Applied Soft Computing 12 (2012) 832–849

Contents lists available at SciVerse ScienceDirect

Applied Soft Computing

journal homepage: www.elsevier.com/locate/asoc

Hybrid bio-inspired techniques for land cover feature extraction: A remote sensing perspective

Lavika Goel[a,*], Daya Gupta[a], V.K. Panchal[b]

[a] Department of Computer Engineering, Delhi Technological University, New Delhi 110042, India
[b] Defense Terrain & Research Lab, DRDO, MetCalfe House, New Delhi 110054, India

ARTICLE INFO

Article history:
Received 8 May 2010
Received in revised form 2 May 2011
Accepted 23 October 2011
Available online 31 October 2011

Keywords:
Biogeography
Remote sensing
Feature extraction
Image classification
Particle Swarm Optimization
Ant Colony Optimization
Kappa Coefficient

ABSTRACT

Recent advances in the theoretical and practical implementations of biogeography have led to the exploration of new bio-inspired techniques which can prove to be the building blocks of hybrid bio-inspired techniques. This aspect was discovered while considering the exploration of bio-inspired intelligence for developing generic optimization algorithms that can be adapted for performing the given land cover feature extraction task at hand. Certain bio-inspired techniques when integrated with the existing optimization techniques can drastically improve their optimization capability hence leading to better feature extraction. In this paper, we propose a generic architectural framework of a hybrid biologically inspired technique that is characterized by its capability to adapt according to the database of expert knowledge for a more efficient, focused and refined feature extraction. Since our hybrid feature extractor possesses intelligence for selective cluster identification for application of either of the constituent techniques which is in turn based on an inefficiency analysis, we term our classifier as the hybrid bio-inspired pattern analysis based intelligent classifier. Our hybrid classifier combines the strengths of the modified BBO Technique for land cover feature extraction with the Hybrid ACO2/PSO Technique for a more refined land cover feature extraction. The algorithm has been tested for for the remote sensing application of land cover feature extraction where we have applied it to the 7-Band carto-set satellite image of size 472 × 546 of the Alwar area in Rajasthan and gives far better feature extraction results than the original biogeography based land cover feature extractor [20] and the other soft computing techniques such as ACO, Hybrid PSO-ACO2, Hybrid ACO-BBO Classifier, Fuzzy sets, Rough-Fuzzy Tie up etc. The 7-band Alwar Image is a benchmark image for testing the performance of a bio-inspired classifier on multi-spectral satellite images since this image is a complete image in the sense that it contains all the land cover features that we need to extract and hence land cover feature extraction results are demonstrated and compared using this image as the standard image.

1. Introduction

Satellite images contain land cover types some of which cover significantly large areas, while some (e.g., bridges and roads) occupy relatively much smaller regions [13]. In the future, new generation of high-resolution satellite sensors will acquire enormous data quantity, so extracting features from a remote sensed image will become even more difficult. Researchers have been continuously searching for new techniques that can extract maximum information from the remotely sensed image [2].

Very recently the concept of swarm intelligence [7,16], an optimized approach of image classification of satellite multi-spectral images, has been introduced in this category. This technique with lower cost and higher degree of classification accuracy, will be able to replace high resolution high cost satellite imageries. Various swarm intelligence techniques such as the Ant Colony Optimization [4], Particle Swarm Optimization [18] and Hybrid ACO2/PSO optimization [5,8,9] have been used for solving the problem of Satellite Image Classification. However, out of these enormous techniques available for image classification, at present, it is not possible to state which classifier is best for all situations as the characteristic of each image and the circumstances for each study vary so greatly. Therefore, it is essential that each analyst

Abbreviations: ACO, Ant Colony Optimization; BBO, Biogeography Based Optimization; PSO, Particle Swarm Optimization; MDMC, Minimum Distance to Mean Classifier; MLC, Maximum Likelihood Classifier; TSP, travelling salesman problem; LISS, linear imaging self scanning; RS1, radarsat 1; RS2, radarsat 2; DN, digital number; *I*, Maximum Immigration Rate; SAR, synthetic aperture radar; GA, genetic algorithm; FCM, fuzzy c-means; RCBBO, real coded biogeography based optimization; HSI, Habitat Suitability Index; SIV, Suitability Index Variables; DPSO, Discrete Particle Swarm Optimization; NIR, near infra red; MIR, middle infra red; DEM, digital elevation model; *E*, Maximum Emigration Rate.

* Corresponding author. Tel.: +91 9899973105, +91 9811052141, +91 0120 2796109.
E-mail addresses: goel_lavika@yahoo.co.in (L. Goel), dgupta@dce.ac.in (D. Gupta), vkpans@ieee.org (V.K. Panchal).

1568-4946/$ – see front matter
doi:10.1016/j.asoc.2011.10.006

understand the alternative strategies for feature extraction so that he or she may be prepared to select the most appropriate technique for the feature extraction task in hand.

Our proposed work is motivated by the very recently introduced swarm intelligence technique called the Biogeography Based Optimization Technique [25] and bases its fundamentals from the research paper titled "Biogeography Based Optimization" [25]. The findings of recent studies are showing strong evidence to the fact that some aspects of biogeography [26] can be adaptively applied to solve specific problems in science and engineering such as in the military application of prediction of enemy base stations [10]. The main characteristic of the BBO Technique that has been modified for application in Satellite Image Classification [20] is that this technique is flexible to classify the desirable features more efficiently than the other features and hence it shows a wide range of efficiencies in classifying different features of an image as was demonstrated in our paper [11]. Also BBO Technique improves its solutions with each iteration and thus probabilistically refines its classification at each iteration. We therefore extend the basic hybrid classification framework in our paper [11] and propose a generic system architecture for feature extraction from a given multi-spectral satellite image.

By the motivation of these characteristic strengths of the above swarm intelligence techniques, we formulated an algorithm which is a hybrid of the ACO, PSO and BBO Techniques for Satellite Image Classification. Hence we propose a swarm intelligence based classification algorithm which integrates the ACO2/PSO Technique with the recently introduced population based swarm intelligence technique called the Biogeography Based Optimization Technique for building an artificially intelligent hybrid classifier.

The organization of the paper is as follows: Section 2 presents a brief review of the related work done in the field of swarm intelligence and our motivation towards the proposed methodology. Section 3 presents a brief review of BBO and Hybrid ACO2/PSO Techniques. Section 4 describes the swarm intelligence techniques of hybrid ACO2/PSO and BBO which have been modified to adapt to the problem of land cover feature extraction. Section 5 presents the proposed framework of our hybrid bio-inspired pattern analysis based intelligent classifier- the system architecture, the mathematical formulation and the detailed functional architecture/algorithm for feature extraction. Section 6 presents the implementation results of the proposed hybrid classifier and compares its efficiency with the BBO Technique Section 7 assesses the accuracy of the proposed algorithm by preparing an error matrix and calculating the user's and the producer's accuracy and also comparing its error matrix with the simple BBO Classifier error matrix. Section 8 presents a comparison of the proposed classifier with the traditional Probabilistic Classifiers such as the MDMC and MLC. Section 9 presents the classified images using other recent Soft Computing Techniques and provides a comparison of the Soft Computing Classifiers vs Probabilistic Classifiers. Section 10 presents Conclusion and future scope of the proposed work.

2. Related work

In remote sensing the problem of land cover feature extraction (or satellite image classification) has been solved by using the traditional classical approaches like Parellelopiped Classification [13], Minimum Distance to Mean Classification [13], Maximum Likelihood Classification [13] etc. However, these techniques show limited accuracy in information retrieval and high resolution image is needed. Also these techniques are insensitive to different degrees of variance in the spectral response data.

To provide a solution to the above problems, soft computing techniques were introduced in remote sensing for image classification. Soft computing techniques differ from the above conventional (hard) computing techniques in that, unlike hard computing, they are tolerant of imprecision, uncertainty and partial truth. In effect, the role model for soft computing is the human mind. The guiding principle of soft computing is: Exploit the tolerance for imprecision, uncertainty and partial truth to achieve tractability, robustness and low solution cost. The principal constituents of soft computing techniques are fuzzy logic [19], rough set theory [22–24], neural network theory, probabilistic reasoning, and Swarm Intelligence Techniques [16] with the latter subsuming belief networks, genetic algorithms, chaos theory and parts of learning theory. However, the soft computing techniques like the fuzzy classifier [19] and the rough set classifier [19,21] were not able to provide good result in case of ambiguity sine the main goal of these techniques was to synthesize approximation of concepts from the acquired data. Hence, these techniques did not provide very much accurate results with low spatial resolution images. Also these techniques were not able to handle the crisp and continuous data separately.

The solution to the above drawbacks was provided by the recently introduced concept of swarm intelligence [16]. Our proposed work bases its fundamentals from Swarm Intelligence. This technique improves the classification of satellite multi-spectral images and is more accurate when working with low spatial resolution images. As the frontiers of space technology advance, the knowledge derived from the satellite data has also grown in sophistication. We need more accurate and efficient image classification algorithms. Also no single classifier can prove to satisfactorily classify all the basic land cover classes of a region. Hence we propose a swarm intelligence based classification architecture which combines the strengths of the ACO2/PSO technique [5,8] with the characteristic strengths of the recently introduced population based swarm intelligence technique called the Biogeography based optimization technique [25]. The main characteristic of the BBO technique for satellite image classification is that this technique is flexible to classify the desirable features more efficiently than the other features and hence it shows a wide range of efficiencies in classifying different features of an image. Also BBO technique improves its solutions with each iteration and thus probabilistically refines its classification at each iteration. This combination gives this classifier artificial intelligence to identify the features efficiently classified by BBO and by ACO2/PSO separately based on an analysis of the training set data distribution graph.

3. A brief review of BBO and Hybrid ACO2/PSO Techniques

This section briefly reviews the bio-inspired techniques of BBO, ACO and PSO used in the development of our hybrid bio-inspired pattern analysis based intelligent classifier.

3.1. Biogeography Based Optimization

BBO is a population based EA wherein optimization is done based on migration of species. As a global optimization method, BBO is an original algorithm based on the mathematical model of organism distribution in biological systems. BBO is an evolutionary process that achieves information sharing by biogeography-based migration operators. In BBO, habitats represent candidate problem solutions, and species migration represents the sharing of features between candidate solutions according to the fitness of the habitats. The idea of BBO was first presented in December 2008 by Simon [25]. One characteristic of BBO is that the original population is not discarded after each generation; it is rather modified by migration. Also for each generation, BBO uses the fitness of each solution to determine its emigration and immigration rate [10,25–27]. In a way, we can say that BBO is an application of biogeography to EAs. In BBO, each individual is considered as a habitat with a HSI

[10,25–27], which is similar to the fitness of EAs, to measure the individual fitness. Also, an SIV which characterizes the habitability of an island is used. A good solution is analogous to an island with a high HSI, and a poor solution indicates an island with a low HSI. High HSI solutions tend to share their features with low HSI solutions. Low HSI solutions accept a lot of new features from high HSI solutions [25,26]. The BBO Algorithm proposed by Dan Simon is presented below [25].

(a) Initialize the BBO parameters. This means deriving a method of mapping problem solutions to SIVs and habitats which is problem dependent.
(b) Initialize the maximum species count S_{max} and the maximum migration rates E and I, the maximum mutation rate m_{max}, and an elitism parameter.
(c) The maximum species count and the maximum migration rates are relative quantities. That is, if they all change by the same percentage, then the behavior of BBO will not change. This is because if E, I, and S_{max} change, then the migration rates immigration rate (μ), emigration rate (λ), and the species count S will change by the same relative amount for each solution.
(d) Initialize a random set of habitats, each habitat corresponding to a potential solution to the given problem.
(e) For each habitat
 Map the HSI to the number of species S, the immigration rate λ, and the emigration rate μ.
 Probabilistically use immigration and emigration to modify each non-elite habitat.
 Re-compute each HSI.
 Update the probability of its species count using (d). Then mutate each non-elite habitat based on its probability and re-compute each HSI.
End For
(f) This loop can be terminated after a predefined number of generations or after an acceptable problem solution has been found.

It is to be noted that after each habitat is modified, its feasibility as a problem solution should be verified. If it does not represent a feasible solution, then some method needs to be implemented in order to map it to the set of feasible solutions [25].

3.2. Hybrid ACO2/PSO optimization

Parpinelli et al. were the first to propose ACO for discovering classification rules [17], with the system *Ant-Miner*. Fernando et al. [9] proposed an extension to Ant-Miner, named cAntMiner, which was able to cope with the continuous values as well.

The Ant Miner and cAntMiner have already been a significant approach for data mining, but an extremely large amount of computation is required with the problem of unusually large amount of attributes and classes. The "standard" binary/discrete PSO algorithm [18] does not deal with categorical values in a natural fashion when compared to ACO. In particular, the standard PSO for coping with binary attributes represents a particle by a bit string, where each binary value such as true or false is encoded as 1 or 0. Later, the standard binary PSO was modified to cope with multi-valued categorical attributes [18], developing a DPSO algorithm for discovering classification rules.

Unlike a conventional PSO the Hybrid PSO-ACO algorithm can directly cope with the nominal attributes, without converting nominal values into numbers in a pre-processing phase. The Hybrid PSO-ACO given by Nicholas and Frietas uses sequential covering approach for rule extraction [6,8]. After that they also proposed a new modified version PSO-ACO2 directly deals with both the continuous and nominal attribute-values [9].

Both the original PSO/ACO algorithm and the new modified version PSO/ACO2 use a sequential covering approach to discover one classification-rule-at-a-time. The hybrid version given by Nicholas and Freitas can be understood as below.

(a) Initially Rule Set is empty (Φ)
(b) For Each class of cases Trs = {All training cases}
 While (Number of uncovered training cases of class A > Maximum uncovered cases per class)
  Run the PSO/ACO algorithm for finding best nominal rule
  Run the standard PSO algorithm to add continuous terms to Rule, and return the best discovered rule BestRule
  Prune the discovered BestRule
  RuleSet = RuleSet U BestRule
  Trs = Trs − {training cases correctly covered by discovered rule}
 End of while loop
End of for lop
(c) Order these rules in RuleSet by descending Quality

It is necessary to estimate the quality of every candidate rule (decoded particle). A measure must be used in the training phase in an attempt to estimate how well a rule will perform in the testing phase. Given such a measure it becomes possible to optimize a rule's quality (the fitness function) in the training phase and this is the aim of the PSO/ACO2 algorithm. In PSO/ACO [7,8] the Quality measure used was Sensitivity × Specificity where TP, FN, FP and TN are, respectively, the number of true positives, false negatives, false positives and true negatives associated with the rule [6–8].

$$\text{Sensitivity} \times \text{Specificity} = \frac{\text{TP}}{\text{TP}+\text{FN}} \times \frac{\text{TN}}{\text{TN}+\text{FP}} \tag{1}$$

Original Quality Measure.

Later it is modified as Eq. (2) given below:

$$\text{Sensitivity} \times \text{Precision} = \frac{\text{TP}}{\text{TP}+\text{FP}} \times \frac{\text{TP}}{\text{TP}+\text{FP}} \tag{2}$$

Quality Measure on Minority Class.

This is also modified with using Laplace correction as;

$$\text{Precision} = 1 + \frac{\text{TP}}{1+k+\text{TP}+\text{FP}} \tag{3}$$

New Quality Measure on Minority Class, where 'k' is the number of classes.

So, PSO/ACO attempted to optimize both the continuous and nominal attributes present in a rule antecedent at the same time, whereas PSO/ACO2 takes the best nominal rule built by PSO/ACO2 and then attempts to add continuous attributes using a standard PSO algorithm.

4. Land cover feature extraction using ACO2/PSO and BBO Techniques

In this section, we shall describe the swarm intelligence techniques of Hybrid ACO2/PSO and BBO which have been modified to adapt to the problem of land cover feature extraction. The details of the dataset used for the purpose of land cover feature extraction are as described in Section 4.1.

4.1. Dataset used

Our objective is to use the above swarm intelligence techniques as an efficient land cover classifier for satellite images. We have taken a multi-spectral, multi resolution and multi-sensor image of size 472 × 546 pixels of Alwar area in Rajasthan, India. The satellite image for 7 different bands is taken. These bands are red, green, near

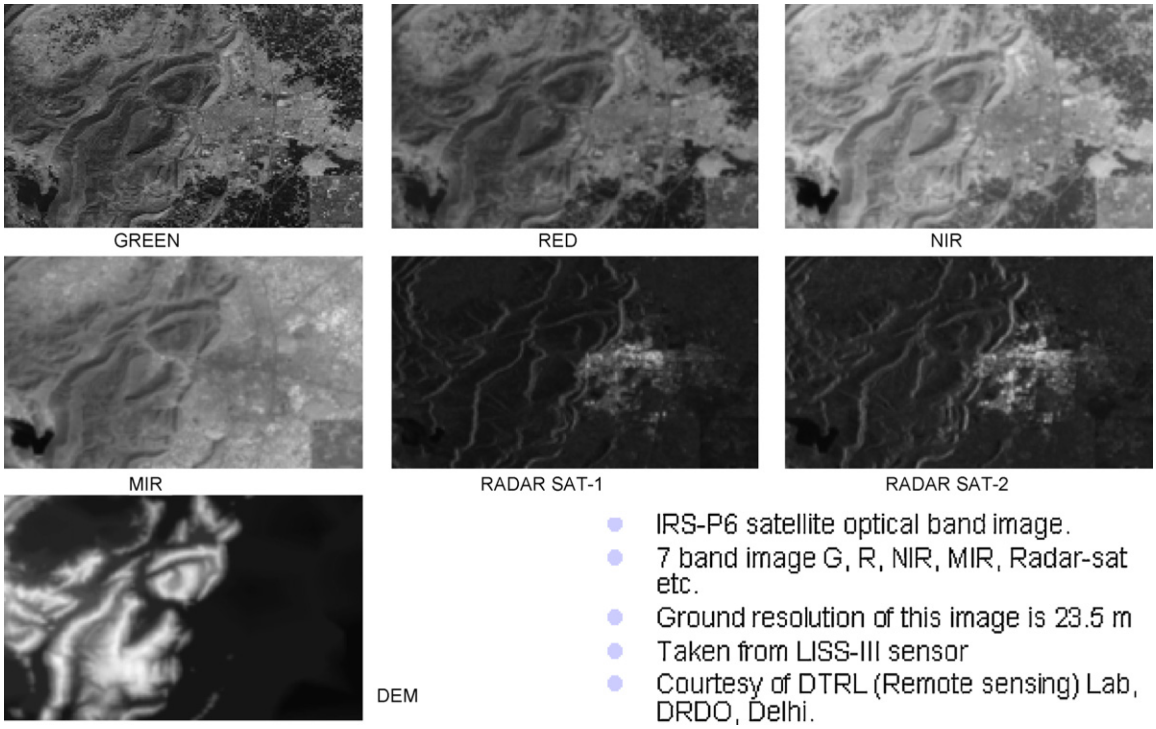


Fig. 1. 7-Band Satellite Image of Alwar Area in Rajasthan, India (Courtesy of Defense Terrain & Research Lab (DTRL), Defense and Research Development Organization (DRDO), India).

infra-red, middle infra-red, radarsat-1, radarsat-2, and digital elevation model. The ground resolution of these images is 23.5 m and is taken from LISS-III sensor. The 7-band satellite image of Alwar area in Rajasthan is given in Fig. 1.

4.2. Hybrid ACO2/PSO land cover feature extraction

The hybrid ACO2/PSO algorithm for land cover feature extraction detailed in [5] is represented diagrammatically in Fig. 2 and is formally presented in the form of an algorithm. The input to the algorithm is the multi-spectral multi-band satellite image and the output is the classified image in .jpeg, .tiff format. By these multi-band images the experts generate the training set. Experts of remote sensing open these images in ERDAS IMAGINE [13,14]. They manually select the pixels of different class regions and set a definition color code using ERDAS Signature Editor. This data set is provided by the experts in the form of digital numbers (intensity value pixel in a digital image). This dataset can be in any Excel Sheet, Access sheet or in Text file. The Decision according to the DN values is also set in the table.

From these DN values we extract rules by applying the PSO/ACO2 algorithm. To apply the Modified Hybrid ACO/PSO2, we use the Open Source Tool for Hybrid ACO/PSO2. Before applying these algorithms for rule-set generation we change the training set into the unified Attribute Relation File Format, i.e., .arff format. We obtain the rule sets from the Open Source Tool for PSO/ACO2 [28].

But this rule set does not provide the classified image. So we code these rules in a MATLAB file. In MATLAB coding the multi-band images are read and then for each pixel of the image, the DN values are checked according to the rule set contained. If the pixel validates any rule then the decision of the rule is set to be the class of the pixel and the pixel is set to a specified color. On executing this MATLAB file we obtain the *Final Classified Image*.

The region matching or classification coding is done in MATLAB and it will give the output classified image on executing.

(a) Retrieve the original multi-spectral multi-band image in .tiff format.
(b) Generate the training set using ERDAS software, according to the different classes generated by experts.
(c) Convert this training data set into the unified Attribute Relation File Format.
(d) Apply the Hybrid PSO/ACO2 on the training data set obtained at step (c), by the Open Source Tool to get the rule set and go to step (e).
(e) For each rule in the rule set perform the following actions:
  For each pixel of the original image
    If the pixel validates the rule
    Then
    The pixel is set to be in the region decided by the rule and the color code is assigned.
    Else
    Move to the next rule.
(f) Finally the classified image is obtained in .jpeg, .jpg, .tiff or any other image format.

4.3. Biogeography based land cover feature extraction

This section describes the details of the biogeography based land cover feature extraction process – the parameters and the definitions used, the assumptions made, the algorithm used and the other architectural details including the derivation of the mathematical formulation of the algorithm. The procedure of biogeography based land cover feature extraction mainly incorporates the following steps:

a. Define the different parameters used in the Biogeography based land cover feature extraction algorithm which are required as the preprocessing steps towards the application of the BBO Technique for the purpose of land cover feature extraction and also state the assumptions made.

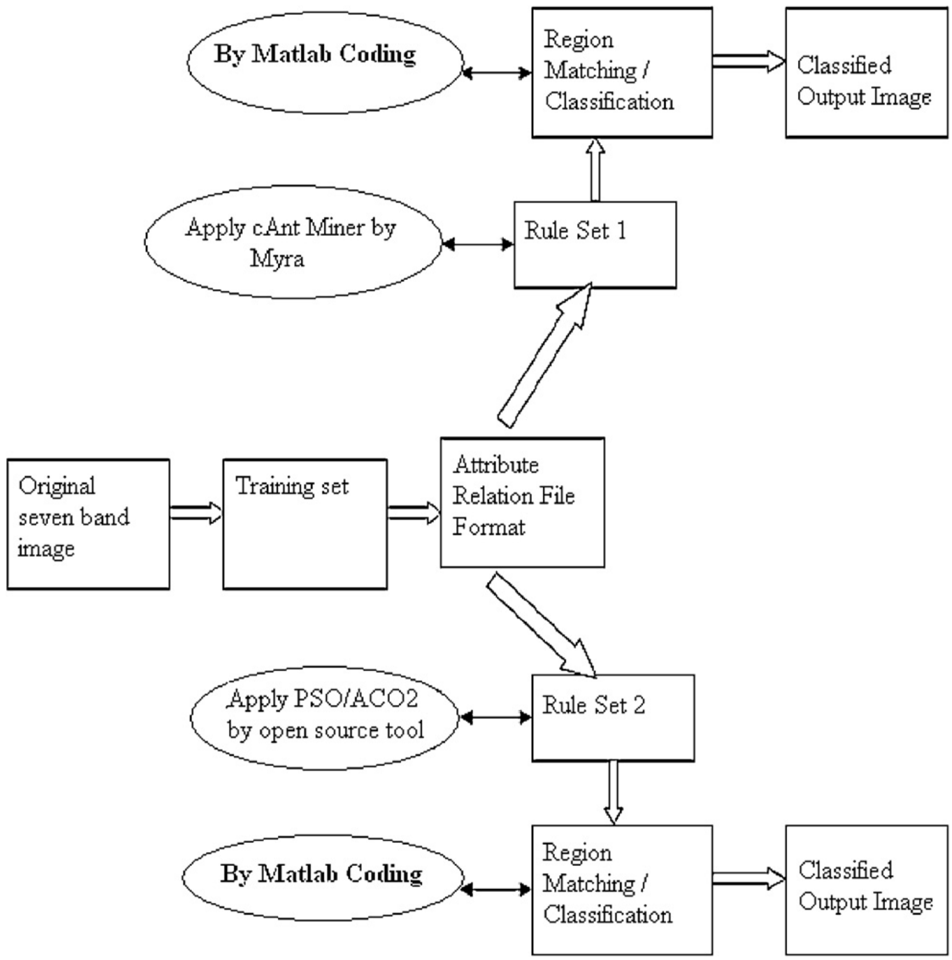


Fig. 2. Diagrammatic representation of the Hybrid ACO2/PSO Algorithm for land cover feature extraction.

b. Initialize the biogeography based land cover feature extraction algorithm by generating unsupervised clusters (elementary classes), using the rough set toolkit, ROSETTA [15], that need to be classified into an appropriate land cover feature upon completion of the algorithm.
c. Modify the BBO Technique proposed by Simon [25] to suit the purpose of land cover feature extraction.
d. Run the Biogeography based land cover feature extraction algorithm on the given multi-spectral satellite image dataset.

4.3.1. Some important relevant definitions for the biogeography based land cover feature extraction algorithm

We now identify some important relevant definitions of the parameters originally proposed by Simon [25] which have been defined to suit our purpose of land cover feature extraction [20]. These parameters define the preprocessing steps of the biogeography based land cover feature extraction algorithm. Following are some important relevant definitions of the parameters and the other concepts used:

Assumption: Species Si in this study refers to the image pixels and the Image *I*, may be considered as a set of species.

$$I = \{Si | \forall i \in [1 \cdots (\text{Size of image} \quad I)]\}$$

Definition 1. *Eco-System*: It is described as under [20]:

(i) A *habitat* is a place where resides some species and that are likely to be migrated to other more suitable habitats. Let the total population of all kinds of species reside in *N* habitats, *Hi*.
(ii) The habitat *Hi*, is not of homogeneous composition of species. It contains different types of species.
(iii) A *Universal Habitat* $\mathbf{\Pi}$ is considered which hosts all the image species. Therefore,

$$\mathbf{\Pi} = \cup_{i=1}^{N} Hi$$

(iv) *Feature Islands*: these are the habitat islands populated by the homogeneous species. These are nothing but the labeled dataset/training dataset generated by the expert. Five types of the feature Islands are considered here, namely: *water*, *vegetation*, *rocky*, *urban* and *barren*.
(v) Therefore, the ecosystem *H6* is a group of 6 habitats (one Universal Habitat and five Feature Islands, Fig. 3).
(vi) The *eco-system* therefore is illustrated in Fig. 3.

Definition 2. *Suitability Index Variable* (SIV): Each of the multi-spectral band of image represents one Suitability Index Variable (SIV) of the habitat. Further, image in each band is a gray image; therefore, SIV $\in C$ is an integer and $C \subseteq [0,255]$.

Definition 3. *Habitat*: A habitat $H \in \text{SIV}^m$ where *m* is 7 = total image bands.

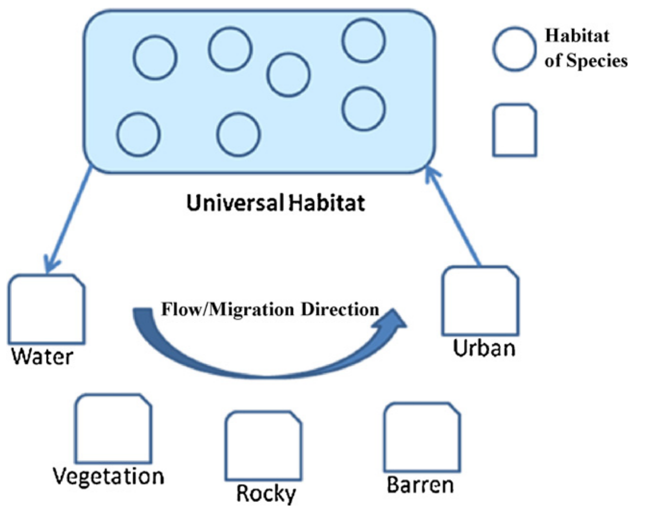


Fig. 3. Eco-system – initial state of Habitats.

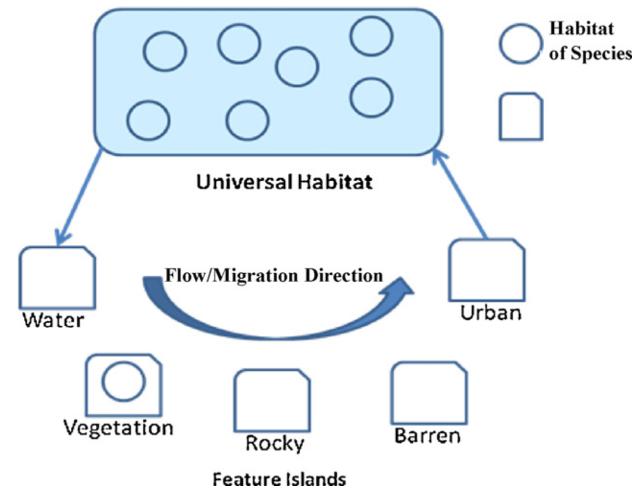


Fig. 4. State after first generation.

Definition 4. *Unsupervised Clustering*: The image I is subjected to 20 simple partitions Pi. The image may be partitioned into any number as per the individual preference.

$$I = \{Pi, \ldots, P20\}$$

Each Pi is subjected to unsupervised clustering so that based on some criteria derived from the Pi statistics, new smaller clusters are generated.

The Rough Sets based clustering [21–23] is performed to Pi, and new, unsupervised, clusters are generated. These are now known as the new habitats Hi.

The preference to the Rough Sets is given because of the following reasons:

i. The concept of *equivalence class* [21,24] automatically generates the clusters based on the criteria of indiscernibility.
ii. This method does not need to know *a priori*, the *number* of clusters required.

For the indiscernibility criteria the attributes are the NIR and MIR bands. It has been observed that NIR and MIR bands contain a good amount of geospatial information [19].

The set of all equivalence class/elementary knowledge granules, i.e., Hi are nothing but the partition of Pi.

$$\frac{Pi}{R} = \{[x]R : x \in Pi\}$$

where the equivalence relation $R = \mathrm{IND}(\{\mathrm{NIR}, \mathrm{MIR}\}) = \{(x, y) \in Pi\}$

Therefore, Hi, may now be considered as *mixed species habitat* and therefore contains uncertainty. The removal of the uncertainty in the type of the species is carried out in this section.

Let $J = |Pi/R|$,

Then J gives the no. of knowledge granules generated from Pi.

From the above discussion we infer the following:

$$\prod = \cup_i^{20} Pi$$

where $Pi = \cup_i^J Hi$. For the sake of simplicity we may define HP, as a single unsupervised cluster of the Pth partition of the image I.

Definition 5. *Feature sharing in BBO*: The original BBO Algorithm [25] proposed the migration of SIV values from a high HSI habitat to a low HSI habitat. The shared features (SIV) remain in the high HSI solutions, while at the same time appearing as new features in the low HSI solutions.

In this algorithm, rather than moving SIV, we are moving species altogether from a Universal Habitat to feature Islands. The species does not remain shared: it is removed from the Universal Habitat and migrated to feature habitat (Fig. 4).

Definition 6. *Habitat Suitability Index (HSI)*: As it is the maiden exploration of the BBO in any classification problem, it was found appropriate to choose standard deviation as the HSI.

Definition 7. *Heuristic Method*: The major steps we followed are image clustering and heuristic method (HSI) implementation. The 7-bands image is clustered into Hi, using rough set theory. The parameters of clustering are NIR and MIR bands. The resultant clusters or the habitats Hi are species in Universal Habitat (Fig. 4).

These species are migrated to the suitable habitats, i.e., feature islands. These habitats are actually the decisions, for the resultant different feature groups.

The heuristic method decides which species are moved to which habitat using a fitness function. Here we have used mean of standard deviation as fitness function. This function for a specific class takes up the corresponding training sets and then calculates the required function values, which further helps to decide the most suitable habitat.

Let σF_i, be the average of the standard deviations of any of the feature islands: water, vegetation, rocky, barren and urban respectively.

$$\sigma F_i = \frac{\sigma_{\mathrm{R}i} + \sigma_{\mathrm{G}i} + \sigma_{\mathrm{NIR}i} + \sigma_{\mathrm{MIR}i} + \sigma_{\mathrm{RS1}i} + \sigma_{\mathrm{RS2}i} + \sigma_{\mathrm{DEM}i}}{7}$$

where $\sigma_{\mathrm{R}i}, \sigma_{\mathrm{G}i}, \sigma_{\mathrm{NIR}i}, \sigma_{\mathrm{MIR}i}, \sigma_{\mathrm{RS1}i}, \sigma_{\mathrm{RS2}i}, \sigma_{\mathrm{DEM}i}$ are the standard deviations of DN values of the red, green, NIR, MIR, RS1, RS2 and DEM bands of the ith feature island.

Similarly, let the average of the standard deviation of the species habitat C_j from Hi of the Universal Habitat $\mathbf{\Pi}$ be represented by:

$$sc_j = \frac{\sigma_{\mathrm{R}j} + \sigma_{\mathrm{G}j} + \sigma_{\mathrm{NIR}j} + s_{\mathrm{MIR}j} + s_{\mathrm{RS1}j} + s_{\mathrm{RS2}j} + s_{\mathrm{DEM}j}}{7}$$

Then the heuristic function Φ be defined as

$\Phi = |\sigma c_j - \sigma F_i| \le \delta$. Then σc_j may be thought of belonging to the feature island class σF_i,
$\Phi = |\sigma c_j - \sigma F_i| \le 0$. Then σc_j is exactly the σF_j class,
$\Phi = |\sigma c_j - \sigma F_i| > \delta$. Then σc_j is of different type than σF_i,

where δ is a threshold.

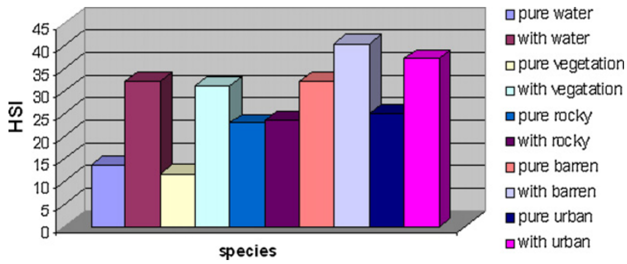


Fig. 5. HSI matching in habitats.

Also, Heuristic function makes use of a specific threshold value. For higher accuracy, the threshold value should be low. But with low threshold value the number of generations rises at a rapid pace. If for a species, no suitable habitat is found in the current generation, it implies that there are a large number of features in the given section of the image (defined as a species). This unclassified cluster migrates to the Universal Habitat. Here it is again classified using rough set theory. This is termed as second generation.

This function decides which value of mean of standard deviation has minimum difference from the original class. If this value is within the threshold then that class (species) will migrate to that habitat.

The recalculated HSI after the migration of species is compared with the original HSI of the habitat (that contained training pixels only). If the compared HSI is within the threshold, i.e., it varies from −1 to +1 from the original HSI then it means that the habitat is suitable for that species and hence it can migrate to it. Fig. 5 shows the comparison of HSI of training sets' pixels with recalculated HSI after migrating the species. If the recalculated HSI is not within the threshold for any of the original habitat then the particular elementary class contained a mixture of species and it is again partitioned into elementary classes.

These classes are migrated into the Universal Habitat and considered as unclassified species. These are further used for the next generation/iteration.

The checking of all the habitats results in the completion of first generation and then we move towards the second generation for further refinement following the similar process. The maximum number of generations can correspond to the situation when we are left with no more unclassified class (species) in the Universal Habitat.

Definition 8. *Migration Rates*: Maximum Immigration Rate and Maximum Emigration Rate are same and equal to number of species in the habitat. The algorithm follows a linear curve ($E = I$). Number of species and thus the Maximum Immigration Rate and Maximum Emigration Rate can vary in each iteration. Maximum species count ($S_{\max}$) and the maximum migration rates are relative quantities. That is, if they all change by the same percentage, then the behavior of BBO will not change [1,25,26].

Definition 9. *Mutation*: Since mutation is not an essential feature of BBO [25], it is not required in the proposed algorithm. Elitism, too, is an optional parameter [25,26], it has not been in the proposed algorithm.

4.3.2. Assumptions made

The assumptions made in this algorithm can be summarized as follows [20]:

- Initially it is considered that there exists a Universal Habitat consisting of all the species.
- It has been assumed that species migrate from one habitat to another habitat as a mixed population.
- Also for the feature extraction purpose, in each generation, all the habitats have been considered exactly once.
- In next generation only the unclassified pixels have been used. Unclassified pixels are those pixels which have not been categorized into one of the five land cover features at the end of one generation and hence do not get absorbed in any of the feature habitats. These are left over in the Universal Habitat (to be considered in the next generation) and marked as unclassified. The mechanism for classification is described in Section 4.3.3. However the total population is still present.
- Also, we have used NIR and MIR band of the image for partitioning the image (done using rough set theory).
- During partitioning, we have used Equal Frequency Binning algorithm for discretization.
- We have used 1 as threshold. Less will be the threshold better will be the results with more number of generations.

4.3.3. Algorithm for biogeography based land cover feature extraction

The modified version of the original BBO algorithm that has been modified for extraction of land cover features from the satellite image proposed by Panchal et al. [20] is presented in this section. The input to the algorithm is the multi-spectral satellite image and the output is the extracted features from the image. The flowchart for the biogeography based land cover feature extraction algorithm is presented in Fig. 6.

(a) Get the multi-spectral satellite image.
(b) Cluster the pixels of image randomly (using rough set theory) and consider each cluster as a species of the Universal Habitat.
Take each of the 5 land cover features as a feature habitat. Hence, we have 5 feature habitats each for water, urban, rocky, barren and vegetation features, having members produced by experts.
(c) Define HSI, $S_{\max}$, immigration rate (λ) and emigration rate (μ). Usually in BBO a high μ (and low lambda) means high fitness, so that individual shares features with other individuals. In our case, species are actually moved from the Universal Habitat to the feature habitat and hence a high fitness specie will migrate to one of the feature habitats and a low-fitness specie will have a low emigration rate and will remain in the Universal Habitat at the end of the iteration.
(d) Calculate HSI of each of the feature habitat.
(e) For each species in the Universal Habitat
  Select a species from the Universal Habitat and migrate it to one of the feature habitat.
  Recalculate the HSI of the feature habitat after the migration of the species to it.
    If the recalculated HSI is within the threshold, then:
      Absorb the species in the feature habitat.
    Else if any unconsidered feature habitat is left then:
      Migrate the species to that feature habitat and recalculate the HSI after migration.
    Else
      Use rough set theory to discretize the species (as it contained mixed pixels) and make random clusters which are then considered as separate species.
      Add these new species to the Universal Habitat.
    End If.
  If no species is left in the Universal Habitat then:
    Stop the process.
  End If
End For.

The detailed working of the above algorithm is described as follows:

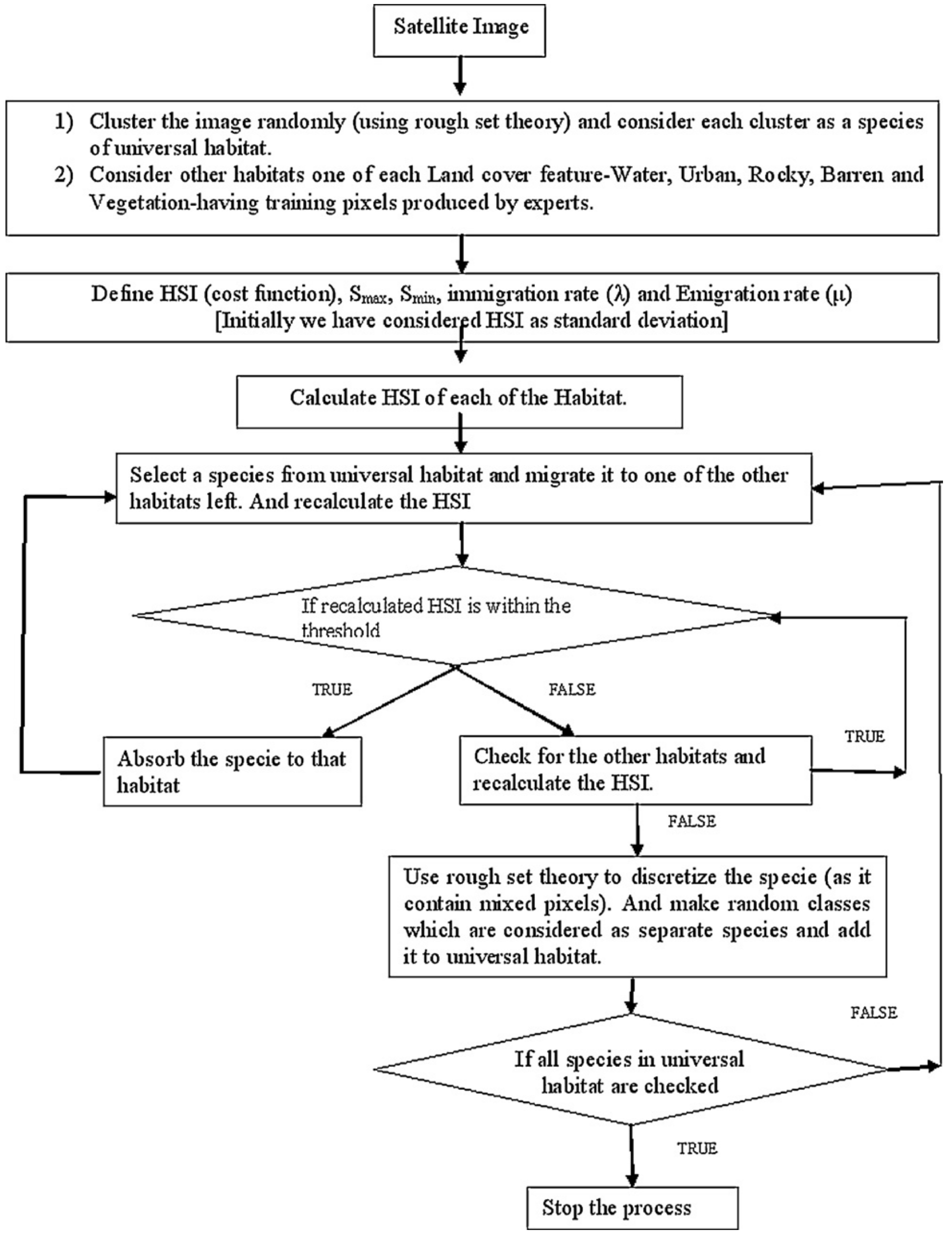


Fig. 6. Flowchart representation for the biogeography based land cover feature extraction algorithm.

Step (i): A 7-band satellite image of Alwar Region of size 472×546 is taken as an input image for Biogeography based feature extraction. The image has 257,712 pixels and each pixel is of 25 m by 25 m spatial resolution. The image is divided into 20 equal parts (however, this value can be varied) and each part corresponds to a class. Each pixel in a class has x and y coordinates along with the 7-band values of each pixel. Each class is considered separately and the following steps are applied to each class.

Step (ii): The classes are discretized using the Rosetta software version 1.4.41 [15]. ROSETTA is *ro*ugh *set* *t*oolkit for *a*nalysis. ROSETTA is a toolkit for analyzing tabular data within the framework of rough set theory [15]. The image data is first read in ROSETTA [15].

The elementary classes are obtained by using discretization followed by partitioning process in ROSETTA. Input read data is further discretized by using naïve discretization algorithm defined in ROSETTA. The no. of elementary classes produced depends upon the clustering of pixels into similarity classes (equivalence classes). The pixels which are similar (the definition of similarity here depends upon the discretization algorithm used) are grouped into one elementary class. In the resulting data, DN value of each pixel is replaced by the discretized interval. Each row in the resulting data represents the x, y coordinates and the DN values in the red, green, NIR, MIR, RS1, RS2 and DEM bands respectively of each pixel in the image.

This discretized data is further partitioned on the basis of NIR and MIR band to obtain the final elementary classes. The band used for partitioning may affect the final results. In our case, we chose the NIR and MIR bands for partitioning. Each elementary class consists of the index of each pixel belonging to it. We define the index of each pixel as the pixel no. of the pixel in each cluster of the image containing 257,712 pixels. Hence, the pixel numbers for a pixel in a cluster will range from 1 to (1/20 × 257,712), i.e., 1 to 12,886 for each of the 20 clusters. The pixel's DN values corresponding to these indexes are obtained before applying the biogeography based feature extraction algorithm.

The elementary classes obtained by ROSETTA are put in the Universal Habitat and each elementary class is treated as a species.

Step (iii): For our process we have considered five feature habitats namely urban, rocky, vegetation, water and barren land along with the Universal Habitat. These feature habitats consist of training sets as species. The training set consists of those pixels whose feature is already known. The training sets are provided by the expert.

Step (iv): The HSI is the standard deviation of all the pixels in the elementary class and is calculated on each band of the image separately. $S_{\max}$ is the maximum number of species and its value varies in each iteration.

Step (v): The feature habitats initially contain the training pixels of the corresponding feature and HSI is calculated on those training pixels. For example, water habitat initially contains pixels of water and standard deviation of all the training pixels of water is calculated in each of seven bands.

Step (vi): Each species is taken from the Universal Habitat and migrated to each of the feature habitat one at a time. The HSI of the habitat is recalculated after the migration of the species to it.

Step (vii): (a) The recalculated HSI after the migration of the species is compared with the original HSI of the habitat (that contained training pixels only). If the compared HSI is within the threshold, i.e., it varies from −1 to +1 from the original HSI, then it means that the habitat is suitable for that species and hence it can migrate to it.

(b) If the recalculated HSI is not within the threshold for any of the original habitat then the particular elementary class contained a mixture of species and it is again partitioned into elementary classes using Rosetta software. These classes are migrated into the Universal Habitat and considered as unclassified species. These unclassified species are further used for the next generation.

Step (viii): The checking of all the habitats results in the completion of first generation and then we move towards the second generation for further refinement following the similar process. The maximum number of generations can correspond to the situation when we are left with no more unclassified class (species) in the Universal Habitat.

5. Proposed framework for the hybrid bio-inspired pattern analysis based intelligent classifier

In this section, first we propose the system architecture for the hybrid bio-inspired pattern analysis based intelligent classifier. We then abstract the mathematical formula used and go on to present the detailed functional architecture/algorithm for feature extraction.

5.1. System architecture

We have implemented BBO algorithm combined with the ACO2/PSO technique for the purpose of features extraction from a satellite image. The generic system architecture of our bio-inspired pattern analysis based intelligent classifier is given in Fig. 7.

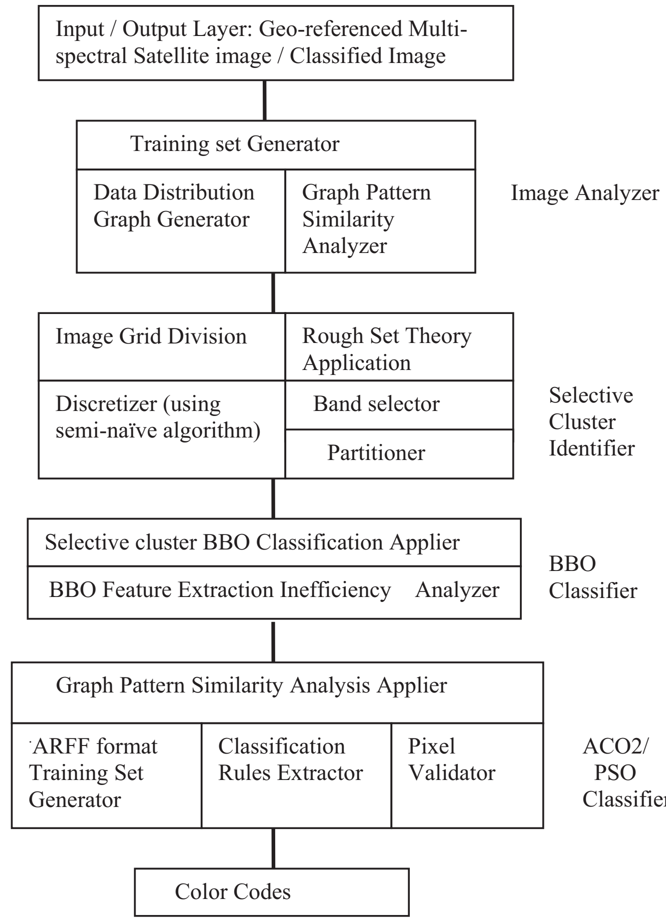


Fig. 7. Software architecture of the proposed intelligent classifier.

The above layers of the system architecture of our proposed hybrid classifier is explained step by step below:

Input/Output layer: This layer is used to input the high resolution multi-spectral satellite image and output the final classified image with the land cover features extracted. In our case, the image used is the 7-band cartoset satellite image of size *472* × *546* of the Alwar Region in Rajasthan and the output image is the classified image in .tiff or in .jpeg format.

Image analyzer layer: This layer is responsible for analysis of the input satellite image to exploit the data distribution of the training set and feed its output to the next layer, i.e., selective cluster identifier layer.

(i) **Training set Generation Layer:** In this layer, the training set is generated by the expert. In our case, we use ERDAS software for generating the training sets.

(ii) **Data Distribution Graph Generation Layer:** This layer takes as input the training set and plots the the data distribution graph between the average of the Standard Deviations of each land cover feature (plotted on the y-axis) for each of the multi-spectral bands of the image (plotted on the x-axis). In our case, the land cover features that are plotted are the water, urban, rocky, vegetation and barren and the multi-spectral bands are the Red, Green, NIR, MIR, RS1, RS2 and DEM bands as shown in Fig. 8. On the *x*-axis, each of the 7-bands of the image are represented as integers from 4 to 10 for red, green, NIR, MIR, RS1, RS2 and DEM bands respectively. We are not considering integers 1–3 in the graph plot on the *x*-axis since

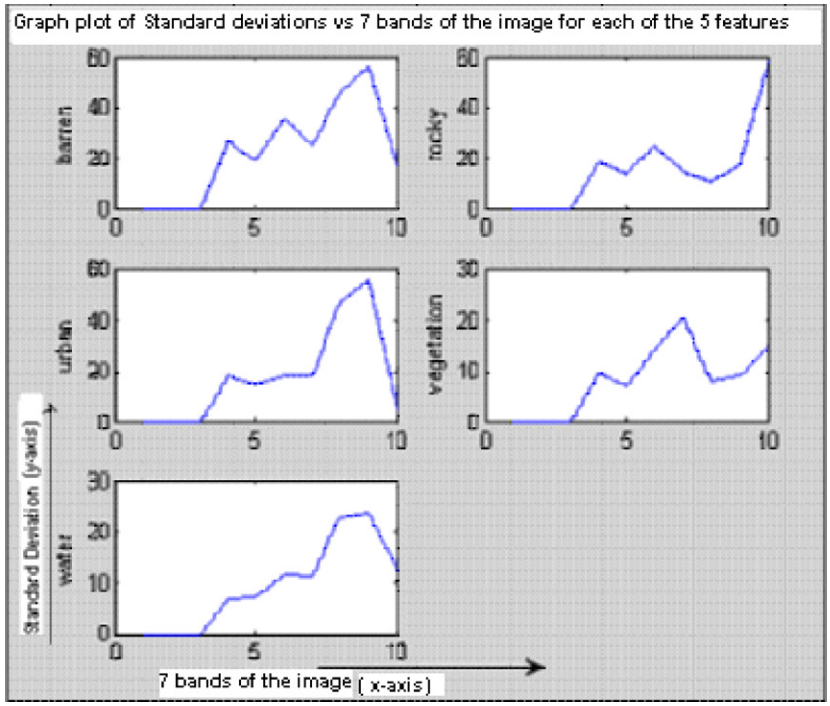


Fig. 8. Graph plot of the standard deviations of each land cover feature vs each of the 7-bands in which the Alwar Image is viewed.

these represent the other attributes of the image namely pixel no. and x and y coordinates which are not required at this stage.

(iii) **Graph Pattern Similarity Analysis Layer:** This layer is responsible for analyzing the data distribution graph and output the similarities in the graph pattern in each of the multi-spectral bands of the image. In our illustration, we observe that the minimum difference between the average standard deviations of the NIR and the MIR bands of the Alwar Image is achieved in particularly two land cover features, those of water and urban area, both of which exhibit the same graph pattern in the NIR and the MIR bands, i.e.

|average of standard deviation of NIR band

− average of standard deviation of the MIR band| lowest

= {water, urban}

Hence, this layer will identify that these are the two features that will be most efficiently classified by our hybrid algorithm which works in the NIR and MIR bands.

Selective Cluster Identifier Layer: This layer is responsible for selecting the appropriate clusters of the image on which BBO technique will show maximum classification accuracy and hence should be applied to. Based on the observation of the adaptive nature of BBO Algorithm for land cover feature extraction [11], we know that BBO is able to classify some particular feature's pixels with greater efficiency than the other features based on the band that we select for creation of equivalence classes in Rosetta. The functionality of this layer is supported by three subunits as described below:

(i) **Image Grid Division Layer:** In this layer, the satellite image is divided into desired number of clusters, say n. In our case, $n = 20$.

(ii) **Band selection Layer:** Depending on our application, in other words depending on which feature we want to extract from the image most efficiently, this layer provides option to choose the band for partitioning. For example, if we want to extract the barren area more efficiently, we choose the green band and for rocky region extraction we choose the MIR band. The RS-1 and RS-2 bands are used to extract the urban area and also for extracting the edges of rocky region from the 7-band image. However, the drainages of rocky region are best viewed in the red band and water and vegetation pixels are best viewed in NIR and MIR bands. This is because BBO follows an adaptive strategy and will classify the desired feature most efficiently since it has its corresponding band selected in the partitioning step. For our illustration, we have chosen the NIR and MIR band of the 7-band image since we want to extract the water pixels effectively and clearly identify the water body in the image and these are the bands in which the water feature is particularly more highlighted and best viewed.

(iii) **Discretization and Partitioning Layer:** This layer is responsible for the discretization and partitioning of each of the clusters of the image obtained in the image grid division layer. In our illustration, since we choose the NIR and the MIR bands in the band selection layer above, we use these bands in the discretization and partitioning step. The algorithm that we used for the discretization of the clusters is the semi –naïve algorithm since this algorithm is capable of discretizing a continuous attribute by sorting the attribute values and hence is an advanced discretization algorithm.

(iv) **Rough Set Theory Application Layer:** This layer is responsible for creating rough set generated equivalence classes, thus creating equivalence classes for each of the clusters. This is what is termed as Unsupervised Classification. Each of these resultant classes are put in the Universal Habitat. In our illustration, we use rough set theory toolkit i.e. *Rosetta software* [15] for the generation of rough sets equivalence classes.

BBO Classifier Layer: This layer applies the BBO technique for feature extraction on the clusters returned by the selective cluster identifier layer. This layer also identifies the clusters on which BBO is inefficient and should not be applied. This layer has two sub units / layers:

(i) **Selective Cluster BBO Classification Applier Layer:** Hence, in this layer, we apply BBO Technique on those clusters of the satellite image which are returned by the selective cluster identifier layer because these clusters show the maximum classification efficiency which is due to the fact that these are the clusters which predominantly show the presence of the feature that is most efficiently classified by the BBO Algorithm. The procedure for applying the biogeography based land cover feature extraction technique was described in Section 4.3.

In our illustration, BBO will be able to extract the water pixels most efficiently since we choose the NIR and MIR bands for partitioning. And therefore, we apply BBO on the *k*th cluster of the Alwar Image since this is the cluster which gives the maximum classification efficiency because it predominantly shows the presence of water body in the image.

(ii) **BBO Feature Extraction Inefficiency Analysis Layer:** This layer identifies the features which cannot be efficiently classified by the BBO technique as compared to the other techniques such as the ACO2/PSO technique. This is because it has been observed that BBO shows a wide range of efficiencies in classifying the rest of the features. In other words, BBO is not able to classify all the 5 features with good classification efficiency and shows poor performance on some of the remaining features. In fact, in our illustration, it shows the poorest performance in classifying the urban pixels. Hence we proceed for the ACO2/PSO classification to improve the image classified by BBO.

ACO2/PSO Classifier Layer: This layer is responsible for application of the ACO2/PSO classifier on the remainder of the clusters, hence, constituting a hybrid classification.

(i) **Graph Pattern Similarity Analysis Application Layer:** This layer analyzes the data distribution graph and selects the clusters which can be appropriately classified by the ACO2/PSO technique. This is done by making a similarity analysis of the graph pattern from the data distribution graph in order to deduce which of the other features will show similar efficiency as the one deduced from the graph pattern similarity analysis layer of the image analyzer layer. The clusters which predominantly contain these features will be the ones that will be more efficiently classified by the ACO2/PSO classifier. In our illustration, we know from the data distribution graph plotted in the Image Analyzer Layer that our hybrid algorithm will show similar efficiency (i.e., maximum) for the urban pixels too as for the water pixels since their graph pattern in the NIR and MIR bands is similar, therefore, we then apply ACO2/PSO Technique on the rest of the image. In our case, BBO performed better than ACO2/PSO only on the water pixels and hence we applied BBO on the *k*th cluster. Although BBO gave good performance on vegetation pixels too (since in NIR and MIR bands these are the 2 features that are best viewed), however, the efficiency was no better than ACO2/PSO classification and hence we did not apply BBO on the vegetation pixels.

(ii) **Training set in .arff format Generator Layer:** This layer applies the ACO2/PSO classifier by taking the training set as input. In our case, in order to apply ACO2/PSO Classifier, we take the training set for the 7-band Alwar Image in .arff format as input to generate the classification rules from it.

(iii) **Classification Rules Extractor Layer:** This layer extracts the classification rules and applies these rules for classification of those clusters of the image on which ACO2/PSO classifier is chosen to be applied. In our case, we use the Hybrid ACO2/PSO open source Tool and then apply the extracted classification rules on each of the remainder clusters of the image.

(iv) **Pixel Validator Layer:** This layer checks for pixel validation on each pixel in the cluster and thus obtain a refined classification of the image. The above procedure of applying ACO2/PSO generated rules for classifying the image is as described in Section 4.2.

Color Codes Assignment Layer: Finally in this layer, color codes are assigned for each pixel of the image corresponding to the classified feature of each pixel. The color codes for water, urban, vegetation, rocky and barren features are assigned as blue, red, green, yellow and black color respectively.

5.2. Functional Architecture of the hybrid classifier

Therefore, the working of our proposed hybrid classifier can be summarized in the mathematical form as equation shown in Fig. 9.where the Universal Habitat contains the rough set classified equivalence classes and the feature habitat consists of the expert generated training set of the original Alwar Image in 7-bands. And, z, input cluster of the 7-band image (z ranges from 1 to 20); j, feature under consideration (j ranges from 1 to 6, $j=1$ for unclassified feature); x, feature most efficiently classified by BBO; k, cluster which shows the maximum classification efficiency since it predominantly shows the presence of feature 'x'

The equation present in Fig. 9 can be explained as follows:

For $z=1$ to 20 /* here, $n=20$, however 'n' may be varied as desired */

If ($z=k$)

/* Apply the BBO Optimizer */

i. For each ith band where 'i' ranges from 1–7, calculate the difference in the standard deviation of the ith band of the Universal Habitat and the ith band of the Feature Habitat containing the expert generated training set of the image.

ii. If this difference is the minimum for the feature 'j' and also less than the pre-specified threshold value of, $-1 < t < +1$, then equivalence class 'z' is classified as the feature 'j' else $j=1$(unclassified).

/* This means that the feature 'x' most efficiently classified by BBO is predominantly present in the cluster 'k' and hence $j=x$ for the cluster $z=k$. */

iii. The process is repeated for each equivalence class until there is no equivalence class left in the universal habitat and the whole process is iterated till there is no unclassified equivalence class left.

End If

If ($z \neq k$)

/* Apply the ACO2/PSO Optimizer */

.i. Training set for the 7-Band Alwar image in .arff format is used as input to generate rules from it using the open source tool for each class of training case and on each iteration, we add continuous terms till the best discovered rule is found.

.ii. The classification rules are then applied on the remainder of the clusters checking for pixel validation on each of them.

End If

End For

Hence, we obtain a more refined classified image with an improved Kappa Coefficient (Kappa Coefficient is a way of measuring the efficiency of a classification algorithm) which is much better than the Kappa Coefficient we get when we apply the original BBO Algorithm on the 7-band image.

This in turn leads us to the improved flexible Hybrid version of the BBO Algorithm for Satellite Image Classification which will classify the particular feature chosen by the band used in the unsupervised classification, most efficiently, which is in turn based on

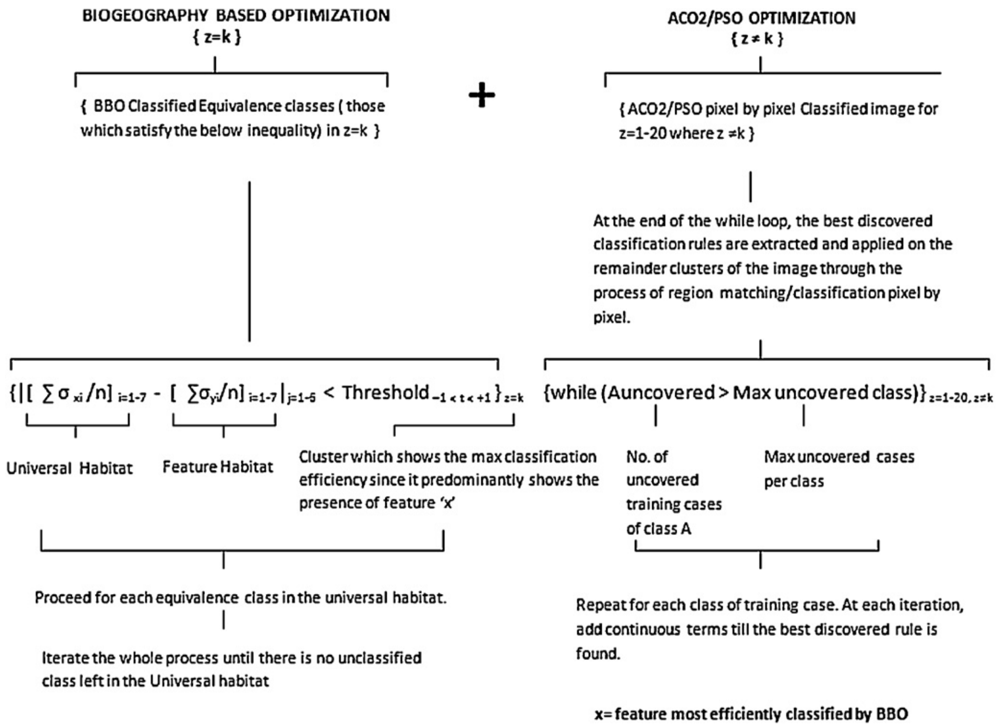


Fig. 9. Mathematical formulation of the ACO2/PSO/BBO Algorithm.

the expert knowledge [11] and the band information contained in the training set of the particular area. Thus, we have efficiently exploited the properties of the BBO Technique to adapt itself to a more focused classification [11] which upon integrating with the ACO2/PSO Technique makes an advanced classifier. Hence, we have obtained a hybrid algorithm which combines the characteristics of embedding expert knowledge for a more flexible classification [11] and the characteristic advantages of a hybrid classification for a more efficient and refined classification, thus resulting in what we term as the hybrid bio-inspired pattern analysis based intelligent classifier. The detailed functional architecture of this Hybrid ACO2/PSO/BBO based classifier is illustrated by means of a flowchart in Fig. 10.

6. Implementation results of the proposed hybrid classifier

Fig. 11 shows the 16th cluster of the Alwar Image after running multiple iterations of Biogeography based land cover feature extraction algorithm on it with the threshold value taken as −1 to +1. The yellow color represents rocky area, green color represents vegetation area, black color represents barren area and red color represents the urban area. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.) The white color in the image represents the unclassified pixels of image in the first iteration. It is analogous to those species that are not absorbed in any feature habitat and are put back to the Universal Habitat. These species represent the unclassified pixels that can be further classified in the next iteration of algorithm.

By taking the threshold value from −1 to +1, we see that no water body (represented in blue color) is extracted in the first iteration (Fig. 11(a)) and water pixels remain unclassified (represented in white color). With the increasing iteration (Fig. 11(b)–(e)), water body is clearly extracted at the end of fifth iteration [20].

After five iterations of the algorithm, each species is absorbed in the corresponding habitat and the Universal Habitat becomes empty. This terminates the biogeography based feature extraction algorithm on the 16th cluster.

Table 1
Error matrix when only BBO is applied.

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	127	9	0	0	2	138
Urban	0	88	1	0	32	121
Rocky	6	2	176	1	17	202
Water	0	0	3	69	0	72
Barren	17	91	20	0	119	247
Total	150	190	200	70	170	780

Kappa Coefficient = 0.6715.

Based on the results obtained on applying the BBO Algorithm to the 7-band of image of Alwar Region for land cover feature extraction which correspond to the image analysis layer of our proposed architecture, we observe that we are able to classify water pixels with the highest efficiency, i.e., 99% efficiency and these are the pixels best viewed in the NIR and MIR bands in the BBO Technique and hence, in the selective cluster identification layer, we identify that we should apply BBO Technique on the 16th cluster of the satellite image of Alwar Region ($z = 16$) since this is the cluster which gives the maximum classification efficiency since it predominantly shows the presence of water body in the Alwar Image and hence we proceed with the selective cluster BBO classification applier layer on this cluster. Next, we proceed with the BBO feature extraction inefficiency layer wherein we observe that, BBO shows poor efficiency, in fact the poorest, in classifying the urban pixels as shown in Fig. 12. Here the encircled region in the BBO Classified Image shows that BBO wrongly classifies the urban pixels as barren ones which is also reflected from Table 1 where BBO classifies 91 urban pixels wrongly out of 190 total urban pixels.

Therefore, based on the observations done in the graph pattern similarity analysis layer of the image analysis layer, in order to classify the urban pixels efficiently, we then apply ACO2/PSO Technique [5] on the remainder of the clusters of the image ($z \neq 16$) by taking the training set for the 7-band Alwar Image in .arff format as input to

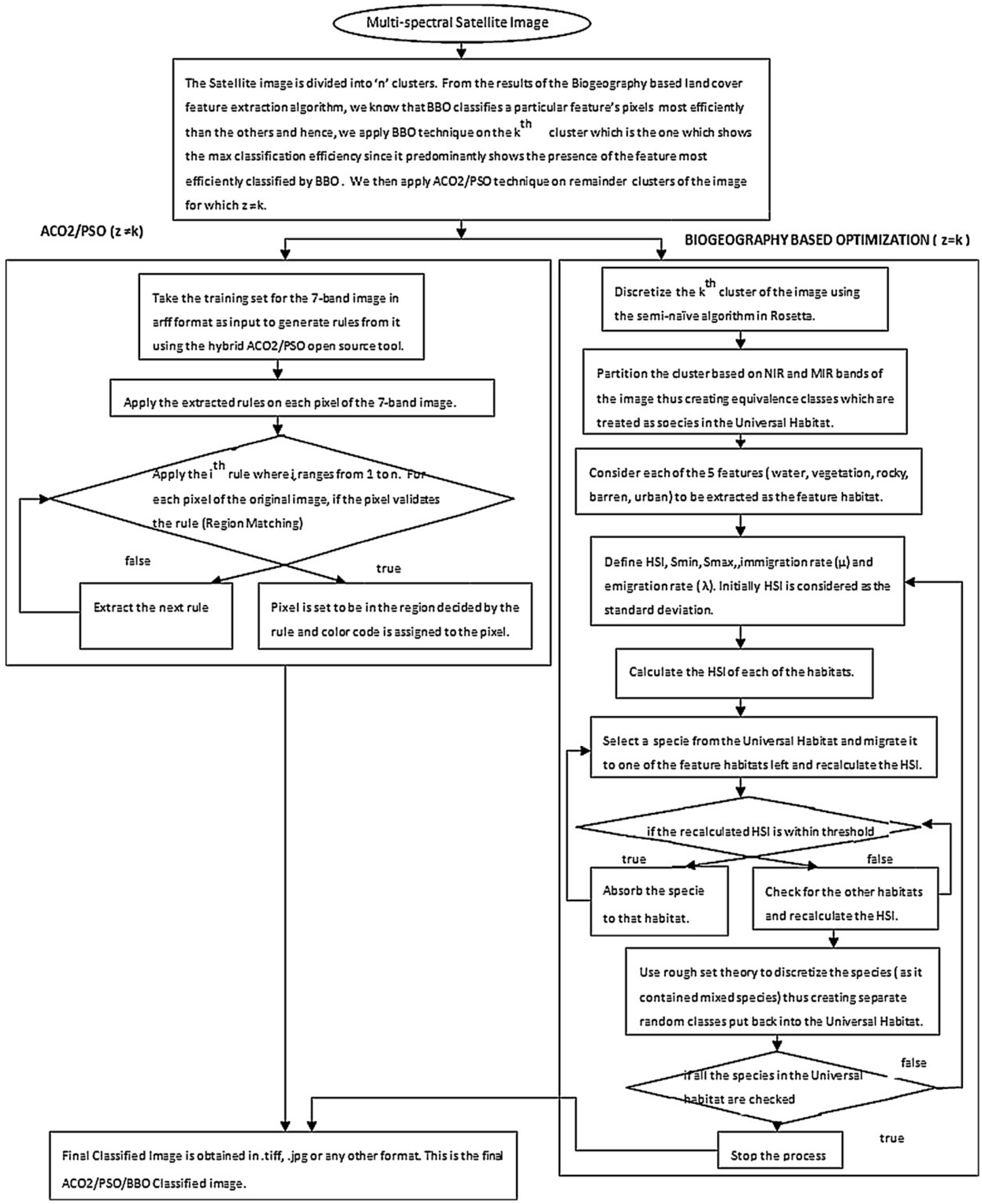


Fig. 10. Detailed Internal architecture of the Hybrid ACO2/PSO/BBO Classifier.

generate rules from it using the Open Source Tool and then applying them on the remainder of the clusters checking for pixel validation for each pixel in the cluster and thus obtain a more refined classification of the image with an improved Kappa Coefficient of 0.9818 which is much better than the Kappa Coefficient of 0.6715 [20] we get, when we apply the original BBO Algorithm on the 7-band image. This in turn leads us to the improved Hybrid version of the original biogeography based land cover feature extraction algorithm where both the urban and the water features are classified with the highest efficiency, i.e., almost 100% with no omission errors followed by rocky with only 1 omission error (column wise error) and thereafter barren and vegetation features, respectively. After applying the proposed algorithm to the 7-band of Alwar Image, the classified image is obtained in Fig. 13. From the figure, it is clearly shown that our proposed ACO2/PSO/BBO Classifier is able to correctly classify the encircled region as urban which was wrongly classified by the simple BBO Classifier. The yellow, black, blue, green, red color represents rocky, barren, water, vegetation, urban region respectively. (For interpretation of the references to color in this figure legend, the reader is referred to the web version

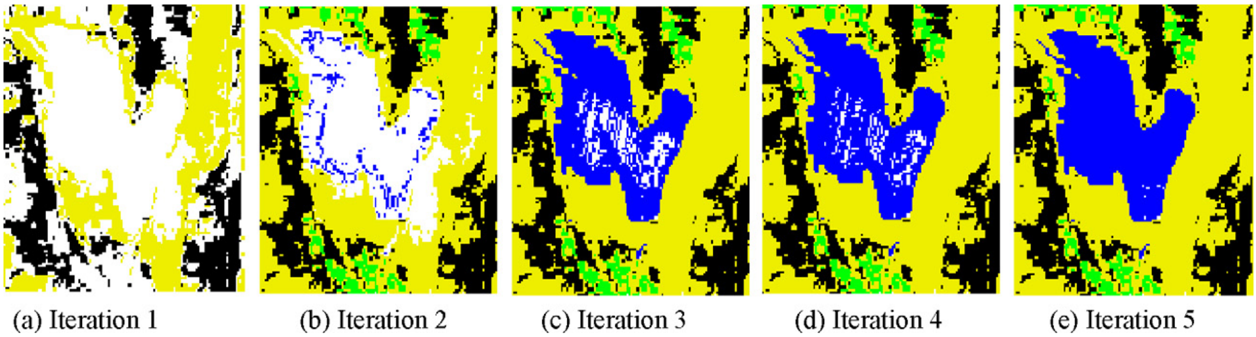


Fig. 11. Water body extraction (in blue color) from the Alwar Image with each iteration. White color portion represents the unclassified image. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

of this article.) As the threshold limit of HSI matching is lowered, the species do not get absorbed in the feature habitat and return to Universal Habitat. Those species are further discretized and classified in next iterations (generation).

From Figs. 12 and 13, it is evident that the Hybrid ACO2/PSO/BBO Technique produces a more refined image as compared to the BBO Classified Image.

7. Accuracy assessment of the proposed algorithm

A classification is not complete until its accuracy is assessed. Accuracy assessment is a general term for comparing the classification to geographical data that are assumed to be true, in order to determine the accuracy of the classification process. Usually, the assumed-true data are derived from ground truth data. It is usually not practical to ground truth or otherwise test every pixel of a classified image. Therefore, a set of reference pixels is usually used. Reference pixels are points on the classified image for which actual features are (or will be) known. The reference pixels are randomly selected. Accuracy assessment is an important step in the classification process. The goal is to quantitatively determine how effectively pixels were grouped into the correct feature classes in the area under investigation.

7.1. Error matrix

Now we proceed to calculate the classification accuracy of our proposed algorithm using the classification *error matrix*. Error matrices compare, on category-by category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. We took 150 vegetation pixels, 190 urban pixels, 200 rocky pixels, 70 water pixels, 170 barren pixels from the training set and the error matrix obtained is shown in Table 3. In Table 3, the training set pixels that are classified into proper land cover categories are located along the major diagonal of the error matrix (running from upper left to lower right). All the non diagonal elements represent the error of omission and commission. Omission errors correspond to the non diagonal column elements. Commission errors are represented by non diagonal row elements.

The error matrix's interpretation along column suggests how many pixels are classified correctly by the algorithm. For example, in the first column, out of total 150 vegetation pixels, 142 pixels were correctly classified as vegetation, 2 vegetation pixels were classified wrongly as barren, and 5 as urban, a total of only 7 misclassified pixels. However, from the simple BBO Error matrix given in Table 1, it is found that 17 vegetation pixels were wrongly classified as barren pixels, and 6 as water, which amounts to a total of 23 wrongly classified pixels. Also, the diagonal elements

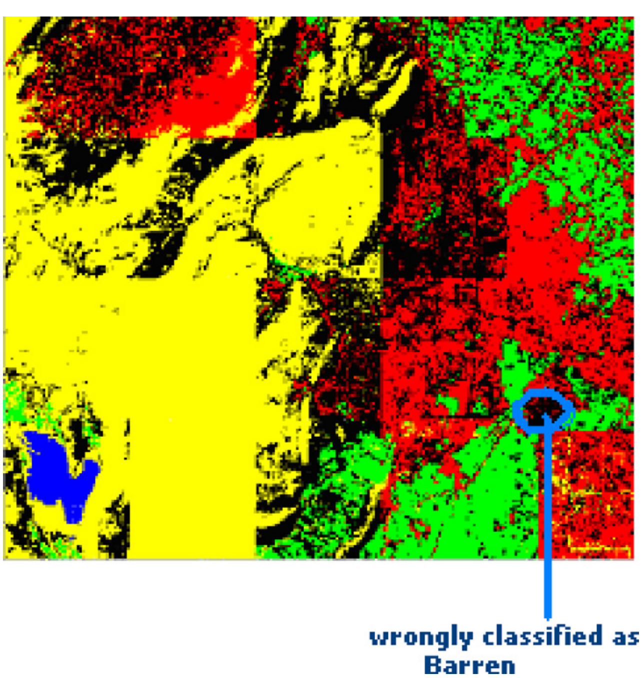


Fig. 12. Classified image after applying BBO.

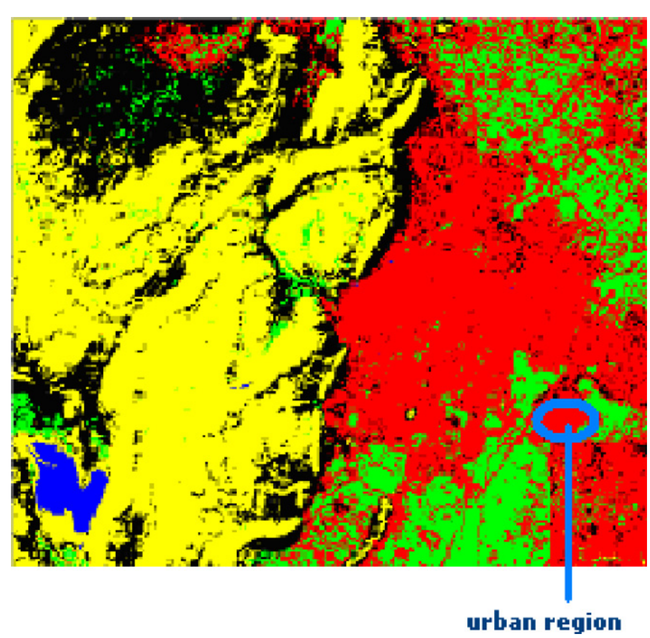


Fig. 13. Hybrid ACO2/PSO/BBO Classified Image.

Table 2
Error matrix when Hybrid ACO2/BBO Technique is applied.

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	142	0	0	0	0	142
Urban	5	190	0	0	0	195
Rocky	0	0	198	0	3	201
Water	0	0	0	70	0	70
Barren	2	0	1	0	163	166
Total	149	190	199	70	166	774

Kappa Coefficient = 0.96699.

Table 3
Error matrix when Hybrid ACO2/PSO/BBO Technique is applied.

	Vegetation	Urban	Rocky	Water	Barren	Total
Vegetation	145	0	0	0	0	145
Urban	2	188	0	0	9	199
Rocky	0	0	194	0	0	194
Water	1	0	3	70	0	74
Barren	1	2	2	0	157	162
Total	149	190	199	70	166	774

Kappa Coefficient = 0.98182.

Table 4
Producer's accuracy.

Feature	Accuracy calculation	Producer's accuracy (%)
Vegetation	142/150	95
Urban	190/190	100
Rocky	198/200	99
Water	70/70	100
Barren	163/170	96

(diagonal elements indicate the no. of correctly classified pixels in that category) of Table 3 (Hybrid ACO2/PSO/BBO Classifier) vary drastically when compared to those in Table 1 (simple BBO Classifier). For example, in Table 3, whole 190 out of 190 pixels were correctly classified as urban pixels whereas simple BBO Classifier in Table 1 could only classify 88 pixels correctly as urban pixels and it classified 91 pixels wrongly as barren ones. This indicates a drastic improvement in the accuracy of the results obtained by BBO through the use of the Hybrid Technique which is able to classify, in particular the urban and the water pixels, with almost 100% efficiency (with no omission errors), which was what was also reflected earlier, from the data distribution graph plotted. The error matrix for the Hybrid ACO2/BBO Classifier is also calculated and presented in Table 2 from which it is reflected that the our proposed classifier is an improvement over the Hybrid ACO2/BBO based intelligent classifier too which is a hybrid of the ACO2 and the BBO based land cover feature extraction algorithm.

7.2. *Producer's accuracy*

Several other descriptive measures can be obtained from error matrix. The accuracy of individual category can be calculated by dividing the number of correctly classified pixels in each category by either the total number of pixels in corresponding row or column. Producer's accuracies (as shown in Table 4) result from dividing the number of correctly classified pixels in each category (on the major diagonal) by the number of training set pixels used for that category (the column total). This figure indicates how well the training pixels of a given cover type are classified. The results in Table 4 show that water and urban pixels have been extracted perfectly followed by rocky pixels which show 99% efficiency.

Table 5
User's accuracy.

Feature	Accuracy calculation	User's accuracy (%)
Vegetation	142/142	100
Urban	190/195	97
Rocky	198/201	98.5
Water	70/70	100
Barren	163/166	98

Table 6
A comparison of the Kappa Coefficients of the Hybrid ACO2/PSO/BBO Classifier with the traditional classifiers.

Minimum Distance Classifier (MDC)	Maximum likelihood Classifier (MLC)	Biogeography Based Optimization (BBO)	Hybrid ACO2/PSO-BBO Classifier
0.7364	0.7525	0.6715	0.98182

7.3. *User's accuracy*

User's accuracies (as shown in Table 5) is computed by dividing the number of correctly classified pixels in each category by the total number of pixels that were classified in that category (the row total). This figure is a measure of commission error and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

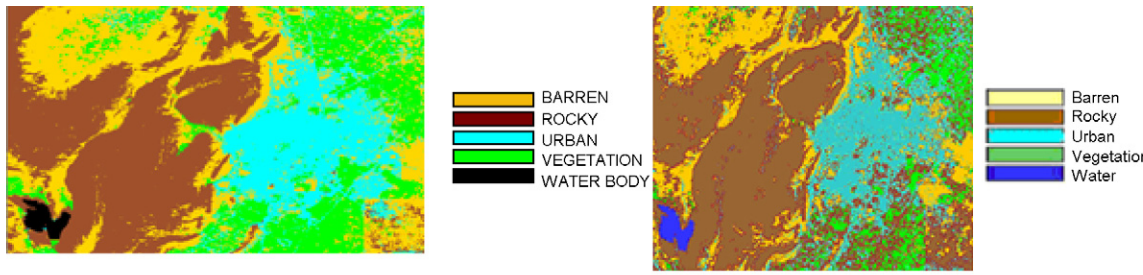
8. Classification comparison of Hybrid ACO2/PSO/BBO Classifier with the traditional Probabilistic Classifiers

Fig. 14 compares the Hybrid ACO2/PSO/BBO Technique with the MDC and MLC. A comparison of the Kappa Coefficients of the Hybrid ACO2/PSO/BBO Classifier with the Traditional Classifiers is given in Table 6.

9. Classification results of other Soft Computing Techniques used for land cover feature extraction

In this section, we present the results of classification of all the Soft Computing Classifiers developed, on the 7-band cartoset satellite image of Alwar Region in Rajasthan, India. The 7-band Alwar Image is a benchmark image for testing the performance of a bio-inspired classifier on multi-spectral satellite images since this image is a complete image in the sense that it contains all the land cover features that we need to extract and hence land cover feature extraction results are demonstrated and compared using this image as the standard image. (This multi-spectral satellite image is provided courtesy of DTRL, Defense and Research Development Organization, New Delhi, India.)

From the discussion in Section 8, it is evident that the Hybrid ACO2/PSO/BBO is a much efficient classifier as compared to the traditional Probabilistic Classifiers such as the MDMC and MLC. However, this Hybrid ACO2/PSO/BBO Technique also produces comparable results with the other recent Soft Computing Classifiers. The Satellite Image Classification results of the other recent Soft Computing Classifiers when applied on the Alwar Region are shown below. Fig. 15(a) shows the fuzzy classification of Alwar Region which has a Kappa-Coefficient of 0.9134 [19]. Fig. 15(b) presents the results of an integrated Rough-Fuzzy Tie-Up Approach which has a Kappa Coefficient of 0.9700 [19]. Fig. 15(c) applies the cAntMiner Algorithm on the Alwar Region which has a Kappa Coefficient of 0.964 [5]. Fig. 15(d) shows the result of applying the Hybrid ACO-BBO Technique on the Alwar Image which has a Kappa-Coefficient of 0.96699. Fig. 15(e) applies the Hybrid ACO2/PSO Classifier which has a Kappa Coefficient of 0.975 [5]. Fig. 15(f)



(a) Minimum Distance Classifier (Kappa Coefficient=0.7364)

(b) Maximum Likelihood Classifier (Kappa Coefficient=0.7525)

Fig. 14. A comparison with the Traditional Probabilistic Classifiers.

presents the results of the Semantic Web Based Classifier on the image with a Kappa Coefficient of 0.9881[12].

It is also evident from the classification results of the various classifiers that the swarm intelligence based classifiers namely cAntMiner, Hybrid ACO-BBO and Hybrid ACO2/PSO are more efficient classifiers as compared to the human-mind modeled Soft Computing Classifiers such as the Rough sets, Fuzzy sets and Rough-Fuzzy Tie-Up Classifiers. This can be verified by considering a portion of the Alwar Image classified by different image classification techniques. The portion that we consider for verifying the correctness of the above results is the area which is encircled. We have chosen this area since we already know that this area is an urban region from the fact that our proposed classifier classified this region with 100% producer's accuracy and 0% omission error which means that our proposed classifier has classified all the urban pixels correctly and no barren pixels have been misclassified as urban pixels and vice versa as is reflected from Table 3.

Now, we see the results of classification of each of the classifiers on the encircled region which we preliminarily know that it is an urban region. From Fig. 15(a) and (b), it is reflected that the Fuzzy set and the Rough-Fuzzy Tie-Up Classifiers are not able to extract the urban land cover feature properly since the Fuzzy sets classifier has classified the encircled region as barren which is actually an urban region. Similar is the case with the Rough-Fuzzy Tie-Up

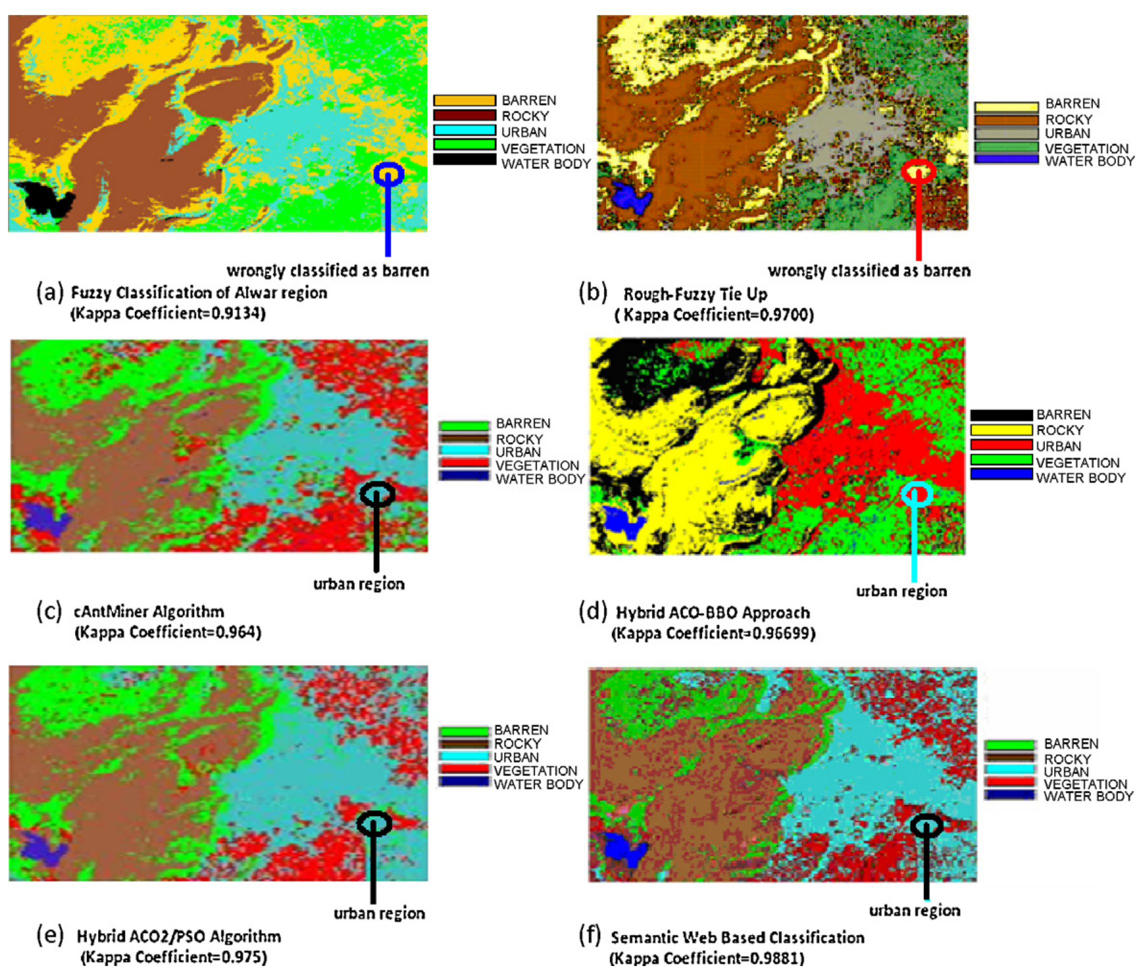


Fig. 15. Classified images of Alwar Region after applying various Soft Computing Techniques.

Table 7
Kappa Coefficient (k) of Soft Computing Classifiers vs Probabilistic Classifiers.

Minimum Distance Mean Classifier (MDMC)	Maximum Likeli-hood Classifier (MLC)	Fuzzy set	Rough-Fuzzy Tie up	cAnt-Miner	Hybrid ACO2/PSO	Semantic Web Based Classifier	Biogeo-graphy Based Classifier	Hybrid ACO-BBO Classifier	Hybrid ACO2/PSO/BBO Classifier
0.7364	0.7525	0.9134	0.9700	0.964	0.975	0.9881	0.6715	0.96699	0.98182

(Probabilistic Classifiers) (Soft Computing Classifiers)
Technology Growth

classifier too. Now, we proceed to see the swarm intelligence based classifiers results. From Fig. 15(c) which shows the results of the cAntMiner classification, we see that the encircled region has been correctly classified as urban region. Fig. 15(d) and (e) which show the results of Hybrid ACO-BBO and Hybrid ACO2/PSO classification also produce a correct classification of the encircled region. Fig. 15(f) is not a swarm intelligence based classifier (the semantic we based computing emphasizes on the structured organization of natural and artificial systems similar to that in granular computing [3] and is added as a special instance of a Soft Computing Classifier), however, correctly classifies the encircled region. Hence, from the above classification results and experimentation on these results, we conclude that the swarm intelligence based classifiers are better classifiers as compared to the other human mind model based classifiers such as the Rough sets and Fuzzy sets.

Table 7 compares the Kappa Coefficients of the Soft Computing Classifiers vs the Traditional Probabilistic Classifiers. From Table 7, it is clearly reflected that Soft Computing Classifiers are much more refined and efficient than the Probabilistic Classifiers. It is also reflected that the swarm intelligence based classifiers are more efficient than the human mind model classifiers namely, Rough sets and Fuzzy sets. Simple BBO Classifier is not a completely swarm based classifier and its results are preliminary in nature since the approach is primarily developed for the purpose of demonstrating the application of BBO to a Satellite Image Classification problem [20] by modifying the original BBO parameters and hence do not produce comparable results with the other swarm intelligence based classifiers demonstrated in Fig. 15 (its Kappa Coefficient being 0.6715). Also, it is reflected from the results of classification that our proposed hybrid classifier is the best known classifier from amongst all the present Soft Computing Classifiers, its performance being only slightly less than the Semantic Web Based Classifier.

10. Conclusion and future scope

At present, it is not possible to state which technique is best for all situations as the characteristic of each image and the circumstances for each study vary so greatly. Therefore, it is essential that each analyst understand the alternative strategies for terrain understanding (here, land cover feature extraction problem) so that he or she may be prepared to select the most appropriate technique for the feature extraction task in hand. To this end, the concept of hybrid bio-inspired computational intelligence which can prove to be an optimized approach of feature extraction from satellite multi-spectral images, has been introduced in this category. These techniques with lower cost and higher degree of classification accuracy, will be able to replace high resolution high cost satellite imageries.

The paper is positioned in the integration and the adaptation phase of the bio-inspired computational models of problem optimization. Since the problem in hand is the land cover feature extraction (or the satellite image classification in remote sensing terms) problem, our focus here was the development of adaptive optimization models from the terrain perspective. In the paper, we analyzed the performance of the existing bio-inspired computational intelligence techniques and developed mathematical model of hybrid algorithm with improved optimization capabilities, and an adaptive framework for the land cover feature extraction problem at hand. To this end, we have made a detailed study and behavioral analysis of the optimization techniques by means of mathematical models governing their performance and hence proposed to refine their results on terrain understanding applications here, land cover feature extraction by extending their basic functionalities by a hybridization of those nature inspired techniques. The Landcover Classification is taken as a case study. It is perceived, from this research, that Kappa coefficient, a well founded metric for assessing the accuracy of classification in remote sensing community, may be used for comparative study of the results from soft computing methods.

The future scope of the research includes proposing certain modification to the algorithm so that the Kappa Coefficient can be improved further. Also, an unsupervised version of the Biogeography based Satellite Image Classification may be explored by including the context information of clusters used in generation of species granules and then evaluating the HSI. The current system is implemented using a simple heuristic technique namely standard deviation; the system performance can be increased by using other heuristic functions. Also, a generic tool representing the generalized system architecture of our hybrid intelligent classifier is under development which takes as input a multi-spectral satellite image and applies the hybrid classifier based on selective cluster identification and inefficiency analysis thus outputting the extracted features from the image.

Further, in this work, we have used rough set theory to obtain the elementary classes (or species). Since the band used in partitioning affect the final results, the rough set theory can be replaced by other swarm based clustering techniques, in this way the proposed feature extraction algorithm will become independent of the bands and hence, the algorithm will become completely 'swarm based'.

The results presented are preliminary and there is a lot of scope for improvement to develop this algorithm as an ideal classifier. In future, the algorithm efficiency can be further improved by lowering the threshold value used in BBO Algorithm thus leading to more iterations and refined results. Also, we can further divide the image into more clusters so that a more accurate comparison can be made and the decision about which of the two techniques to be applied on the particular cluster, can be further streamlined. The system performance can be further increased by using better unsupervised classifications and better training sets.

Acknowledgements

We acknowledge the efforts undertaken by Dan Simon (Professor, Department of Electrical Engineering, Cleveland State

University, Ohio and Senior Member, IEEE), the pioneering author and the founder of the core technique of Biogeography Based Optimization (BBO), in reviewing this paper. His comments and suggestions were invaluable in bringing this paper from its original version to the current form.

This paper has been a dedicated effort towards the development of highly autonomous artificial intelligence. This work has been carried out at Defense and Terrain Research Lab (DTRL), Defense and Research Development Organization (DRDO), New Delhi, India which provided us the invaluable satellite data for the experimental study. The proposed work is carried out on the 7-band cartoset satellite image of the Alwar area in Rajasthan, India which is a standard input multi-spectral satellite image for all the work carried out at DRDO and has now become a benchmark image for testing the performances of various Soft Computing Classifiers for remote sensing applications.

References

[1] Agnar Aamodt, Enric Plaza, Case based reasoning: foundational issues, methodological variations, and system approaches, Artificial Intelligence Communications (IOS Press) 7 (1) (1994) 39–59.
[2] Antonio Plaza, C.-I. Chang, Computer architectures for remote sensing overview and case study, in: Antonio Plaza, C.-I. Chang (Eds.), High Performance Computing in Remote Sensing, Chapman & Hall/CRC Press, Computer & Information Science Series, 2007, pp. 9–41.
[3] A. Bargiela, W. Pedrycz, Granular Computing, Kluwer, Boston, 2003.
[4] Dorigo, T. Stuetzle, Ant Colony Optimization, MIT Press, 2004.
[5] Shelly Bansal, Daya Gupta, V.K. Panchal, Shashi Kumar, Remote sensing image classification by improved swarm inspired techniques, in: International Conference on Artificial Intelligence and Pattern Recognition (AIPR-09), Orlando, FL, USA, 2009.
[6] J. Hand, Construction and Assessment of Classification Rules, Wiley, 1997.
[7] Holden, A.A. Freitas, A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data, in: IEEE Swarm Intelligence Symposium (SIS-05), 2005, pp. 100–107.
[8] Holden, A.A. Freitas, Hierarchical classification of Gprotein-coupled receptors with a PSO/ACO algorithm, in: Proc. IEEE Swarm Intelligence Symposium (SIS-06), 2006, pp. 77–84.
[9] Fernando E.B. Otero, Alex A. Freitas, Colin G. Johnson, cAnt-Miner: an ant colony classification algorithm to cope with continuous attributes, Springer-Verlag, Berlin, Heidelberg, 2008.
[10] Lavika Goel, V.K. Panchal, Daya Gupta, Rajiv Bhola, Hybrid ACO-BBO approach for predicting the deployment Strategies of enemy troops in a military terrain application, in: 4th International Multi-Conference on Intelligent Systems & Nanotechnology (IISN-2010), 2010, pp. 118–122.
[11] Lavika Goel, V.K. Panchal, Daya Gupta, Embedding expert knowledge to hybrid bio-inspired techniques—an adaptive strategy towards focused land cover feature extraction, International Journal of Computer Science & Information Security 8 (2) (2010) 244–253, ISSN: 1947-5500.
[12] Sonal Kumar, Daya Gupta, V.K. Panchal, Shashi Kumar, Enabling web services for classification of satellite images, in: International Conference on Semantic Web and Web Services (SWWS'09), Orlando, FL, USA, 2009.
[13] T.M. Lillesand, R.W. Kiefer, Jonathan Chipman, Remote Sensing and Image Interpretation, 6th edition, Wiley, 2008.
[14] W. Long III, S. Srihann, Land cover classification of SSC image: unsupervised and supervised classification using ERDAS Imagine, in: Proceedings of Geoscience and Remote Sensing Symposium, Unsupervised and Supervised Classifications (IGARSS'04), vol. 4, 2004, pp. 20–24.
[15] A. Ǿhrn, J. Komorowski, ROSETTA "A rough set tool kit for analysis of data", in: Proc. 3rd International Joint Conference on information Sciences, Durham, NC, 1997.
[16] S.N. Omkar, Kumar Manoj, Muley Dipti, Urban satellite image classification using biologically inspired techniques, Proceedings of IEEE (2007).
[17] S. Parpinelli, H.S. Lopes, A.A. Freitas, Data mining with an Ant Colony Optimization Algorithm, IEEE Transactions On Evolutionary Computation 6 (4) (2002) 321–332 (special issue on Ant Colony algorithms).
[18] Bratton, J. Kennedy, Defining a standard for particle swarm optimization, in: Proceedings of the 2007 IEEE Swarm Intelligence Symposium, Honolulu, Hawaii, USA, 2007.
[19] V.K. Panchal, Naresh Singhal, Shashi Kumar, Sonam Bhakna, Rough-fuzzy sets tie-up for geospatial information, in: ISRO's International Conference on Emerging Scenario in Space Technology & Applications (ESSTA' 2008), vol. 1, Chennai, 2008.
[20] V.K. Panchal, Parminder Singh, Navdeep Kaur, Harish Kundra, Biogeography based satellite image classification, International Journal of Computer Science and Information Security (IJCSIS) 6 (2) (2009) 269–274.
[21] Z. Pawlak, Rough set theory and its applications to data analysis, Cybernetics and Systems 29 (7) (1998) 661–688.
[22] Z. Pawlak, Rough sets, International Journal of Computer and Information Science 11 (1982) 341–356.
[23] Z. Pawlak, Rough Set: Theoretical Aspects of Reasoning about Data, Kluwer Academic Publishers, 1991.
[24] L. Polkowski, Rough Sets, Mathematical Foundations, Physics-Verlag, New York, 2002.
[25] D. Simon, Biogeography-based optimization, IEEE Transactions on Evolutionary Computation (IEEE Computer Society Press) 12 (6) (2008) 702–713.
[26] D. Simon, D. Du, M. Ergezer, Biogeography-based optimization combined with evolutionary strategy and immigration refusal, in: Proceedings of IEEE International Conference on System, Man and Cybernetics, San Antonio, USA, 2009.
[27] D. Simon, A probabilistic analysis of a simplified biogeography-based optimization algorithm, Evolutionary Computation 19 (2) (2011) 167–188.
[28] Open Source Tool for Hybrid ACO2/PSO rule extraction. <http://www.PSOACO2.sourceforgenet.com/>.

Lavika Goel is a research scholar in the Department of Computer Engineering at Delhi Technological University (DTU), Delhi, India. She worked at Oracle India Private Ltd. for six months through campus recruitment and then left it to pursue research in the field of Nature Inspired Computational Intelligence, her interest area. Securing a GATE percentile of 98.7%, she did her M.E. (Master of Engineering with *Distinction*) in Computer Technology & Applications from Delhi Technological University (formerly Delhi College of Engineering), Delhi and B-Tech (with *Honours*) in Computer Science & Engineering from JSS Academy of Technical Education, U.P. Technical University and has always been among the top rank-holders of the university in these courses.

Dr. Daya Gupta is a Professor in the Department of Computer Engineering, Delhi Technological University, Delhi, and is currently the head of the department there. She has done Ph.D. in Computer Engineering from University of Delhi and M.Sc. (Maths), Post M.Sc. Diploma (Computers Sc.) from IIT, Delhi. She is a senior member of IEEE and a life member of CSI. Her research interests are in the field of requirement engineering, network security, ad hoc networks and image classification in remote sensing. She has published several research papers in international journals and conferences and has chaired sessions and delivered invited talks at many national and international conferences.

Dr. V.K. Panchal is the Associate Director at Defense Terrain & Research Lab, Defense and Research Development Organization (DRDO), Delhi. He is an Associate Member of IEEE (Computer Society) and Life Member of Indian Society of Remote Sensing. He is also the Vice-Chairman of the Indian Society of Remote Sensing, Delhi Chapter. He has done Ph.D in Computational Intelligence from Jawarharlal Nehru University, Delhi and M.S. in Information Systems from BITS, Pilani and is currently working as Scientist 'G' at DRDO, Delhi. His research interests are in the synthesis of terrain understanding model based on incomplete information set using bio-inspired intelligence and remote sensing. He has chaired sessions & delivered invited talks at many national & international conferences.