or premature convergence. Therefore, how to develop algorithms well balancing the exploration and exploitation, and find the suitable bands precisely is still a challenge. In this paper, a new hybrid global optimization algorithm, which is based on the Wind Driven Optimization (WDO) and Cuckoo Search (CS) is proposed to solve hyperspectral band selection problems. Both WDO and CS have strong searching ability and require less control parameters, but easily suffer from premature convergence due to loss of diversity of population. The proposed approach uses the Chebyshev chaotic map to initialize the population at initial step. The population is divided into two subgroups and WDO and CS are adopted for these two subgroups independently. By division, these two subgroups can share suitable information and utilize each other's pros, thus avoid premature convergence, and obtain best optimal solution. Furthermore, the Levy flight step size in CS algorithm is adaptively adjusted based on fitness value and current iteration number, which helps in boosting the convergence speed of algorithm. The experimental results on three standard benchmark datasets namely, Pavia University, Botswana and Indian Pines, prove the superiority of the proposed approach over standard WDO and CS approaches as well as the other traditional approaches in terms of classification accuracy with fewer bands.

Keywords Hyperspectral image · Band selection · Wind driven optimization · Cuckoo search algorithm · Chebyshev chaotic map

✉ Prabukumar Manoharan
mprabukumar@vit.ac.in

Shrutika Sawant
shrutika.sawant1@gmail.com

¹ School of Information Technology & Engineering, Vellore Institute of Technology, Vellore, India

1 Introduction

Recent advances in imaging techniques allow hyperspectral sensors capable of gathering spectral information in hundreds of narrow and contiguous bands, covering a broad range of wavelength in the spectrum. The increasing availability of hyperspectral images has gained huge success in the field of remote sensing. Hyperspectral image is three-dimensional (3-D) image cube (the third dimension denotes spectral band) comprises huge amount of spectral as well as spatial information to recognize and discriminate spectrally unique materials [30]. The enormous spectral bands of hyperspectral images provide redundant information, further entail huge storage cost and heavy computational burden. Besides, spectral bands are strongly correlated, i.e., more number of bands may not always result in significant discriminating capability for classification purpose. Therefore, reducing the dimensionality is a vital process to discard the redundant bands and make it less time-consuming for classification [4, 29, 32].

Generally, reduction in the dimensionality can be achieved with either feature extraction or feature selection (also known as band selection). Feature extraction is the process of reducing the dimension by mapping the data from original higher dimensional space to new lower dimensional space. Approaches like, Principal component analysis (PCA), Independent component analysis (ICA), Discriminant analysis, and transform based approaches have been widely used [14, 24, 43, 44]. However, these approaches generally alter the physical meaning of original data since the data in the low-dimensional space correspond to their linear combinations but do not original individual bands. Band selection approaches to pick set of most suitable bands from pool of bands may be preferred when the physical meaning of bands needs to be preserved [28, 31, 36]. Recently, many band selection approaches have been proposed [15, 45]. There exists two most commonly used ways for selection of optimal band subset. The first strategy includes the individual band evaluation, and the second strategy includes the band subset evaluation. Individual evaluation consists of clustering based approaches and ranking based approaches. In the individual band evaluation, the score (which signifies the importance of the band) of an individual band is measured according to its degree of relevance. Band clustering approaches form cluster of similar bands and select representative band [25, 46], whereas, in ranking based band selection approaches, the bands are selected based on the ranking [8, 47]. A hierarchical clustering approach is most commonly used band clustering approach, which is based on Ward's linkage [18]. This technique employs mutual information (MI) between the bands to pick most informative bands. Kamandar and Ghasseman [36] proposed a novel band selection technique that maximizes relevance of selected bands and minimizes redundancy between them (mRMR), which further leads to excellent classification performance. Jia et al. [12] proposed a new hyperspectral band selection approach based on Relief-F algorithm and discards less informative bands.

Intelligent algorithms try to evaluate a complete subset of bands rather than individually selecting the bands. Generating such subsets is known to be an NP-hard problem. Therefore, such problems are solved by employing nature-inspired algorithms. In band selection, popular nature-inspired algorithms include Firefly algorithm [38], Artificial Bee Colony(ABC) [48], Wind Driven Optimization (WDO) [27], Genetic Algorithms (GA) [16], Particle Swarm Optimization (PSO) [37], Cuckoo Search (CS) optimization algorithm [33, 34], Gray Wolf Optimization (GWO) [22], and so on. These approaches consider the band selection problem as a combinatorial optimization problem which is solved by formulating an appropriate fitness function or objective function. The objective function assesses the band subsets and provides the degree of their goodness. Objective function needs careful determination as it has an influence on the performance of the system. It can

be dependant or independent on the on learning algorithm. Hence, the objective function can be modelled by dependant evaluation criteria (depends on a predefined learning algorithm) [20, 39] or independent evaluation criteria such as such as, information measures (divergence, entropy or mutual information) [7, 42], distance measures (Bhattacharya distance, Kullback–Leibler divergence, Jeffries–Matusita distance, Hausdorff distance, and Spectral Angle Mapper (SAM)) [13, 22], and dependency measures (correlation measures, similarity measures) [6, 53]. Selection of an effective search strategy is a crucial task in the band selection problem. To optimize the defined objective function, an appropriate optimization algorithm must be chosen to converge to the global optimal solution and escape out of the local optimum.

Tschannerl et al. [42] proposed a modified discrete gravitational search algorithm based a novel unsupervised band selection approach employing maximum information minimum redundancy (MIMR) criterion which outperformed original MIMR with clonal selection algorithm. ABC approach is utilized to address problem of band selection [48]. An improved subspace decomposition is suggested to obtain intuitive band subspaces and ABC is employed as a search strategy. Xu et al. proposed a hybrid band selection approach by combining differential evolution (DE) and PSO in order to effectively combine characteristics of PSO and DE [49]. Sawant et al. [27] used a novel metaheuristic algorithm known as WDO, which is inspired atmospheric motion. WDO has excellent exploitation ability, but loses population diversity as gravitational force used in WDO pulls particles back to the origin. To address this limitation, PSO is combined with WDO and a novel band selection approach based on WDO and PSO is presented in [27]. Medjahed [21] proposed a binary version of CS algorithm to ensure that the selected band subset had a significant classification capability. Medjahed et al. [22] utilized a highly efficient metaheuristic algorithm called as GWO to reduce the dimensionality of hyperspectral image. This procedure has found a better band combination by optimizing five different objective functions. However, these algorithms also have their own limitations. A random population initialization strategy used in above algorithms may lead to the sluggish optimization response and fail to obtain the global optimum solution.

Recently, developing hybrid meta-heuristic algorithms has been gaining increasing attentions for researchers [10, 41, 49]. By combining two or more meta-heuristics, hybrid optimization approach can expand the performance of individual algorithm, and they seem to be more effective in solving complex optimization problems [1, 5, 9]. WDO and CS are population based meta-heuristic algorithms with different search mechanisms. WDO inspires from atmospheric motion and utilizes actual physical equations to search the solution space, whereas CS uses the Levy flight random walk to generate new solutions. The internal search mechanisms results in different characteristics of the two algorithms, i.e., WDO favours global exploration while CS is good at local exploitation. Considering the ability of these two algorithms, a hybrid approach based on WDO and modified CS, called WDOMCS, is proposed to speed up the convergence rate and enhance the optimization performance of the hybrid algorithm, aimed for selecting suitable bands with better classification performance. The main contributions of this manuscript are summarized as follows:

- 1) The proposed hybrid optimization algorithm utilizes the Chebyshev chaotic map for generation of initial population which helps circumventing the problem of premature convergence.
- 2) By population division, WDO and CS employ information sharing mechanism which decreases the risk of falling into a local optimal solution.
- 3) An adaptive step size is employed in Lévy flight random walk which helps enhancing the global searching ability of the standard CS algorithm.

- 4) A band selection approach for hyperspectral image classification based on WDOMCS is proposed, and the classification accuracy is the optimal with fewer bands.

The remaining part of the manuscript is organized as follows, a brief description about wind driven optimization, cuckoo search algorithm and chaotic map is provided in Section 2. Section 3 discusses the proposed WDOMCS algorithm in detail. Experimental analysis of the standard datasets is discussed in Section 4 and lastly, Section 5 presents the conclusions.

2 Technical background

In this section, the technical background of wind driven optimization, cuckoo search algorithm and chaotic map is discussed.

2.1 Wind driven optimization

WDO is population based nature inspired metaheuristic optimization algorithm proposed by Bayraktar et al. [2, 3]. It is easy to implement and highly effective in solving multi-modal and multidimensional problem. The inspiration for WDO comes from atmospheric motion where wind blows in order to balance the imbalances in air pressure. Air flows from high pressure point to low pressure point at a velocity which is proportional to pressure gradient. Based on the above theory, WDO algorithm is derived. It starts with Newton's second law of motion which states that the acceleration of an air particle as produced by a total force (F_t) is directly proportional to the magnitude of the total force, in the same direction as the total force according to Eq. (1).

$$\rho a = \sum \vec{F}_t \quad (1)$$

Where, ρ is the air density for an infinitesimal air particle and a is the acceleration of an air particle. The relationship of air pressure with the air particle's density and temperature is given by the ideal gas law,

$$P = \rho RT \quad (2)$$

Where, P is the pressure, R is the universal gas constant and T is the temperature.

There are various forces that are responsible for movement of air particles mainly including pressure gradient force (F_{PG}), friction force (F_F), gravitational force (F_G) and Coriolis force (F_C). The physical equations of the above mentioned forces are expressed as,

$$F_{PG} = -\nabla P \delta V \quad (3)$$

$$F_F = -\rho \alpha_f \mathbf{u} \quad (4)$$

$$F_G = \rho \delta V g \quad (5)$$

$$F_C = -2\Omega \times \mathbf{u} \quad (6)$$

Where, ∇P represents the pressure gradient, δV is finite volume of the air particle, \mathbf{u} represents the velocity vector of the wind, and α_f is the friction coefficient, g represents the gravitational acceleration, and Ω is rotation of the Earth.

By including all four forces $F_G, F_{PG}, F_C,$ and F_f in the right hand side of Eq. (1) and Eq. (1) can be expressed as,

$$\rho \frac{\Delta \mathbf{u}}{\Delta t} = \rho \delta V g + (-\nabla P \delta V) + (-2\Omega \times \mathbf{u}) + (-\rho \alpha_f \mathbf{u}) \tag{7}$$

The acceleration a in Eq. (1) is expressed as, $a = \frac{\Delta \mathbf{u}}{\Delta t}$. By assuming $\Delta t = 1$ for simplicity and for dimensionless and weightless air particle, $\delta V = 1$, the Eq. (7) can be expressed as,

$$\rho \Delta \mathbf{u} = \rho g + (-\nabla P) + (-2\Omega \times \mathbf{u}) + (-\rho \alpha_f \mathbf{u}) \tag{8}$$

By substituting Eq. (2) in Eq. (8),

$$\frac{P}{RT} \Delta \mathbf{u} = \frac{P}{RT} g + (-\nabla P) + (-2\Omega \times \mathbf{u}) + \left(-\frac{P}{RT} \alpha_f \mathbf{u}\right) \tag{9}$$

For simplicity, pressure P is replaced by pressure value of air particle at current location P_{curr} , and then Eq. (9) can be represented as,

$$\Delta \mathbf{u} = g + \left(-\nabla P \frac{RT}{P_{curr}}\right) + \left(-2\Omega \times \mathbf{u} \frac{RT}{P_{curr}}\right) + (-\alpha_f \mathbf{u}) \tag{10}$$

In WDO algorithm, velocity and position of air particle update their corresponding values at each iteration. Hence, change in the velocity $\Delta \mathbf{u}$ can be written as, $\Delta \mathbf{u} = \mathbf{u}_{new} - \mathbf{u}_{curr}$, where, \mathbf{u}_{curr} is the velocity of air particle at current iteration and \mathbf{u}_{new} is the velocity of air particle at next iteration. Then the velocity update equation will be,

$$\mathbf{u}_{new} = ((1-\alpha_f)\mathbf{u}_{curr}) - g + \left(-\nabla P \frac{RT}{P_{curr}}\right) + \left(-2\Omega \times \mathbf{u} \frac{RT}{P_{curr}}\right) \tag{11}$$

Consider, x_{curr} is current location of air particle, x_{new} is new location of air particle, $g \rightarrow$ and ∇P are vectors, which can be broken down in direction and magnitude as $g = |g|(0 - x_{curr}), -\nabla P = |P_{opt} - P_{curr}|(x_{new} - x_{curr})$, P_{opt} is optimum pressure, P_{curr} is the pressure of current location. The influence of the Coriolis force ($\Omega \times \mathbf{u}$) is replaced by the velocity influence from another dimension, $\mathbf{u}_{curr}^{other \ dim}$. Then, all the coefficients are combined together, $c = -2RT$ and the velocity update equation in Eq. (11) will be expressed as,

$$\mathbf{u}_{new} = ((1-\alpha_f)\mathbf{u}_{curr}) - g(x_{curr}) + \left[RT \left|1 - \frac{1}{i}\right| (x_{new} - x_{curr})\right] + \left[\frac{-c \mathbf{u}_{curr}^{other \ dim}}{i}\right] \tag{12}$$

Then the air particle location update equation will be,

$$x_{new} = x_{curr} + (\mathbf{u}_{new} \times \Delta t) \tag{13}$$

Where, i is rank among all air parcels based on their pressure values and x_{new} is the new updated location for next iteration. WDO is similar to other nature inspired optimization algorithms with few control variables that need to be adjusted. When standard WDO is applied to solve the hyperspectral band selection problem, the description of WDO parameters used in the algorithm is given in Table 1 and the hyperspectral band selection using standard WDO is given as Algorithm 1.

Algorithm 1: Hyperspectral band selection optimized by WDO

Input: Hyperspectral data χ

Output: The set of n selected bands (Global best solution)

Step 1. Initialization of parameters

WDO parameters: P_s (Population size), T (number of iterations), RT (RT coefficient), α_f (The friction coefficient), $\max V$ (Maximum allowed speed), g (Gravitational constant), c (Constant in the update equation)

Step 2. Initialize populations

Generate random initial population X_i (initial set of bands) with P_s air particles,

Assign random position x_{curr} and velocity u_{curr} to each particle

Evaluate the objective function f_i

Step 3. Identify the best particle P_{gbest} along with its position among all particles in initial population.

Step 4. Rank the air particles

Step 5. **while** ($t < T$)

Update the velocity u_{new} according to (12)

Update the position x_{new} by (13)

Evaluate the objective function f_i

Identify the best solution P_{lbest} among all particles

Rank the air particles

Update global best

If ($P_{\text{lbest}} > P_{\text{gbest}}$)

$$P_{\text{gbest}} = P_{\text{lbest}}$$

Store the corresponding position as best position

End if

End while

Step 6. Print the global best solution (set of n selected bands).

Table 1 Description of WDO parameters

Term	Description
Air particle	An individual member, whose coordinate values indicates solution to the optimization problem or one group of band combination
Population	An assembly of a predetermined number of diverse air particles
Position	The coordinates of an air particle
Velocity	The amount of position displacement per iteration
Pressure	A number assigned to air particles, which establishes how well an air particle meets the desired design performance.
Ranking	Here it act as objective function or fitness function Shorting the air particle at each iteration based on their pressure values

2.2 Cuckoo search algorithm

CS is population based bio-inspired metaheuristic algorithm proposed by Yang and Deb [51] as the solution for the global optimization problems. The inspiration for CS comes from the breeding behavior of some cuckoo species. Cuckoo bird put sits eggs in the nests of other bird. If a cuckoo egg is discovered by host bird, then host bird either throws the foreign eggs away or abandons the nest and constructs a new nest. In the standard CS algorithm, each egg of the host bird in a nest represents a solution, and a cuckoo egg denotes a new solution. If a new solution is better than the one in the nest, the worse one will be replaced with new solution. Therefore, the development of the CS algorithm is based the following three main rules: (1) Each cuckoo places one egg in an arbitrarily selected nest. (2) The best nest with good quality eggs is transferred to the next generation. (3) The number of available host's nests is fixed. The host bird determines the cuckoo egg by a probability of $p_a \in [0, 1]$. If the host bird finds a cuckoo egg, the host bird can abandon the nest or throw the egg away, and construct a completely new nest.

In the standard CS algorithm, initially m number of random nest locations are generated representing potential solution, X_i for a cuckoo i . A new solution X_{i+1} is generated by adopting Lévy flight random walk as follows:

$$X_{i+1} = X_i + \alpha \oplus \text{levy}(\lambda) \quad (14)$$

Where, $\alpha > 0$ is a step size scaling factor; the product of \oplus means entry wise multiplication; $\text{levy}(\lambda)$ is a random walk with the random step length which is drawn from a Lévy distribution. The Levy flight step size coefficient α , which controls the scale of the flight by multiplication and is related to the size of the solution space of the objective function. When a Lévy step is generated using a random number generator, it is first multiplied by α before it is used to generate a new egg. The cuckoo laying eggs corresponds to a new solution set and this is analogous to generating new solutions for a cuckoo i . The $\text{levy}(\lambda)$ is calculated using Mantegna's algorithm as follows:

$$\text{levy}(\lambda) = \frac{\mu}{|\nu|^{1/\lambda}} \quad (15)$$

Where, λ is fixed parameter with value 3/2. Furthermore, μ and ν are drawn from a normal distribution with zero means and an associated variance, as follows:

$$\sigma_{\mu} = \left[\frac{\Gamma(1 + \lambda)\sin(\pi\lambda/2)}{\Gamma((1 + \lambda/2)\lambda 2^{(\lambda-1)/2})} \right]^{1/\lambda}, \sigma_{\nu} = 1 \quad (16)$$

Where, $\Gamma(\cdot)$ is the standard Gamma function.

As standard CS is applied to solve the hyperspectral band selection problem, each cuckoo represents a band of hyperspectral data and each nest represents a solution of the problem (set of selected bands). Discovery probability p_a indicates each band has probability of p_a to be successful. The hyperspectral band selection using standard CS is given as Algorithm 2:

Algorithm 2: Hyperspectral band selection optimized by CS

Input: Hyperspectral data χ , Max_iteration, number of nests, Discovery probability

Output: The set of n selected bands (Global best solution)

Step 1. Set the initial parameters: Max_iteration T , m number of nests, Discovery probability of p_a

Step 2. Generate random initial population X_i (initial set of bands).

Step 3. Choose the current best nest by evaluating the objective function f_i

Step 4. **While** ($t < T$)

Get a cuckoo egg arbitrarily via Levy flights

Evaluate the objective function f_j

Select a nest amongst m random nests

If ($f_i > f_j$)

Replace X_i with the new solution X_i'

End if

Abandon the worst nests according to probability p_a

Preserve the best solutions

Rank the solutions and determine the current best solution

End while

Step 5. Print the global best solution (set of n selected bands).

2.3 Chaotic map

Chaos is the behavior of nonlinear systems which exhibits randomness and ergodic behavior. Chaotic map is an exceptional solution to the dynamic behavior of the systems that are highly sensitive to the random initial seeds [35]. Most of metaheuristic optimization algorithms involve random initialization of seeds and are highly dependent on initial conditions. So, if the initial solutions are not well selected, they lead to local minimum and poor results. The randomness at initial step leads the algorithm to premature convergence and easily gets trapped into a local optimal solution. Yang proposed a chaotic map based global optimization algorithm which employed chaotic variables instead of random variables, and improved the efficiency of the algorithm [50]. Ergodicity and non-repetition of chaos enables the chaotic map in maintaining the diversity in the population and tackles the problem of premature convergence.

A chaotic map is a mathematical function that shows some kind of chaotic behavior over time characterized by,

$$X_{t+1} = f(X_t) \quad (17)$$

In the chaotic sequence given in equation (4), chaotic variable X_{t+1} at time $t + 1$ depends on chaotic variable X_t at time t . There are several chaotic maps [40], such as iterative chaotic map with infinite collapses (ICMIC), tent map, logistic map, Chebyshev map, gauss map. The Chebyshev chaotic map used in this study is based on preliminary trials. Chebyshev map is defined as:

$$X_{t+1} = \cos(a \cdot \cos^{-1}(X_t)) \quad (18)$$

Where, a is finite constant.

3 Proposed band selection approach

Consider the hyperspectral image dataset which is represented as, $\chi \in \mathbb{R}^{H \times W \times N}$, where, H and W are the height and width of hyperspectral image respectively and N is the total number of spectral bands or the feature dimensions. Assume that k classes in the hyperspectral data are denoted as, $\Omega = [\Omega_1, \Omega_2, \Omega_3, \dots, \Omega_k]$ and n be the selected number of bands denoted as, $X_i = [b_1, b_2, b_3, \dots, b_n]$. The following subsection discusses in detail the proposed WDOMCS band selection algorithm, and the objective function.

3.1 Objective function or fitness function

The goal of the proposed band selection approach is to explore a finite set of possible subset of bands that maximize the classification accuracy. In this study, the classification accuracy of the classifier is used as an objective function. Classification accuracy is defined as follows [17]:

$$\frac{E(i)}{\text{total number of pixels}} = \text{Accuracy} = \frac{\sum_{k=1}^{|H \times W|} \text{Evaluate}(p_k)}{\text{total number of pixels}} \quad (19)$$

Where, assess (p_i) is the function used to classify pixels p_k .

$$Evaluate(p_k) = \begin{cases} 1 & \text{if } classify(p_k) = \Omega \\ 0 & \text{otherwise} \end{cases} \quad (20)$$

Classify (p_k) is the function that gives the class of p_k . For the pixel p_k with true class, the function assess (p_k) = 1, and 0 otherwise.

3.2 The proposed WDOMCS band selection algorithm

Although the standard WDO and CS have their own features and benefits, they also have some limitations. CS has relatively strong local exploration capabilities, but algorithm suffers from premature convergence and gets trapped into local optimal solution. The WDO algorithm has strong exploitation ability, thus combining the searching capabilities of both algorithms seem to be good approach which will certainly lead to the development of better performance algorithms.

The standard WDO and CS start with random population initialization which may cause repeated calculations and can easily fall into local optimal solution. For the optimization problem, the final solution is very sensitive to the initial population. By using the chaotic map, the property of the non-repetition and ergodicity accelerate the search speed by exploring all search space efficiently. Furthermore, a scaling factor α of Lévy flight distribution used in Eq. (14) is fixed in a standard CS algorithm. This parameter has a great influence on the global optimal solution. A large value of α allows cuckoo to explore search space rigorously at the risk of slower convergence. A smaller value of scaling factor causes an algorithm to get stuck into a local optimum solution, thereby leading to premature convergence. Therefore, the selection of the appropriate step size is indispensable to avoid local optimal solution and to enhance the speed of convergence. The scaling factor α_{new} can be formulated as,

$$\alpha_{new} = \begin{cases} \alpha_L + \frac{(\alpha_U - \alpha_L)(f_i(t) - f_w(t))}{f_{avg}(t) - f_w(t)}, & f_i(t) \leq f_w(t) \\ \frac{\alpha_U}{t}, & f_i(t) > f_w(t) \end{cases} \quad (21)$$

Where,

t is the iteration number.

α_L and α_U are predefined lower and upper limits of step size, respectively.

$f_{avg}(t)$ is the average fitness value of all nests.

$f_i(t)$ is the fitness value of i^{th} nest.

$f_w(t)$ is the worst fitness value.

The new nest location is computed from Lévy flight and Eq. (1) is re-written as,

$$X_i(t+1) = X_i(t) + \alpha_i(t+1) \oplus Levy(\lambda) \quad (22)$$

It is evident from this adaptation strategy, the moment current solution reaches a near optimum solution, the cuckoos upgrade their exploration around the current best solution using a smaller step size. On the other hand, a large step size is used when the current solution is less satisfactory than the average fitness value in order to inspire

the cuckoos to explore a search space more rigorously. The term t is involved to encourage more exploitation. Also, the position and fitness value of the new cuckoo egg remain unchanged if it goes out of the search space. In this manner, CS can make full use of the previous excellent host bird's nests. Therefore, inclusion of adaptive Lévy flight in the standard CS can help the algorithm quickly converge to the global optimal solution.

In this manuscript, a hybrid metaheuristic algorithm called WDOMCS is proposed aiming at combining both the exploration of CS and exploitation of WDO. In general, hybrid optimization algorithms can utilize the characteristics of the original algorithm. The main goal of hybridization is to complement each other's characteristics. The hybridization approach is as follows: In the beginning, the initial population is generated using Chebyshev chaotic map, and the whole population is divided into two equal subgroups, Group A and Group B. WDO and modified version of CS are adopted for these two subgroups independently. By division, these two subgroups can share suitable information and utilize each other's pros, thus avoid premature convergence, and obtain best optimal solution. The steps involved in the WDOMCS are presented in Algorithm 3. The flowchart of WDOMCS process is shown in Fig. 1.

Algorithm 3: Hyperspectral band selection optimized by WDOMCS

Input: Hyperspectral data χ

Output: The set of n selected bands (Global best solution)

Step 1. Initialization of parameters

P_s (Population size), T (number of iterations);

WDO parameters: RT (RT coefficient), α (The friction coefficient), $\max V$ (Maximum allowed speed), g (Gravitational constant), c (Constant in the update equation)

CS parameters: m number of nests, Discovery probability of p_a

Step 2. Initialize population

Generate initial population (initial set of bands) by the Chebyshev chaotic map using equation (18)

Step 3. Fitness evaluation

Calculate fitness values of all individuals according to the fitness or objective function $E(i)$ using equation (19) and find the global best.

Step 4. Population division

The whole population is equally divided into two subgroups (Group A and Group B). Individuals in each subgroup are randomly assigned by the population to utilize for the Steps 5 and 6.

Step 5. WDO process

$j = 1: Ps/2$ (all individuals in Group A)

Update the velocity and position of each individual (each air particle) using equation (12) and equation (13), respectively

Evaluate the $E_{new}(\mathbf{i})$ using equation (19)

Identify the best solution among all individuals and rank the individuals

Update global best

If ($E_{new}(i) > E(i)$)

 Replace current solution with the new solution

End if

Get the current global best solution

Step 6. MCS process

$j = 1: Ps/2$ (all individuals in Group B)

Generate a new solution via adaptive Levy flights using equation (22)

Evaluate the objective $E_{new}(\mathbf{i})$ using equation (19)

If ($E_{new}(i) > E(i)$)

 Replace current solution with the new solution

 Discard the worst nests according to probability p_a

End if

Get the current global best solution

Step 7. New solution update

After the two subgroups are updated, merge them into one population and store the global optimal value.

Step 7: Stop criteria

Determine whether the maximum number of iterations is reached. If not, return to step 2, otherwise Print the global best solution (set of n selected bands).

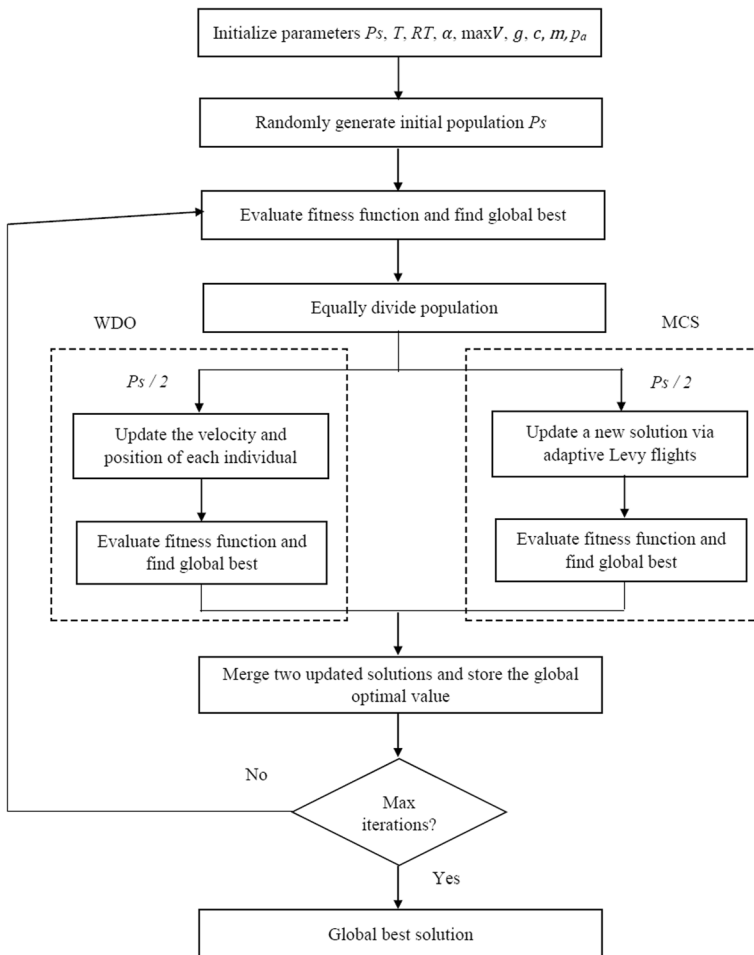


Fig. 1 Flowchart of WDOMCS process

4 Result and discussion

The proposed WDOMCS algorithm was tested on three different hyperspectral remote sensing datasets, Indian pines, Pavia University and Botswana [11] for assessment of the effectiveness of the proposed band selection approach. Details of the datasets and comprehensive results are discussed in following subsections.

4.1 Dataset description and experimental setup

Table 2 summarizes the details of the standard bench-mark hyperspectral datasets, Indian pines, Pavia University and Botswana.

All the experiments are conducted using MATLAB platform version R2018b on Intel Xeon processor 2.90 GHz CPU along with 128 GB RAM in Windows 10 (64 bit) environment. The performance of the proposed method is compared with those of the other five metaheuristic band selection approaches, i.e. GA, PSO, GWO, WDO, and CS.

Table 2 Details of Hyperspectral datasets

Dataset	Indian Pines	Pavia University	Botswana
The sensor used to capture the scene	AVIRIS	ROSIS	Hyperion
Type of the Sensor	Airborne	Airborne	Space borne
Size of the image	145 X 145	610 X 340	1476 X 256
Number of spectral bands	224	115	242
Number of bands existing when water absorption bands are removed	200	103	145
Wavelength range	400–2500 nm	430–860 nm	400–2500 nm
Spectral resolution	10 nm	3.7 nm	10 nm
Spatial resolution	20 m	1.3 m	30 m
Number of classes	16	9	14
Number of labelled samples	10,249	42,776	3248

In the beginning, 10% number of samples of each land cover class from the Indian pine dataset, Pavia University and Botswana datasets are randomly chosen as training samples for validation of the efficacy of the proposed technique with less amount of labelled data. The rest of the samples are considered as testing data. The experimentation is conducted ten times for the evaluation of the average of Overall Accuracy (OA), Average Accuracy (AA), class-wise accuracy, and kappa coefficient κ . To perform the classification task using selected bands, one-against-all Support Vector Machine (SVM) is employed as classifier [23]. The SVM classifier is selected due to its excellent performance when dealing with small-sized training data. To compare fairly, we have conducted the experiments on the final bands searched by their methods, but using the same classifier of SVM. In SVM, Radial Basis Function (RBF) is used as its kernel function. The parameters C and γ in SVM are determined using fivefold cross validation method. The parameters of each algorithm are summarized in Table 3.

Table 3 Parameters setting of each algorithm

Algorithm	Parameters	Value
GA	Number of genetics	20
	Selection ratio	0.8
	Crossover ratio	0.9
	Mutation ratio	0.01
PSO	Population size	20
	Number of iterations	100
	Cognitive constant ($c1$)	2
	Social constant ($c2$)	2
	Inertia weight	0.2
GWO	Number of search agents	20
	Number of iterations	100
CS	Population size	20
	Number of iterations	100
WDO	Discovery probability p_a	0.25
	Population size	20
	Maximum number of iterations	100
	RT coefficient	3
	gravitational constant	0.15
	constants in the update equation	0.4
	Coriolis effect	0.3
maximum allowed speed	0.3	

Table 4 Classification accuracies (%) obtained by the proposed method and competing approaches for Indian pines dataset

Class Name	GA	PSO	GWO	WDO	CS	WDOMCS
Alfalfa	0.00	63.41	26.19	0	39.02	39.02
Corn-no till	52.93	63.34	58.39	76.88	73.38	80.15
Corn-min till	28.16	53.54	59.57	59.70	65.19	67.73
Corn	14.74	31.92	40.65	32.86	44.13	55.39
Grass-pasture	82.17	82.71	78.62	88.94	90.55	80.41
Grass-tree	94.35	92.84	92.99	94.52	94.06	95.73
Grass-pasture- mowed	17.39	0	76.92	0	0	20
Hay-windrowed	96.87	97.67	96.75	96.27	98.13	98.83
Oat	0.00	0	50.00	0	0	0
Soybean-no till	59.38	37.41	70.05	76.88	67.27	74.25
Soybean-min till	74.13	83.70	71.67	84.51	86.69	86.28
Soybean-clean	32.00	62.10	50.93	67.35	73.17	71.29
Wheat	94.51	94.56	94.59	98.91	97.82	96.73
Woods	92.98	97.10	89.11	96.30	93.32	94.72
Buildings-grass-trees- drives	11.65	41.49	45.40	49.56	52.16	53.89
Stone-steel- towers	73.33	77.10	83.33	93.97	81.92	86.74
OA	64.86	72.59	71.28	79.75	79.81	81.67
AA	51.54	61.19	67.82	63.55	66.05	68.83
κ	58.96	68.23	66.53	76.72	76.79	78.96

Bold value shows the better performance of the proposed method compare to the competing methods in terms of OA, AA, and κ

4.2 Experimental analysis

This section discusses the classification results obtained by the proposed approach and commonly used metaheuristic approaches, comparison of the proposed band selection approach with commonly used band selection approaches, WDOMCS parameter analysis and convergence analysis.

4.2.1 Classification performance

The classification results obtained by the proposed band selection approach WDOMCS are compared with other metaheuristic approaches to assess the effectiveness of the WDOMCS. GA,

Table 5 Classification accuracies (%) obtained by the proposed method and competing approaches for Pavia University dataset

Class Name	GA	PSO	GWO	WDO	CS	WDOMCS
Asphalt	82.28	90.84	87.80	94.01	94.97	93.24
Meadows	96.91	96.49	95.55	97.35	97.77	98.12
Gravel	43.75	60.61	72.96	82.00	76.33	74.74
Trees	79.73	92.05	83.50	90.71	91.69	93.21
Painted metal sheets	99.16	98.34	99.25	98.51	99.58	98.76
Bare soil	58.25	80.97	72.14	89.63	88.75	90.96
Bitumen	75.47	74.93	84.29	81.11	82.53	84.37
Self-blocking bricks	83.47	85.27	82.64	87.08	90.40	92.42
Shadows	100	99.88	99.88	99.17	99.88	100
OA	84.57	90.21	86.83	93.38	93.79	94.17
AA	79.89	86.60	86.45	91.07	91.33	91.76
κ	78.56	86.96	82.29	91.20	91.73	92.25

Bold value shows the better performance of the proposed method compare to the competing methods in terms of OA, AA, and κ

PSO, GWO, WDO, CS, and WDOMCS are executed by selecting 20 bands on the Indian Pines, Pavia University and Botswana dataset. The OA, AA and κ are measured in every case alongside the individual class accuracies. Results are presented in Tables 4, 5 and 6 and the best value for each metric is recorded in bold. It is observed from Table 4 that, GA performs worst on Indian Pines dataset, whereas PSO and GWO show similar performances. It is observed from Table 5 that, GA and GWO perform worst on Pavia University dataset; however, PSO shows satisfactory performance. On the other hand, Table 6 shows that, performance of GA and PSO is inferior to that of GWO. Moreover, the performance of GWO is almost similar to that of CS.

According to the results in Tables 4, 5 and 6, the optimization abilities of WDO and CS are better than GA, PSO and GWO and the OA value exceeded 80% for Indian Pines dataset, 94% for Pavia University dataset and 92% for Botswana dataset. WDO and CS require a few control parameters; however, the exploration ability of both algorithms is restricted to the initial setting. In WDO, air particles simultaneously walk around the universe, and explore the entire search space without getting trapped at the boundary. In MCS, adaptive Lévy flight strategy helps to maintain balance between exploration and exploitation and enhances the searching ability to obtain global best solution. Although, the WDO and CS give better results, they have not obtained theoretical optimal value. Thus, WDOMCS obtains the highest fitness value for all datasets, which illustrates that the excellent optimization ability of WDOMCS in comparison with other metaheuristic approaches. In contrast to this excellent performance, the classification performance obtained by the standard WDO and CS algorithm with the random initialization is inferior to that of proposed WDOMCS approach. This substantial difference indicates that, the classification performance obtained by the standard WDO and CS algorithm is deteriorated due to the randomness introduced in the selection of initial population. It is evident from Tables 4, 5 and 6, the proposed WDOMCS algorithm has strong capability to obtain an optimal solutions. Among all metaheuristic approaches, the proposed WDOMCS algorithm has achieved excellent classification performance in terms of OA, AA and κ . The proposed WDOMCS algorithm

Table 6 Classification accuracies (%) obtained by the proposed method and competing approaches for Botswana dataset

Class Name	GA	PSO	GWO	WDO	CS	WDOMCS
Water	99.58	100	100	100	100	100
Hippo grass	92.22	84.44	92.22	83.33	82.22	88.75
Floodplain grasses1	91.55	95.11	96.44	96.88	99.55	94.50
Floodplain grasses2	90.15	75.64	96.37	95.85	97.92	99.41
Reeds1	76.44	77.27	83.05	74.38	67.35	88.83
Riparian	68.59	62.80	85.53	72.72	85.12	82.32
Firescar2	93.56	96.13	93.99	96.56	96.13	98.06
Island interior	92.30	84.61	95.60	100	97.80	94.44
Accacia woodlands	87.94	75.53	79.07	94.68	92.19	89.24
Accaciashrublands	80.26	79.37	88.78	89.23	91.92	84.84
Accacia grasslands	94.52	94.16	87.22	94.16	87.59	93.03
Short mopane	96.29	95.06	93.20	87.65	95.06	91.66
Mixed mopane	87.55	90.45	84.64	91.28	85.89	95.79
Exposed soils	91.76	90.58	96.47	100	92.94	100
OA	88.20	85.47	90.06	91.02	90.70	92.64
AA	88.77	85.80	90.90	91.20	90.84	92.92
κ	87.22	84.24	89.23	90.26	89.93	92.02

Bold value shows the better performance of the proposed method compare to the competing methods in terms of OA, AA, and κ

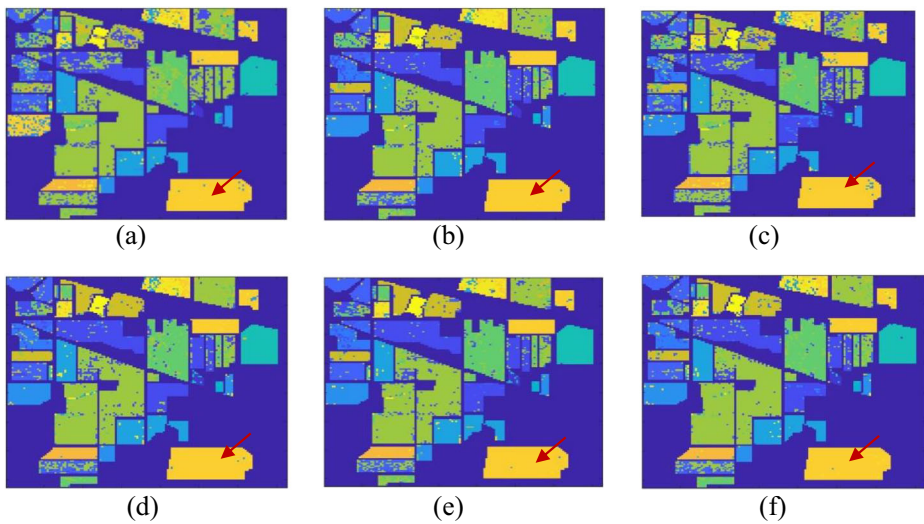


Fig. 2 Classification map obtained by various methods for Indian Pines dataset. **a** GA **(b)** PSO **(c)** GWO **(d)** WDO **(e)** CS **(f)** WDOMCS

has achieved the maximum OA of 81.67% for Indian Pines dataset, 94.17% for Pavia University dataset, and 92.64% for Botswana dataset. It can be inferred from Tables 4, 5 and 6 that, the use of the initialization strategy, adaptive Lévy flight strategy and hybridization of WDO and CS have greater influence on the optimization ability of an algorithm.

The classification maps produced by all the above mentioned band selection approaches on Indian Pines, Pavia University and Botswana dataset are shown in Figs. 2, 3 and 4, respectively. The proposed approach WDOMCS shows a smooth classification map than other competing band selection approaches. For instance, the region of class “Woods” as shown

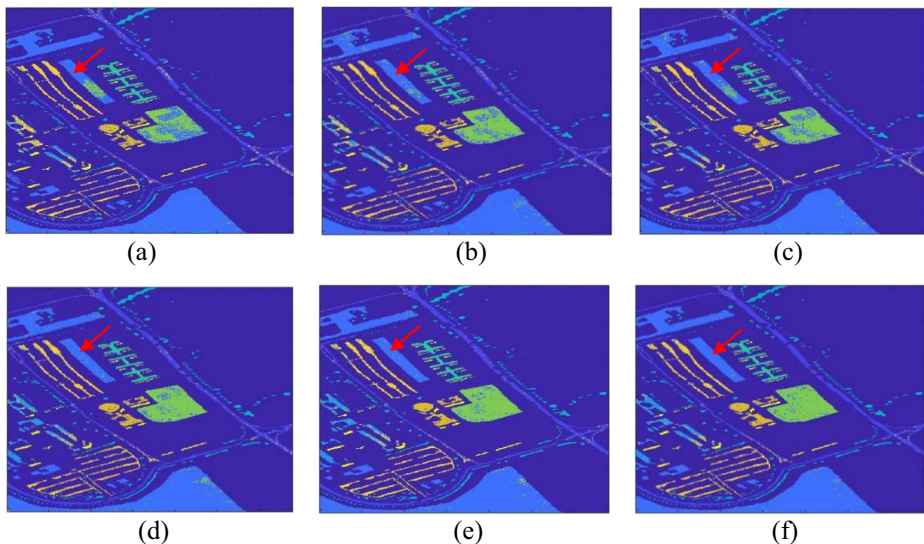


Fig. 3 Classification map obtained by various methods for Pavia University dataset. **a** GA **(b)** PSO **(c)** GWO **(d)** WDO **(e)** CS **(f)** WDOMCS

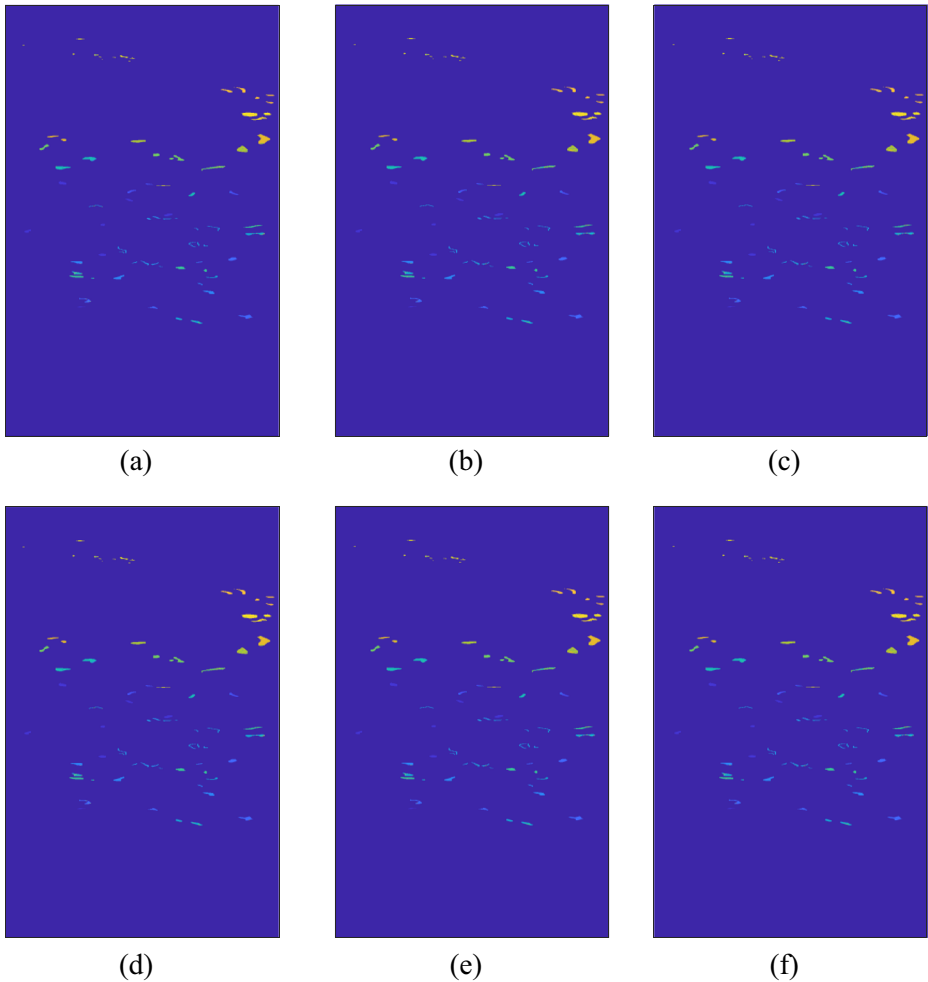
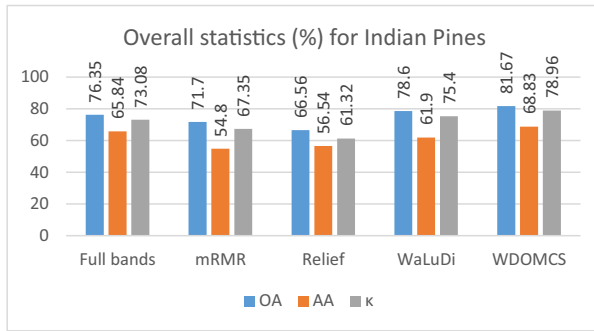


Fig. 4 Classification map obtained by various methods for Botswana dataset. **a** GA **(b)** PSO **(c)** GWO **(d)** WDO **(e)** CS **(f)** WDOMCS

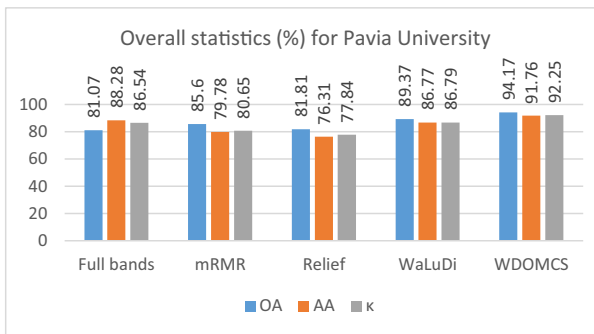
in Fig. 2f (denoted by brown arrow) is smoother than those of the other approaches. Similarly, the region of class “Meadows” as shown in Fig. 3f (denoted by red arrow) is smoother than other approaches.

4.2.2 Comparison of the proposed band selection approach with commonly used band selection approaches

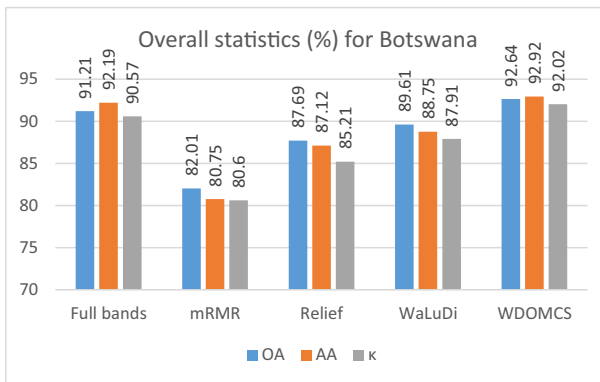
To validate the effectiveness of the proposed WDOMCS approach, the classification performance of the WDOMCS is compared with commonly used band selection approaches such as Ward’s linkage strategy using divergence (WaLuDi) [18], minimum-redundancy maximum-relevance (mRMR) [19], and Relief algorithm [12]. Figure 5a, b and c present the OA, AA and κ for three hyperspectral datasets using above mentioned band selection approaches. It is observed that the classification accuracy by the proposed WDOMCS shows significant



(a)



(b)



(c)

Fig. 5 Overall statistics (OA, AA, κ in %) for (a) Indian Pines dataset (b) Pavia University dataset (c) Botswana dataset

improvement that by using the full spectral bands, which is higher than 4% for Indian Pines, 13% for Pavia University and 1% for Botswana dataset. More importantly, the number selected of bands is greatly reduced in comparison with original spectral bands. Among all the band selection approaches, WDOMCS attains excellent classification performance by selecting most discriminative bands. Although the number of selected bands is the same, OA is lower than 80% for Indian Pines, less than 90% for Pavia University and Botswana

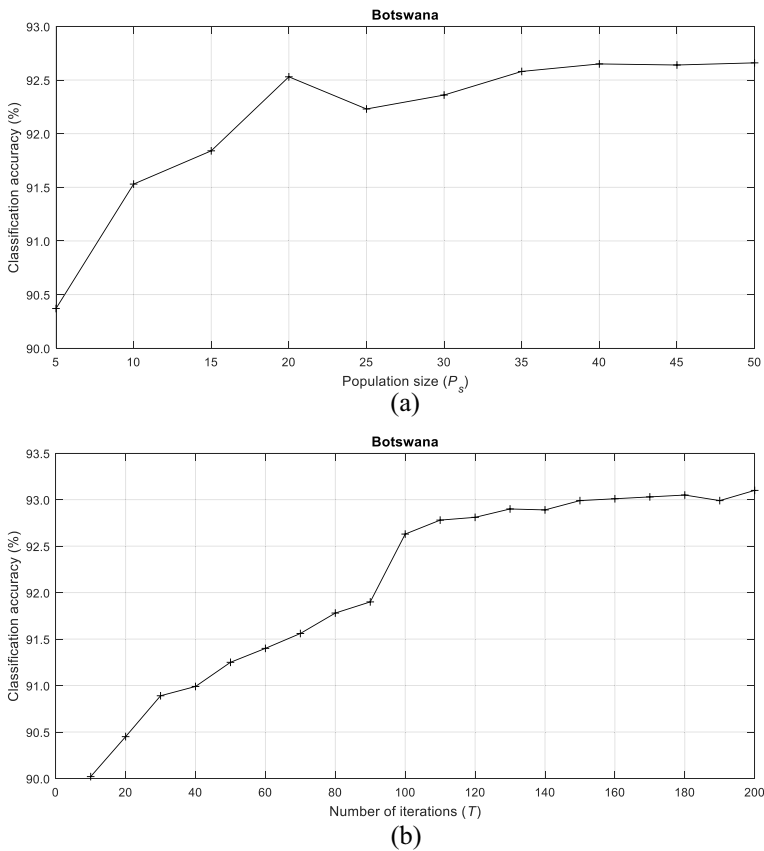


Fig. 6 Sensitivity of WDOMCS parameters (a) population size (b) number of iterations

datasets when using Relief, mRMR, and WaLuDi approaches. Similarly, AA is less than 62% for Indian Pines, lower than 89% for Pavia University and Botswana datasets when using Relief, mRMR, and WaLuDi approaches. Furthermore, kappa coefficient is lower than 76% for Indian Pines, less than 88% for Pavia University and Botswana datasets when using Relief, mRMR, and WaLuDi approaches. In short, we can say that WDOMCS has the excellent searching ability and optimization performance, which makes it better choice to solve the band selection problem.

4.2.3 WDOMCS parameter analysis

The population size P_s , number of iterations T are the two main parameters influencing the performance of WDOMCS. Both parameters have been analysed in Fig. 6a and b for a fixed number of 20 bands on the Botswana dataset. As expected, classification accuracy increases with increase in P_s and T . They are chosen to be as minimal as possible to attain maximum optimisation capacity with minimal computational effort. Based on these results, P_s is set to 20 and T is set to 100.

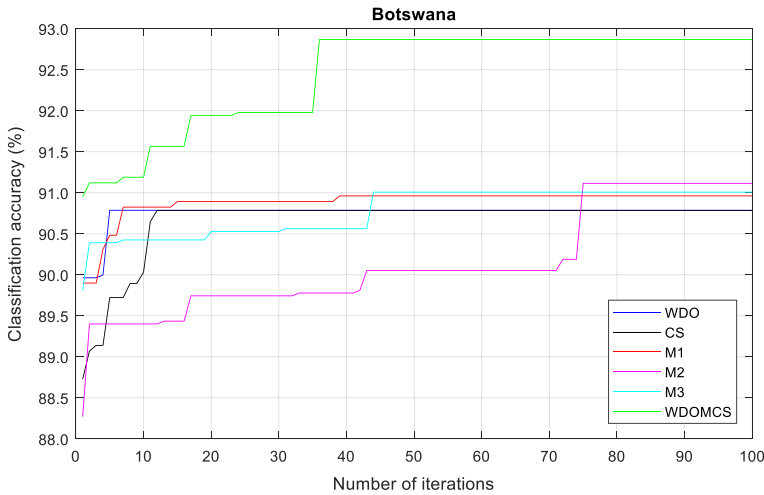


Fig. 7 Convergence behaviour of WDOMCS algorithm on Botswana dataset

4.2.4 Convergence analysis

In this section, optimization performance of the WDOMCS is discussed with respect to convergence analysis. The proposed WDOMCS includes hybridization of the standard WDO and CS algorithm. In addition, Chebyshev chaotic map based population initialization phase and one modification to the standard CS algorithm were adopted to speed up the convergence process. The ergodicity and stochastic properties exhibited by chaos helps the algorithm to reach its global best optimal solution. A modification to the standard CS algorithm includes the use of an adaptive Lévy flight strategy based on the fitness value of previous iteration, which will help in increasing the convergence rate by exploring around the most likely solution space. In this section, the influence of each modification on the standard algorithms (M1: WDO algorithm with Chebyshev chaotic initialization, M2: CS algorithm with Chebyshev chaotic initialization, M3: CS algorithm with adaptive Lévy flight strategy) are studied. The convergence plots using the above three modifications and the standard WDO, CS and WDOMCS algorithm for Botswana dataset are shown in Fig. 7. It is evident from Fig. 7, the use of Chebyshev chaotic map based population initialization have reached a near attainable global solution without bringing any substantial enhancement to the convergence rate in comparison with the standard WDO. Similarly, CS algorithm Chebyshev chaotic map based population initialization attained maximum fitness value without getting trapped in the local optimum. Whereas, the convergence rate of M2 is slower than that of standard CS algorithm. The Chebyshev chaotic map based population initialization strategy can improve the searching ability because it generate non-repeated solution and improve the quality of the initial feasible solution. The use of adaptive Lévy flights strategy in M3 provided more stable convergence rate. The adaptive step size encourages the cuckoos to explore more rigorously to find a new solution. However, in comparison with modification M2, it failed to attain the global optimal solution. Therefore, the inclusion of all three modifications to the standard WDO and CS algorithm lead to a more stable and reliable optimization performance as expected.

All the approaches except WDOMCS have obtained almost same solution at the end iteration; however convergence speed slightly varies. As can be clearly seen from Fig. 7, WDOMCS algorithm converges quickly and reaches global optimal solution. The Chebyshev chaotic map generates non-random population and significantly increases the convergence rate. Adaptive step size helps

algorithm to come out of local optima, overcomes premature convergence and improves quality of solution significantly. Furthermore, hybridization of the WDO and CS leads to exchange of useful information across the search space. Therefore, with these modifications, WDOMCS obtains optimal solution quickly in comparison with other metaheuristic approaches. The fitness curve of WDOMCS proved that, the exploitation capability of the algorithm has been enhanced which succeed in reducing premature convergence problem as well as improving the convergence rate.

5 Conclusion

In this paper, a novel hybrid optimization approach is proposed based on WDO and MCS algorithm for solving the problem of band selection in hyperspectral image. In WDOMCS, the randomness of initial population is modified through use of the chaotic map. Furthermore, the searching ability of CS algorithm is improved by adaptively adjusting the Levy flight step size. WDOMCS can enhance the ability of exploitation and avoid getting trapped into local optimal solution by dividing the population, grouping update and combining the population to share the useful information between individuals. The experimental results have been compared with other band selection techniques optimized by GA, PSO, GWO, WDO and CS. Compared to GA, PSO, GWO, WDO and CS algorithm, the WDOMCS is more effective in obtaining better solutions and convergence speed of WDOMCS is higher. This hybridization presented a better optimization performance and robustness and significantly enhances the optimization capability of original WDO and CS algorithms.

Although in this paper the hybrid WDOMCS algorithm was implemented for solving the problem of hyperspectral band selection, there are still many aspects worthy of our study. Firstly, we will focus on various initialization strategies in order to further improve the performance of optimization. Secondly, multi-objective optimization model is needed to find an optimal trade-offs between conflicting issues such as, information preserving and redundancy reducing for choosing optimal band subsets. Also, we will further explore various strategies for combining spectral and spatial information to improve the classification performance. Lastly, we will further explore our proposed approach to various multimedia applications such as, image segmentation [52], image detection [26], and image dehaze [54].

Acknowledgments The authors thank the Council of Scientific & Industrial Research (CSIR), New Delhi, India for the award of CSIR-SRF and VIT for providing a VIT seed grant for carrying out this research work.

References

1. Abed-algumi BH, Alkhateeb F (2018) Intelligent hybrid cuckoo search and b-hill climbing algorithm. *J King Saud Univ - Comput Inf Sci*. <https://doi.org/10.1016/j.jksuci.2018.05.003>
2. Bayraktar Z, Komurcu M, Werner DH (2010) Wind Driven Optimization (WDO): a novel nature-inspired optimization algorithm and its application to electromagnetics, 2010 IEEE Antennas Propag. Soc Int Symp: 1–4. <https://doi.org/10.1109/APS.2010.5562213>
3. Bayraktar Z, Komurcu M, Bossard JA, Werner DH (2013) The wind driven optimization technique and its application in electromagnetics. *IEEE Trans Antennas Propag* 61:2745–2757. <https://doi.org/10.1109/TAP.2013.2238654>
4. Boggavarapu LNP, Manoharan P (2020) Classification of hyper spectral remote sensing imagery using intrinsic parameter estimation. Springer International Publishing, Berlin. <https://doi.org/10.1007/978-3-030-16660-1>
5. Bouyer A, Hatamlou A (2018) An efficient hybrid clustering method based on improved cuckoo optimization and modified particle swarm optimization algorithms. *Appl Soft Comput J* 67:172–182. <https://doi.org/10.1016/j.asoc.2018.03.011>

6. Chang CI, Wang S (2006) Constrained band selection for hyperspectral imagery. *IEEE Trans Geosci Remote Sens* 44:1575–1585. <https://doi.org/10.1109/TGRS.2006.864389>
7. Feng J, Jiao LC, Zhang X, Sun T (2014) Hyperspectral band selection based on trivariate mutual information and clonal selection. *IEEE Trans Geosci Remote Sens* 52:4092–4115. <https://doi.org/10.1109/TGRS.2013.2279591>
8. Feng S, Itoh Y, Parente M, Duarte MF (2017) Hyperspectral band selection from statistical wavelet models. *IEEE Trans Geosci Remote Sens* 55:2111–2123. <https://doi.org/10.1109/TGRS.2016.2636850>
9. Garg H (2019) A hybrid GSA-GA algorithm for constrained optimization problems a hybrid GSA-GA algorithm for constrained optimization problems. *Inf Sci (Ny)* 478:499–523. <https://doi.org/10.1016/j.ins.2018.11.041>
10. Ghamisi P, Benediktsson JA (2015) Feature selection based on hybridization of genetic algorithm and particle swarm optimization. *IEEE Geosci Remote Sens Lett* 12:309–313. <https://doi.org/10.1109/LGRS.2014.2337320>
11. http://www.ehu.es/ccwintco/index.php/Hyperspectral_Remote_Sensing_Scenes (n.d.)
12. Jia J, Yang N, Zhang C, Yue A, Yang J, Zhu D (2013) Object-oriented feature selection of high spatial resolution images using an improved relief algorithm. *Math Comput Model* 58:619–626. <https://doi.org/10.1016/j.mcm.2011.10.045>
13. Kailath T (1967) The divergence and Bhattacharyya distance measures in signal selection. *IEEE Trans Commun Technol* 15:52–60. <https://doi.org/10.1109/TCOM.1967.1089532>
14. Kumar BLNP (2020) Hyperspectral image classification using fuzzy-embedded hyperbolic sigmoid non-linear principal component and weighted least squares approach. *J Appl Remote Sens* 14(2):024501. <https://doi.org/10.1117/1.JRS.14.024501>
15. Kumar BLNP, Manoharan P (2020) A new framework for hyperspectral image classification using Gabor embedded patch-based convolution neural network. *Infrared Phys Technol* 110:103455. <https://doi.org/10.1016/j.infrared.2020.103455>
16. Li S, Wu H, Wan D, Zhu J (2011) Knowledge-based systems an effective feature selection method for hyperspectral image classification based on genetic algorithm and support vector machine. *Knowledge-Based Syst* 24:40–48. <https://doi.org/10.1016/j.knosys.2010.07.003>
17. Liu Y, Wang G, Chen H, Dong H, Zhu X, Wang S (2011) An improved particle swarm optimization for feature selection. *J Bionic Eng* 8:191–200. [https://doi.org/10.1016/S1672-6529\(11\)60020-6](https://doi.org/10.1016/S1672-6529(11)60020-6)
18. Martínez-UsóMartínez-Usó A, Pla F, Sotoca JM, García-Sevilla P (2007) Clustering-based Hyperspectral band selection using information measures. *IEEE Trans Geosci Remote Sens* 45:4158–4171. <https://doi.org/10.1109/TGRS.2007.904951>
19. Max-dependency C (2005) *Feat Select Based Mutual Info* 27:1226–1238
20. Medjahed SA, Ouali M (2018) Band selection based on optimization approach for hyperspectral image classification. *Egypt J Remote Sens Sp Sci*. <https://doi.org/10.1016/j.ejrs.2018.01.003>
21. Medjahed SA, Saadi TA, Benyettou A, Ouali M (2015) Binary cuckoo search algorithm for band selection in hyperspectral image classification. *IAENG Int J Comput Sci* 42:1–9
22. Medjahed SA, Ait Saadi T, Benyettou A, Ouali M (2016) Gray wolf optimizer for hyperspectral band selection. *Appl Soft Comput J* 40:178–186. <https://doi.org/10.1016/j.asoc.2015.09.045>
23. Melgani F, Bruzzone L (2004) *Classification Hyperspect Remote Sens* 42:1778–1790
24. Manoharan Prabukumar, Shrutika Sawant, Sathishkumar Samiappan, Loganathan Agilandeeswari(2018) Three-dimensional discrete cosine transform-based feature extraction for hyperspectral image classification, *J. Appl. Remote Sens.* 12(4):046010 (2018). <https://doi.org/10.1117/1.JRS.12.046010>
25. Prabukumar M, Sawant SS (2018) Band clustering using expectation – maximization algorithm and weighted average fusion-based feature extraction for hyperspectral image classification. *J Appl Remote Sens* 12. <https://doi.org/10.1117/1.JRS.12>
26. Quan Q, He F, Li H (2020) A multi-phase blending method with incremental intensity for training detection networks. *Vis Comput*. <https://doi.org/10.1007/s00371-020-01796-7>
27. Sawant SS, Manoharan P (2019) New framework for hyperspectral band selection using modified wind-driven optimization algorithm. *Int J Remote Sens* 00:1–22. <https://doi.org/10.1080/01431161.2019.1607609>
28. Sawant SS, Manoharan P (2020) Unsupervised band selection based on weighted information entropy and 3D discrete cosine transform for hyperspectral image classification. *Int J Remote Sens* 41:3948–3969. <https://doi.org/10.1080/01431161.2019.1711242>
29. Sawant SS, Prabukumar M (2017) Semi-supervised techniques based hyper-spectral image classification: a survey, in: 2017 *Innovations in Power and Advanced Computing Technologies (i-PACT)*, Vellore, 2017, pp 1–8. <https://doi.org/10.1109/IPACT.2017.8244999>
30. Sawant Shrutika, Manoharan Prabukumar (2020) A Review on Graph-Based Semi-Supervised Learning Methods for Hyperspectral Image Classification, *The Egyptian Journal of Remote Sensing and Space Sciences*, 23: 243–248. <https://doi.org/10.1016/j.ejrs.2018.11.001>

31. Sawant S, Prabukumar M (2020) A survey of band selection techniques for hyperspectral image classification. *J Spectr Imaging* 1:1–18. <https://doi.org/10.1255/jsi.2020.a5>
32. Sawant S, Manoharan P, Samiappan S (2018) Ranking and Grouping based Feature Selection for Hyperspectral Image Classification, *Proceedings Asian Conference on Remote Sensing*, pp. 2305–2313
33. Sawant S, Manoharan P, Samiappan S (2019) A band selection method for hyperspectral image classification based on cuckoo search algorithm with correlation based initialization, in: 10th workshop on Hyperspectral imaging and signal processing: Evolution in remote sensing (WHISPERS), Amsterdam, Netherlands, pp. 1–4.
34. Sawant S, Prabukumar M, Samiappan S (2020) A modified cuckoo search algorithm based optimal band subset selection approach for hyperspectral image classification. *J Spectr Imaging* 1:1–20. <https://doi.org/10.1255/jsi.2020.a6>
35. Science N, Phenomena C, Walton S, Hassan O, Morgan K, Brown MR (2011) Chaos, Solitons & Fractals Modified cuckoo search: A new gradient free optimisation algorithm, *Chaos, Solitons Fractals Interdiscip. J Nonlinear Sci Nonequilibrium Complex Phenom* 44:710–718. <https://doi.org/10.1016/j.chaos.2011.06.004>
36. Senthil V, Ganesan K, Vasuki S (2018) Maximin distance based band selection for endmember extraction in hyperspectral images using simplex growing algorithm, 7221–7237. <https://doi.org/10.1007/s11042-017-4630-0>
37. Su H, Du Q, Chen G, Du P (2014) Optimized hyperspectral band selection using particle swarm optimization. *IEEE J Sel Top Appl Earth Obs Remote Sens* 7:2659–2670. <https://doi.org/10.1109/JSTARS.2014.2312539>
38. Su H, Cai Y, Du Q (2017) Firefly-algorithm-inspired framework with band selection and extreme learning machine for Hyperspectral image classification. *IEEE J Sel Top Appl Earth Obs Remote Sens* 10:309–320. <https://doi.org/10.1109/JSTARS.2016.2591004>
39. Sui C, Tian Y, Xu Y, Xie Y (2015) Unsupervised band selection by integrating the overall accuracy and redundancy. *IEEE Geosci Remote Sens Lett* 12:185–189. <https://doi.org/10.1109/LGRS.2014.2331674>
40. Tavazoei MS, Haeri M (2007) Comparison of different one-dimensional maps as chaotic search pattern in chaos optimization algorithms, 187: 1076–1085. <https://doi.org/10.1016/j.amc.2006.09.087>
41. Taylor P (2013) Hybrid genetic algorithm for feature selection with hyperspectral data 37–41. <https://doi.org/10.1080/2150704X.2013.777485>.
42. Tschannerl J, Ren J, Yuen P, Sun G, Zhao H, Yang Z, Wang Z, Marshall S (2019) MIMR-DGSA : unsupervised hyperspectral band selection based on information theory and a modified discrete gravitational search algorithm. *Inf Fusion* 51:189–200. <https://doi.org/10.1016/j.inffus.2019.02.005>
43. Vaddi R, Manoharan P (2020) Probabilistic PCA based hyperspectral image Classification for remote sensing applications. Springer International Publishing, Berlin. <https://doi.org/10.1007/978-3-030-16660-1>
44. Vaddi R, Manoharan P (2020) Hyperspectral image classification using CNN with spectral and spatial features integration. *Infrared Phys Technol* 107:103296. <https://doi.org/10.1016/j.infrared.2020.103296>
45. Vaddi R, Manoharan P (2020) CNN based hyperspectral image classification using unsupervised band selection and structure-preserving spatial features, *Infrared Phys Technol* 10:103457. <https://doi.org/10.1016/j.infrared.2020.103457>
46. Veera Senthil Kumar G, Vasuki S (2017) Clustering based band selection for endmember extraction using simplex growing algorithm in hyperspectral images, 8355–8371. <https://doi.org/10.1007/s11042-016-3420-4>
47. Wang Q, S. Member, Lin J, S. Member, Yuan Y, S. Member (2016) Salient Band Selection for Hyperspectral Image Classification via Manifold Ranking 27: 1279–1289.
48. Xie F, Li F, Lei C, Yang J, Zhang Y (2019) Unsupervised band selection based on artificial bee colony algorithm for hyperspectral image classification. *Appl Soft Comput J* 75:428–440. <https://doi.org/10.1016/j.asoc.2018.11.014>
49. Xu M, Sun Q, He Z, Shi J (2016) Band selection for hyperspectral images based on particle swarm optimization and differential evolution algorithms with hybrid encoding, 16: 629–640. <https://doi.org/10.3233/JCM-160645>
50. Yang D, Li G, Cheng G (2007) On the efficiency of chaos optimization algorithms for global optimization, 34: 1366–1375. <https://doi.org/10.1016/j.chaos.2006.04.057>
51. Yang X, Deb S, A.C.B. Behaviour (2009) Cuckoo Search via Lévy Flights 210–214
52. H. Yu, F. He, Pan Y (2019) A scalable region-based level set method using adaptive bilateral filter for noisy image segmentation
53. Zhang W, Li X, Zhao L (2018) A fast Hyperspectral feature selection method based on band correlation analysis. *IEEE Geosci Remote Sens Lett* 15:1–5. <https://doi.org/10.1109/LGRS.2018.2853805>
54. Zhang J, He F, Chen Y (2020) A new haze removal approach for sky / river alike scenes based on external and internal clues, 2085–2107