A
Dissertation
On

IMAGE CLASSIFICATION
USING ARTIFICIAL BEE COLONY ALGORITHM

Submitted in partial Fulfilment of the requirement
For the award of Degree of

Master of Technology

In Computer Technology and Application

Submitted By

SRIDEEPA BANERJEE

College Roll No. 15/CTA/2010

Under the esteemed guidance of

Dr. Daya Gupta

Department of Computer Engineering

DELHI TECHNOLOGICAL UNIVERSITY

2010-2012

DELHI TECHNOLOGICAL UNIVERSITY
DELHI- 110042

CERTIFICATE

This is to certify that the work contained in this dissertation entitled "**Image Classification Using Artificial Bee Colony Algorithm**" submitted in the partial fulfilment, for the award for the degree of M.Tech in Computer Technology and Applications at **DELHI TECHNOLOGICAL UNIVERSITY** by **SRIDEEPA BANERJEE, Roll No. 15/CTA/10,** is carried out by her under my supervision. This matter embodied in this project work has not been submitted earlier for the award of any degree or diploma in any university/institution to the best of our knowledge and belief.

Date: ___________

(Dr. DAYA GUPTA)

HOD & Project Guide

Department of Computer Engineering

Delhi Technological University

ACKNOWLEDGEMENT

It is a great pleasure to have the opportunity to extent my heartfelt gratitude to everybody who helped me throughout the course of this project.

First of all, I thank my parents who have always motivated me and given their blessings for all my endeavours.

It is distinct pleasure to express my deep sense of gratitude and indebtedness to my learned supervisor Dr. Daya Gupta, HOD, Department of Computer Engineering, Delhi Technological University, for her invaluable guidance, encouragement and patient review.

I would like to extend my earnest and heartfelt thankfulness to my DRDO supervisor Dr.Vinod Kumar Panchal, Associate Director, DTRL (Defence Terrain Research laboratory), Defence Research Development Organization (DRDO). I am very thankful to him for his timely advice, support, comments, guidance and encouragement throughout the research work.

I take this opportunity to appreciate all my friends at DTU for their support, in particular Akanksha Bharadwaj, Pooja Arora, Teeja Marry Sebastian, Samidha and Indu Sirohi for their constant support, help and making this work duration jovial.

Srideepa Banerjee

College Roll no: 15/CTA/10

M.E. (Computer Technology and Application)

Department of Computer Engineering

Delhi Technological University

Delhi-110042.

ABSTRACT

Remote sensing is the most important provider of the various data sources that are used in GIS. It has been globally used for knowledge elicitation of earth's surface and atmosphere. Land cover mapping, one of the widely used applications of remote sensing is a method for acquiring geo-spatial information from satellite data.

We have attempted here to solve the land cover problem by image classification using one of the newest and most promising Swarm techniques of Artificial Bee Colony optimization.

Artificial Bee Colony algorithm (ABC) is a recent development and has emerged as an efficient global optimization technique. It is inspired from the intelligent foraging behaviour of honey bees. In this paper we propose an implementation of ABC for satellite image classification. Classification plays a very important role in image processing and with the increase in images being acquired and archived, optimal classification tool for different application domains is needed. The objective here is to utilize the bee communication and food search method of information exchange and hence achieve maximum classification accuracy.

We also present a new dimension to analyze the efficiency of classification techniques. It is the feature extraction from heterogeneous regions. It is very important for any method to correctly identify all features in such regions. We have used the Heterogeneity factor of an image to analyse the actual efficiency of our method. Various techniques overlook this for overall classification accuracy. The results produced by artificial bee colony algorithm are compared with the results obtained by other traditional and soft computing techniques to show the effectiveness of our proposed implementation. We have analysed the feature extraction of Heterogeneous regions in an image for different soft computing techniques like

fuzzy, cAntMiner, hybrid ACO-BBO and shown that the accurate classification of such regions is independent of kappa coefficient as the accuracy assessment parameter.

The results produced by artificial bee colony (ABC) algorithm are compared with the results obtained by other techniques like Minimum Distance Classifier (MDC), Maximum Likelihood Classifier (MLC), Biogeography Based Optimization (BBO), and Membrane Computing (MC), hybrid Flower pollination/ Bacterial Foraging (FPAB/BFO) and Fuzzy classifier to show the effectiveness of our proposed implementation.

We go beyond the multispectral regime to the hyperspectral images and their dimensionality reduction problem. The scope of ABC has not been explored much yet but we can utilize the bee optimization in reducing the dimension of hyperspectral images. We have discussed the dimensionality reduction problem with analysis of various existing reduction techniques and propose a theory of Artificial Bee Colony dimension reduction for the same.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

In remote sensing, satellite based sensors are burgeoning as a major facilitator of geo-spatial information providing different manifestations of the terrain. Remote sensing is the most important provider of the various data sources that are used in GIS. The satellite image is one of the main sources for capturing the geo-spatial information [1]. It has been globally used for knowledge elicitation of earth's surface and atmosphere. Image classification is applied to remotely sensed data for land cover mapping of multispectral images of Alwar (Rajasthan) and Saharanpur (UP) region. We have applied here one of the most recently introduced swarm-based algorithm, Artificial Bee Colony. ABC simulates the intelligent social behaviour of a honeybee swarm.

We have made use of ABC's optimization potential and applied it in Remote sensing Image classification field. The proposed algorithm once applied to different dataset gives excellent results. As it works on single unit object thus heterogeneity factor can easily be detected and identified in an image.

In this work, ABC is used for optimal classification of images for mapping the land-usage efficiently. The results produced by ABC algorithm are compared with the results obtained by other Swarm techniques. We go beyond the multispectral regime also study the dimensionality reduction problem of Hyperspectral Images. As a future extension we propose a theory of utilizing Artificial Bee Colony algorithm for solving this problem.

1.1 MOTIVATION

Image classification using artificial bee colony algorithm is proposed in this work. Classification is undoubtedly the most significant digital image processing step in remote sensing. Land cover mapping is an important application area of remote sensing discipline and Classification is the preferential step for producing thematic spatial information from satellite image data. Classification plays a very important role in image processing and with the increase in images being acquired and archived, optimal classification tool for different application domains is needed.

For over the last two decades, Artificial Intelligence (AI) optimization algorithms such as Genetic Algorithm (GA) and Differential Evolution (DE) have been used widely for different optimization problems in many disciplines successfully. GA is derived from natural evolutionary process while DE is a GA-based intuitional algorithm.

Scientists have been trying to model complex optimization problems by employing natural metaphors. This is due to the inefficiency of classical optimization algorithms in solving complex non-linear and large scale combinatorial problems. Swarm Intelligent techniques present a new promising domain in this nature inspired paradigm.

Swarm Intelligence is an innovative Artificial Intelligence technique that was born out of biological insights of the incredible abilities of social insects to solve their simple food/shelter related problems and is it now widely accepted as one of the most efficient optimization technique. Ant Colony Optimization (ACO) and Particle Swarm Optimization have been the most popular swarm techniques in the past. ACO is motivated by the organized behaviour of ants [3] and PSO is inspired from the social behaviours of bird and fish shoals. These AI optimization algorithms are especially preferred when the classical deterministic methods are inadequate because of too many parameters and data sets are not homogenous. These AI tools are effectively used in remote sensing, as well. [4]

As an AI optimization algorithm, Artificial Bee Colony (ABC) was recently proposed [2], which is inspired from intelligent social behaviour of honey bees. ABC tries to model the intelligent social behaviour of honey bees in food foraging. Bees use several mechanisms like waggle dance to optimally locate nectar sources and to search new ones. This makes them a good candidate for developing new intelligent search algorithms.

In remote sensing we have a huge database, provided by experts so we need a technique which efficiently searches all possible cases. Many of the previous classification techniques like ACO, BBO, PSO and their hybrid variants work on a cluster basis and hence not been able to achieve good accuracy. ABC however classifies the image on a pixel basis and overcomes this disadvantage. In this paper we propose an implementation of ABC for satellite image classification.

In an image all the features are usually not totally distinguished, some features are intertwined with other features. In some regions many features coexist together. Due to mixture of feature at some portion of images the concept of heterogeneous regions came into existence. Since our computation technique i.e. Artificial Bee Colony, works on the whole image taking each pixel and hence can correctly identify all features in the heterogeneous portions in image and can classify each portion perfectly.

It’s always a challenge to select, extract most significant and meaningful information from Hyperspectral Imagery. With images being acquired and archived, the present existing techniques are becoming inadequate for analyzing such large volumes of data. Dimension reduction of is very critical for extracting information from these images. There are various techniques for this process, we have analysed them strategically and presented a proposal which uses artificial bee colony for optimal dimension reduction.

1.2 RELATED WORK

With huge amount of digital images being continuously acquired and archived, large volumes of raw data is being generated. Extracting relevant information from these is becoming a challenging work. Classification is an important tool in feature retrieval from large datasets. Research community has been on continuous search for best classification techniques.

The problem of satellite image classification in remote sensing has been solved by using the traditional classical approaches like Parallelopiped Classification [5], Minimum Distance to Mean Classification [5], Maximum Likelihood Classification [5] 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.

Nature inspired computing techniques are a recent trend and were introduced in remote sensing for image classification. The principal constituents of soft computing techniques are fuzzy logic [6], rough set theory [7], neural network theory, and probabilistic reasoning and Swarm Intelligence techniques [8]. However the soft computing techniques like the fuzzy classifier [9], and the rough set classifier [7] were not able to provide good result in case of ambiguity since the aim 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 recently introduced concept of swarm intelligence [7] [10] which come under natural computation. A hybrid BBO/ACO2/PSO was formulized [11]. This technique improves the classification of multi-spectral satellite images but the heterogeneous regions were not classified correctly. Recently membrane computing was implemented [28] which was a pixel wise approach so could identify the heterogeneous region to some extent but its overall accuracy was low and it could accurately classify only water and vegetation region for Alwar dataset.

To overcome the above mentioned difficulties we have proposed an Artificial Bee Colony approach for image classification.

In [4] most popular AI optimization algorithms, PSO, GA, DE and ABC are used for the unsupervised classification of remotely sensed image data and some benchmark data. The results show both the potential of these algorithms against the conventional KM, FCM and superiority of the ABC over other AI tools like GA, PSO, DE [4].

D. Chandrakala and S. Sumathi [12] have proposed an image retrieval system with Artificial Bee Colony optimization algorithm by fusing similarity score based on color and texture features of an image thereby achieving very high classification accuracy and minimum retrieval time.

We have developed Artificial Bee Colony Algorithm for land cover feature recognition. it is a pixel wise approach like membrane computing [28] but it uses less parameters and is able to achieve better accuracy.

1.3 PROBLEM STATEMENT

Image classification can also be seen as an optimization problem. Here out of several features available we have to classify a pixel (or cluster) to its most suitable class. Hence determine the most optimum match.

The aim of this work is to highlight the potential of nature inspired swarm techniques in building efficient optimization models. We propose an Artificial Bee Colony algorithm as a new technique for classifying remote sensing satellite image. Karaboga developed Artificial Bee Colony (ABC) Algorithm by modelling foraging behaviours of honey bees [2].

We justify our Artificial Bee Colony algorithm as an efficient new technique for Land Cover feature extraction for classifying remote sensing satellite image by applying it on two datasets.

We go further and analyse the performance of many existing techniques such as rough set, fuzzy, BBO, Hybrid ACO/PSO, Membrane computing and others with our classifier based on heterogeneous region feature extraction criteria.

We also go beyond the multispectral and extend the scope of Artificial Bee Colony for Dimension Reduction of Hyperspectral Images.

Our problem statement can be proposed as follows-

"To Develop Nature inspired Artificial Bee Colony algorithm for extracting land cover terrain features for both homogeneous and heterogeneous regions"

1.4 SCOPE OF THE WORK

In this project, a new image classification technique in the nature inspired domain has been proposed. The technique is based on population heuristics of Artificial Bee Colony Algorithm.

The algorithm is been applied to 7 band Indian Resourcesat Satellite image of dimension 472 X 546 of Alwar Region in Rajasthan (since it contains a variety of land cover features) and to 6 band Indian Resourcesat Satellite image of dimension 641 X 641 of Saharanpur region in Uttar Pradesh.

The accuracy parameter considered in our work is Kappa Coefficient, which is accepted as a universal standard for assessing classification accuracy. This statistic describes how much better or worse our classification is as compared to a chance alone. It is based on the error matrix which compares on a category by category basis the relationship between reference data and corresponding results of our classifier.

The performance analysis supports our theory when compared to some already implemented other land cover classification techniques.

Broadly the scope of this work can be summarized as:

- To develop ABC classification algorithm for remote sensing satellite image.
- To test our classification algorithm on Alwar dataset for recognising its various land cover features namely water, urban, barren vegetation and rocky.
- To test our classification algorithm on Saharanpur dataset for recognising its various land cover features namely water, urban, barren, dense vegetation, medium vegetation and sparse vegetation.
- Evaluate the classification accuracy using Error matrix and Kappa coefficient.
- To compare the efficiency of our algorithm with other techniques.
- Critical analysis of ABC classification on Heterogeneity factor.
- Analyse the various techniques for Hyperspectral image dimension reduction and give an Artificial Bee Colony proposal for it.

1.5 ORGANIZATION OF THE THESIS

The remainder part of this thesis is organized in the following chapters:

Chapter 2: Remote Sensing satellite Imagery and Image Classification

This section gives an overview of the remote sensing concepts that we have used in our work. It explains the electromagnetic spectrum and multispectral satellite image characteristics. Basic concepts of spectral signatures, reflectance and digital numbers is given. Here we also introduce Image Classification as the preferential step for land cover mapping problem of remote sensing, its concepts and stages.

Chapter 3: Classification Techniques

In this section we move to the different classification techniques. We first start by some traditional supervised and unsupervised techniques and then move to nature inspired algorithms domain. We discuss the weakness of the traditional techniques and reason why these natural computing methods are better. We present taxonomy of the major nature inspired techniques and give some of those techniques implemented in remote sensing classification.

Chapter 4: The New Swarm Technique of Artificial Bee Colony Algorithm

In this chapter we introduce the artificial bee colony algorithm, its biological inspiration and the original algorithm description and finally some strengths of ABC.

Chapter 5: Artificial Bee Colony Classification Algorithm

This is where we introduce our proposed classification technique based on ABC. We present the architecture of our classifier and the algorithm with detailed step by step explanation and flowchart.

Chapter 6: Experiments and Results

The experimental datasets and their characteristics are discussed here. We explain the training dataset of both Alwar and Saharanpur regions. Then show the final classified results of both with the land features extracted. We then show the heterogeneous regions of the dataset and how our classifier has identified each feature from them.

Chapter 7: Evaluation and Analysis

We evaluate and judge the performance of our experimental results. Accuracy assessment and kappa calculations are made. Finally a comparative analysis is made with many other soft computing techniques with heterogeneity factor.

Chapter 8: Application of Artificial Bee Colony on Hyperspectral Images

We present a study of Hyperspectral imaging and its dimensionality problem. Analysis of various techniques along with taxonomy is given. A theory based on Artificial Bee Colony algorithm which can be used for dimension reduction is proposed here.

Chapter 9: Publication from Thesis

This section gives the details of publication from the thesis and the conference details.

Chapter 10: Conclusion and Future Scope

In this section the conclusion of the thesis work and the future scope of the work are presented.

References: This section gives the reference details of the thesis.

Appendix A: Abbreviations

Appendix B: Introduction to Image Processing in MATLAB

Appendix C: Introduction to ERDAS

2. REMOTE SENSING SATELLITE IMAGERY AND IMAGE CLASSIFICATION

In this chapter we understand the fundamentals of remote sensing and image classification. The important concepts related to our work are explained. We discuss the stages of remote sensing and the electromagnetic spectrum. The image characteristics of spectral signatures and digital numbers are explained. Then image classification, its types and steps are explained.

2.1 INTRODUCTION TO REMOTE SENSING

Remote sensing is the most important provider of the various data sources that are used in GIS. Remote sensing is a technique to observe the earth surface or the atmosphere from out of space using satellites (space borne) or from the air using aircrafts (airborne). It records the electromagnetic energy reflected or emitted by the earth's surface. The main concept in remote sensing with multispectral satellite imagery is that different features/objects constituting the land cover reflect electro-magnetic radiations over a range of wavelengths in its own characteristics way according to its chemical composition and physical state. A multi-spectral remote sensing system is operated in a limited number of bands and measures radiations in series of discrete spectral bands. The spectral response is represented by the discrete digital number (DN) spectral signatures of an object may be used for identification much like a fingerprint [1]. There's a lot of uncertainty and ambiguity in categorizing geographic objects based on the geospatial information received from remote sensing satellites.

2.1.1 STAGES OF REMOTE SENSING

• Emission of electromagnetic radiation, or **EMR** (sun/self- emission): energy source which illuminates or provides electromagnetic energy to the target of interest.

• Transmission of energy from the source to the surface of the earth, as well as absorption and scattering: the energy travels from its source to the target, it will come in contact with and interact with the atmosphere it passes through. This interaction may take place a second time as the energy travels from the target to the sensor.

• Interaction of **EMR** with the earth's surface: reflection and emission:

Once the energy makes its way to the target through the atmosphere, it interacts with the target depending on the properties of both the target and the radiation.

• Transmission of energy from the surface to the remote sensor: after the energy has been scattered by, or emitted from the target, we require a sensor (remote - not in contact with the target) to collect and record the electromagnetic radiation.

• Sensor data output: the energy recorded by the sensor has to be transmitted, often in electronic form, to a receiving and processing station where the data are processed into an image (hardcopy and/or digital).

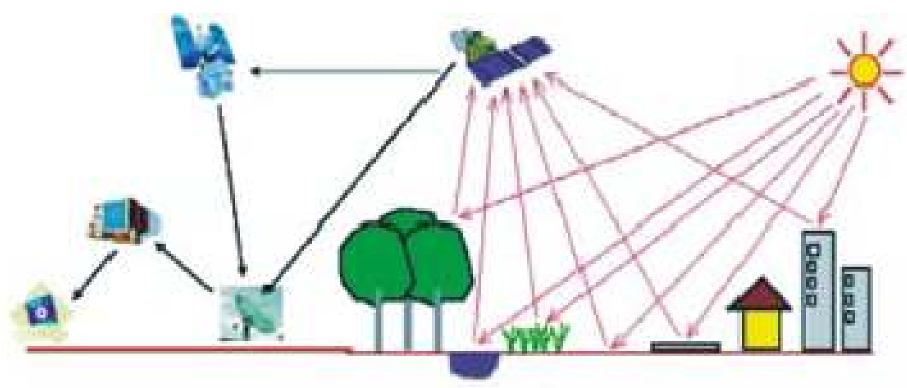


Fig 2.1 Remote Sensing process [13]

2.1.2 ELECTROMAGNETIC RADIATION AND SPECTRUM

Electromagnetic radiation consists of an electrical field (E) which varies in magnitude in a direction perpendicular to the direction in which the radiation is travelling, and a Magnetic field (M) oriented at right angles to the electrical field as depicted in fig 2.2 Both these fields travel at the speed of light (c). Two characteristics of electromagnetic radiation are particularly important for understanding remote sensing. These are the **wavelength and frequency.**

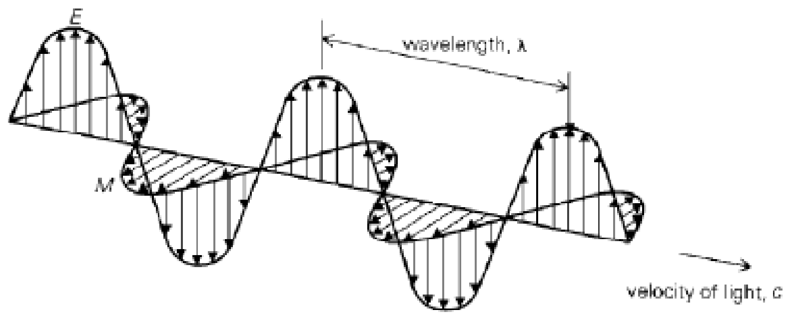


Fig 2.2 Electromagnetic wave [13]

Electromagnetic Spectrum (EMS)

Each photon of the electromagnetic spectrum has a wavelength determined by its energy level. Light and other forms of electromagnetic radiation commonly are described in terms of their wavelengths. EMS ranges from the shorter wavelengths (including gamma and x-rays) to the longer wavelengths (including microwaves and broadcast radio waves). There are several regions of the electromagnetic spectrum which are useful for remote sensing.

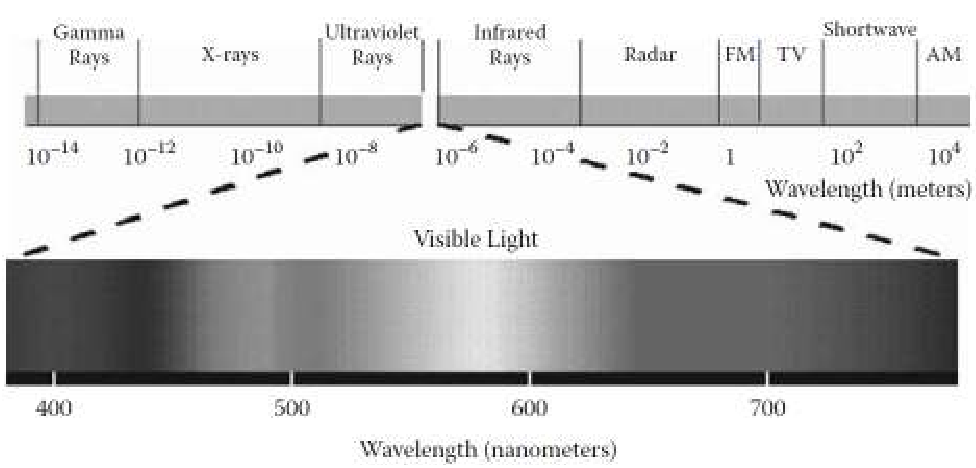


Fig 2.3 Electromagnetic spectrum

2.1.3 CHARACTERISTICS OF REMOTE SENSING IMAGES

Electromagnetic energy may be detected either photographically or electronically. The photographic process uses chemical reactions on the surface of light- sensitive film to detect and record energy variations. A photograph could also be represented and displayed in a digital format by subdividing the image into small equal-sized and shaped areas, called picture elements or pixels, and representing the brightness of each area with a numeric value or digital number.

A. Digital Numbers (DN value)

It is a positive integer value, which represents the relative reflectance or emittance of an object in a digital image. For 8 bit images, the DN or digital number lies in the range of

0-255. In a most generalized way, a digital image is an array of numbers depicting spatial distribution of a certain field parameters (such as reflectivity of EM radiation, emissivity, temperature or some geophysical or topographical elevation. Digital image consists of discrete picture elements called pixels. Associated with each pixel is a number represented as DN (Digital Number) that depicts the average radiance of relatively small area within a scene. The range of DN values being normally 0 to 255. The size of this area effects the reproduction of details within the scene. As the pixel size is reduced more scene detail is preserved in digital representation.

B. Spectral Response Patterns

A spectral response pattern is sometimes called a *signature*. It is a description (often in the form of a graph) of the degree to which energy is reflected in different regions of the spectrum. Most humans are very familiar with spectral response patterns since they are equivalent to the human concept of color.

Finding distinctive spectral response patterns is the key to most procedures for computer-assisted interpretation of remotely sensed imagery. Spectral reflectance, [$\rho(\lambda)$], is the ratio of reflected energy to incident energy as a function of wavelength. Various materials of the earth's surface have different spectral reflectance characteristics. Spectral reflectance is responsible for the color or tone in a photographic image of an object.

The spectral reflectance is dependent on wavelength, it has different values at different wavelengths for a given terrain feature. The reflectance characteristics of the earth's surface features are expressed by spectral reflectance, which is given by:

$$\rho(\lambda) = [E_R(\lambda) / E_I(\lambda)] \times 100$$

Where,

$\rho(\lambda)$ = Spectral reflectance (reflectivity) at a particular wavelength.

$E_R(\lambda)$ = Energy of wavelength reflected from object

$E_I(\lambda)$ = Energy of wavelength incident upon the object

The plot between $\rho(\lambda)$ and λ is called a spectral reflectance curve. This varies with the variation in the chemical composition and physical conditions of the feature, which results in a range of values. The spectral response patterns are averaged to get a generalized form, which is called generalized spectral response pattern for the object concerned.

Spectral signature is a term used for unique spectral response pattern, which is characteristic of a terrain feature. Figure 2.4 shows a typical reflectance curves for three basic types of earth surface features, healthy vegetation, dry bare soil (grey-brown and loamy) and clear lake water.

C. Heterogeneous and Homogeneous regions

In a remote sensing image there are two types of regions based on how the features are distributed. Image features are the different land terrain characteristics like water, vegetation, rocky area or urban and so on. Land formation is a natural phenomenon and hence these features are unevenly distributed.

Homogeneous Regions: The regions of an image where there is prominently one feature existing.

Heterogeneous Region: The regions where a mixture of features are coexisting together.

D. Mixed Pixels

Mixed pixels are ones having a signature representative of more than one land cover class (as with boundary pixels) or pixels saturated by reflectance or emittance of sub-pixel size features.

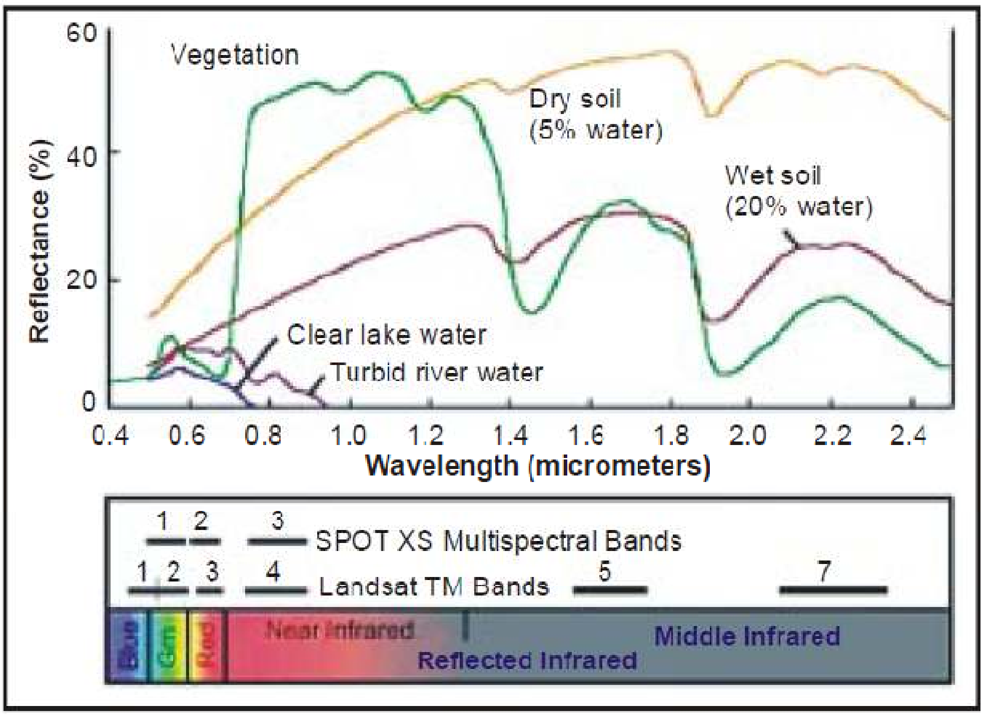


Figure 2.4 Typical Spectral Reflectance curves for vegetation, soil and water [13]

2.2 MULTISPECTRAL REMOTE SENSING

For the visual interpretation of remotely sensed images, a variety of image characteristics are considered: colour, texture, size, shape, pattern, and context. However, for computer-aided interpretation, it is most often simply colour (i.e., the spectral response pattern) that is used. It is for this reason that a strong emphasis is placed on the use of multispectral sensors (sensors that, like the eye, look at more than one place in the spectrum and thus are able to gauge spectral response patterns), and the number and specific placement of these spectral *bands*.

Figure 2.5 illustrates the multispectral bands from the LANDSAT satellite, which is a commercial system providing multi-spectral imagery in seven spectral bands at a 30 meter resolution [14].

It can be shown through analytical techniques such as Principal Components Analysis, that in many environments, the bands that carry the greatest amount of information about the natural environment are the near-infrared and red wavelength bands. Water is strongly absorbed by infrared wavelengths and is thus highly distinctive in that region. In addition, plant species typically show their greatest differentiation here. The red area is also very important because it is the primary region in which chlorophyll absorbs energy for photosynthesis. Thus it is this band which can most readily distinguish between vegetated and non-vegetated surfaces.

Given this importance of the red and near-infrared bands, it is not surprising that sensor systems designed for earth resource monitoring will invariably include these in any particular multispectral system. Other bands will depend upon the range of applications envisioned. Many include the green visible band since it can be used, along with the other two, to produce a traditional false color composite—a full color image derived from the green, red, and infrared bands (as opposed to the blue, green, and red bands of natural color images). This format became common with the advent of color infrared photography, and is familiar to many specialists in the remote sensing field. In addition, the combination of these three bands works well in the interpretation of the cultural landscape as well as natural and vegetated surfaces. However, it is increasingly common to include other bands that are more specifically targeted to the differentiation of surface materials [14].

For example, LANDSAT TM Band 5 is placed between two water absorption bands and has thus proven very useful in determining soil and leaf moisture differences. Similarly, LANDSAT TM Band 7 targets the detection of hydrothermal alteration zones in bare rock surfaces. By contrast, the AVHRR system on the NOAA series satellites includes several thermal channels for the sensing of cloud temperature characteristics [14].

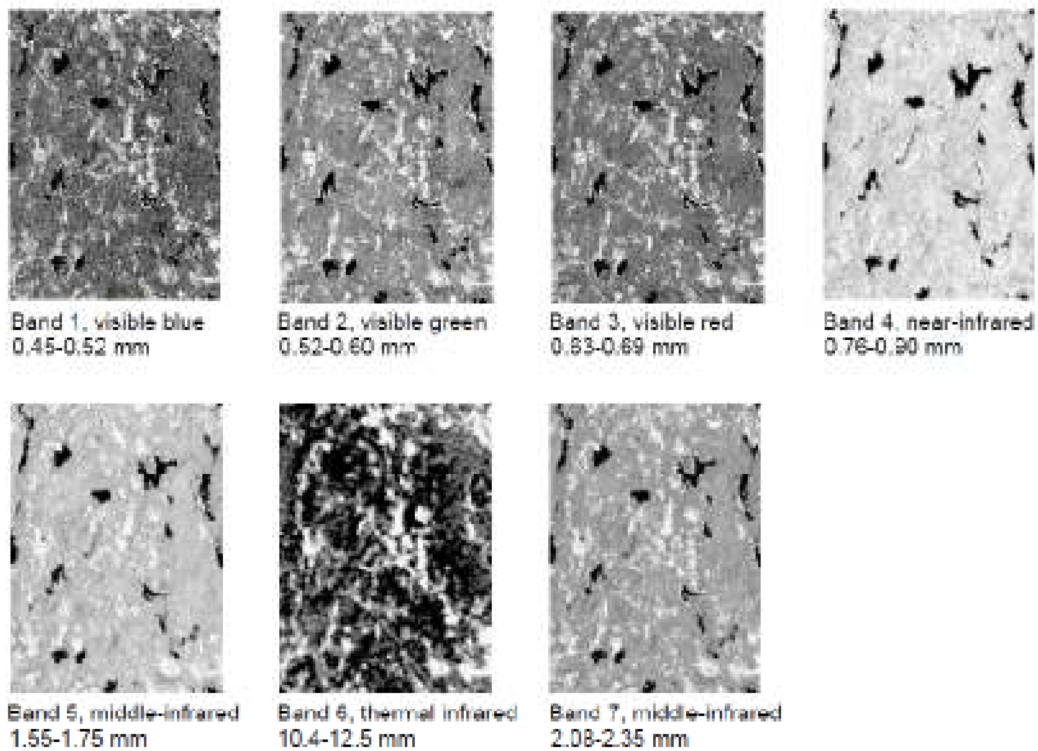


Fig 2.5 Spectral bands of the LANDSAT Thematic Mapper (TM) system [14]

2.3 OVERVIEW OF DIGITAL IMAGE PROCESSING

In today's world of advanced technology where most remote sensing data are recorded in digital format, virtually all image interpretation and analysis involves some element of digital processing. Digital Image Processing is largely concerned with four basic operations: ***image restoration***, ***image enhancement, image classification, image transformation.*** *Image restoration* is concerned with the correction and calibration of images in order to achieve as faithful a representation of the earth surface as possible—a fundamental consideration for all applications. *Image enhancement* is predominantly concerned with the modification of images to optimize their appearance to the visual system. Visual analysis is a key element, even in digital image processing, and the effects of these techniques can be dramatic. *Image classification* refers to the computer-assisted interpretation of images—an operation that is vital to GIS. Finally, *Image transformation* refers to the derivation of new imagery as a result of some mathematical treatment of the raw image bands.

2.4.1 Image Classification

A broad group of digital image-processing techniques is directed toward image classification, the automated grouping of all or selected land cover features into specified categories [15].

Classification is the process of assigning classes to the pixels in a remotely sensed image. Digital image classification uses the spectral information represented by the digital numbers (DN) in one or more spectral bands, and attempts to classify each individual pixel based on this spectral information. This type of classification is termed spectral pattern recognition. In either case, the objective is to assign all pixels in the image to particular classes or themes (e.g. water, coniferous forest, deciduous forest, corn, wheat, etc.). The resulting classified image is comprised of a mosaic of pixels, each of which belong to a particular theme, and is essentially a thematic map of the original image [16]

There are two major approaches to multispectral classification:

1. *Supervised classification.* The analyst defines on the image a small area, called a training site, which is representative of each terrain category, or class. Spectral values for each pixel in a training site are used to define the decision space for that class. After the clusters for each training site are defined, the computer then classifies all the remaining pixels in the scene.

2. *Unsupervised classification.* The computer separates the pixels into classes with no direction from the analyst.

With supervised classification, we identify examples of the Information classes (i.e., land cover type) of interest in the image. These are called "training sites". The image processing software system is then used to develop a statistical characterization of the reflectance for each information class. This stage is often called "signature analysis" and may involve developing a characterization as simple as the mean or the rage of reflectance on each bands, or as complex as detailed analyses of the mean, variances and covariance over all bands. Once a statistical characterization has been achieved for each information class, the image is then classified by examining the reflectance for each pixel and making a decision about which of the signatures it resembles most. (Eastman, 1995). Figure 2.6 describes this process.

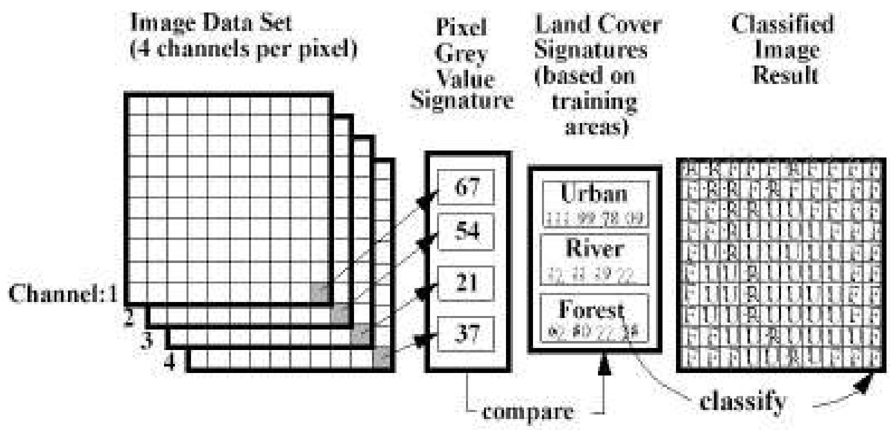


Figure 2.6 Supervised Classifications [16]

In **unsupervised classification** we provide the algorithm with basic information on how many classes we expect to be present on the image, and the algorithm attempts to identify those classes fig 2.7. Some unsupervised algorithms are also known as *clustering* algorithms. Unsupervised classifiers do not utilize training data as the basis for classification, rather this family of classifiers involves algorithm that examine the unknown pixel in an image and aggregate them into a number of classes based on the natural grouping or clusters present in the image values. The basic premise is that values within a given cover type should be closed together in the measurement space whereas data in different classes should be comparatively well separated. The classes that results from unsupervised classification are spectral classes because they are based solely on natural grouping in the image values, identity of the spectral classes will not be initially known. The analyst must compare the classified data within some form of referenced data to determine the identity and informational value of spectral classes.

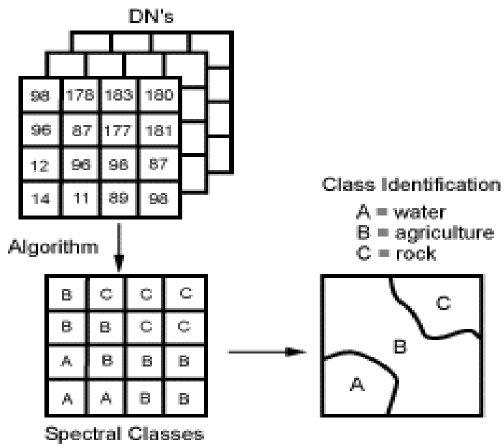


Fig 2.7 Unsupervised Classification

2.4.2 Steps of Image Classification

Image Classification procedure consists of three steps. These should be performed with accuracy so that the produced classified image will have better accuracy. The steps are 8defined as follows Fig 2.7

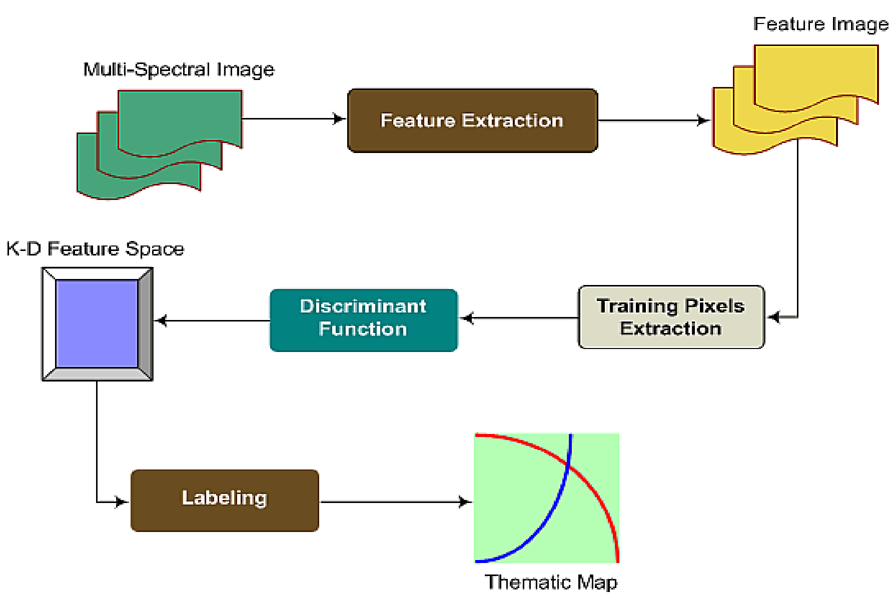


Fig 2.8 Steps in Image Classification [17]

2.8.1 Feature extraction.

This is an optional step of the classification process, which serves as a low-level pre-processing of the image to reduce its spectral or spatial dimensionality. It can be accomplished by using any type of spatial filter(s) or spectral transform(s) to reduce the data and/or enhance its multi-spectral features. In this stage, the multi-spectral image is transformed into a feature image.

2.8.2 Extraction of Training pixel. In this step, pixels from the image are extracted to train the classifier to recognize patterns that help to differentiate the classes. Based on these patterns, the classifier creates discriminator functions to assign each pixel to a class in the feature space. The training of the pixels can be either Supervised or Unsupervised.

2.8.3 Classes labelling.

In this final stage of the image-classification process, the discriminated functions are used to label all the pixels in the entire feature image. If the training of the pixels was supervised, then a previous knowledge of the classes' spatial distribution allows the labelling of classes to be carried out upon the application of the discriminate functions to the feature space. If the training was unsupervised, the assignment of a label to each discriminated class is subject to the analyst's own labelling criteria.

3. CLASSIFICATION TECHNIQUES

In this chapter we talk about the various classification techniques. We discuss some traditional techniques, its weaknesses and then move to nature inspired techniques.

3.1 TRADITIONAL CLASSIFICATION TECHNIQUES

The traditional classification techniques can be divided in two domains, supervised and unsupervised. In supervised classification training set is provided by the experts knowledge, while unsupervised no such training data is given (as explained in section 2.7)

3.1.1 Supervised Classification Techniques

Some of the traditional supervised classification techniques are:

- *Parellopiped Classifier*
- *Minimum distance to Mean Classifier*
- *Maximum Likelihood/Gaussian classifier*

A. Parallelopiped Classifier :

The Parallelopiped classifier is a supervised classifier that uses signature creation. It works by creating intervals of pixels values of bounded region of an area to determine whether a pixel belongs to a class or not. The values of the pixels of samples for the class give the intervals‘ bounding points.

Although parallolepiped classification is an accurate classification technique, it is not widely used because it has several disadvantages. The most important disadvantage is that it can leave many unclassified pixels. Another disadvantage of this classification method is that it can have overlap between training pixels.

B. Minimum Distance to Mean Classification

This technique uses the 'mean' as the basis for class labelling. In this classification first of all mean or average or spectral in each band for each category is determined. For each category there is a mean vector. Those pixels which are unidentified can be classified by computing the distance between the value of unknown pixel and each of category mean. The unknown pixel is assigned to the class to which it closely belongs to.

The minimum-distance-to-means strategy is mathematically simple and computationally efficient, but it has certain limitations like it is insensitive to different degrees of variance in the spectral response data. However this classifier is not widely used in applications where spectral classes are close to one another in the measurement space and have high variance.

C. Gaussian Maximum Likelihood Classification

The maximum likelihood classifier quantitatively evaluates both the variance and covariance of the category spectral response patterns when classifying an unknown pixel. Works on the Gaussian principle. For this, an assumption is made that the distribution of the cloud of points forming the category training data is Gaussian (normally distributed). This assumption of normality is generally reasonable for common spectral response distributions. In the Dimensional graph the vertical axis is associated with the probability of a pixel value being a member of one of the classes. The resulting bell-shaped surfaces are called probability density functions, and there is one such function for each spectral category. The probability density functions are used to classify an unidentified pixel by computing the probability of the pixel value belonging to each category.

The major disadvantage of this technique is that the maximum likelihood decision rule is based on the probability that a pixel belongs to a particular class. The basic equation assumes that these probabilities are equal for all classes, and that the input bands have normal distributions. The maximum likelihood algorithm assumes that the histograms of the bands of data have normal distributions, however this not true and hence not applicable in all applications.

3.1.2 Unsupervised Classification Techniques

Some of the traditional unsupervised classification techniques are:

- *K-means Clustering(KM)*
- *Expectation-Maximization (EM)*
- *Fuzzy C Means(FCM)*

A. K- Means Clustering

The basic step of K-means clustering is given in Fig 3.1 [18]. To initialize we determine number of cluster k and we assume center of this cluster. We can take any random object as initial centroid.

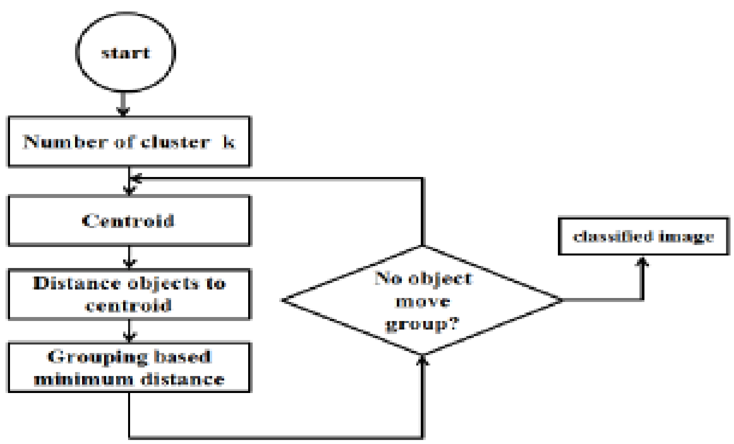


Fig. 3.1 K means algorithm [18]

B. Expectation-Maximization (EM) Algorithm

It is standard method to fit finite mixture models in to observed data. EM is an iterative procedure which converges to a (local) maximum of the marginal a posteriori probability function [19]. EM is a general method of estimating the features of a given data set, when the data are incomplete or have missing values. Being an iterative procedure, the EM method can present high computational cost.

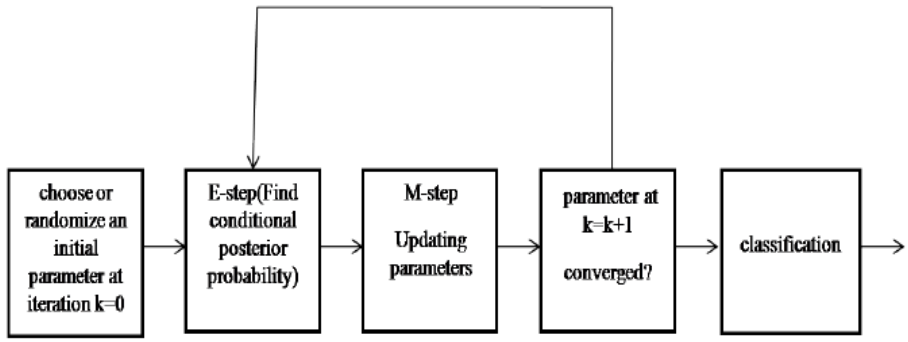


Fig. 3.2 EM algorithm [20]

C. Fuzzy C Means Clustering

In **fuzzy clustering**, which is an unsupervised classification each point has a degree of belonging to clusters, as in fuzzy logic, rather than belonging completely to just one cluster. Thus, points on the edge of a cluster may be in the cluster to a lesser degree than points in the centre of cluster. The fuzzy c-means algorithm is very similar to the k-means algorithm, just differ that it is based on the fuzzy logic. Its main steps are:

- Choose a number of clusters.
- Assign randomly to each point coefficients for being in the clusters.
- Repeat until the algorithm has converged (that is, the coefficients' change between two iterations is no more than ε, the given sensitivity threshold) :
    - Compute the centroid for each cluster, using the formula above.
    - For each point, compute its coefficients of being in the clusters, using the formula above.

The algorithm minimizes intra-cluster variance as well, but has the same problems as *k*-means; the minimum is a local minimum, and the results depend on the initial choice of weights. The expectation-maximization (EM) algorithm is a more statistically formalized method which includes some of these ideas: partial membership in classes. Fuzzy c-means has been a very important tool for image processing in clustering objects in an image.

3.1.3 Weakness of Traditional Techniques

The research community has been trying to model and solve complex optimization problems by employing natural metaphors. This is due to inefficiency of classical optimization algorithms in solving complex non linear and large scale combinatorial problems.

It is a known fact that classical optimization techniques impose several limitations on solving mathematical programming and operational research models. This is mainly due to inherent solution mechanisms of these techniques. Solution strategies of classical optimization algorithms are generally dependent on the type of objective and constraint functions like linear, non-linear etc. and the type of variables used in the problem modelling like integer, real etc.. Their efficiency is also gets effected by the size of the solution space, number of variables and constraints used in the problem modelling, and the structure of the solution space like convex, non-convex, etc. They also do not offer general solution strategies that can be applied to problem formulations where, different type of variables, objective and constraint functions are used. As an example the simplex algorithm can be used to solve models with linear objective and constraint functions; geometric programming can be used to solve non-linear models with a polynomial or signomial objective function etc. However, most of the optimization problems require different types of variables, objective and constraint functions simultaneously in their formulation [21]. Therefore, classic optimization procedures are generally not adequate or easy to use for their solution. Researchers have spent a great deal of effort in order to adapt many optimization problems to the classic optimization procedures. It is generally not easy to formulate a real life problem that suits a specific solution procedure.

In order to achieve this, it is necessary to make some modifications or assumptions on the original problem parameters. This affects the solution quality greatly.

3.2 NATURE INSPIRED ALGORITHMS

A new set of problem and model independent nature inspired heuristic optimization algorithms were proposed by researchers to overcome drawbacks of the classical optimization procedures. These techniques are efficient and flexible. They can be modified and/or adapted to suit specific problem requirements. Research on these techniques is still continuing all around the globe.

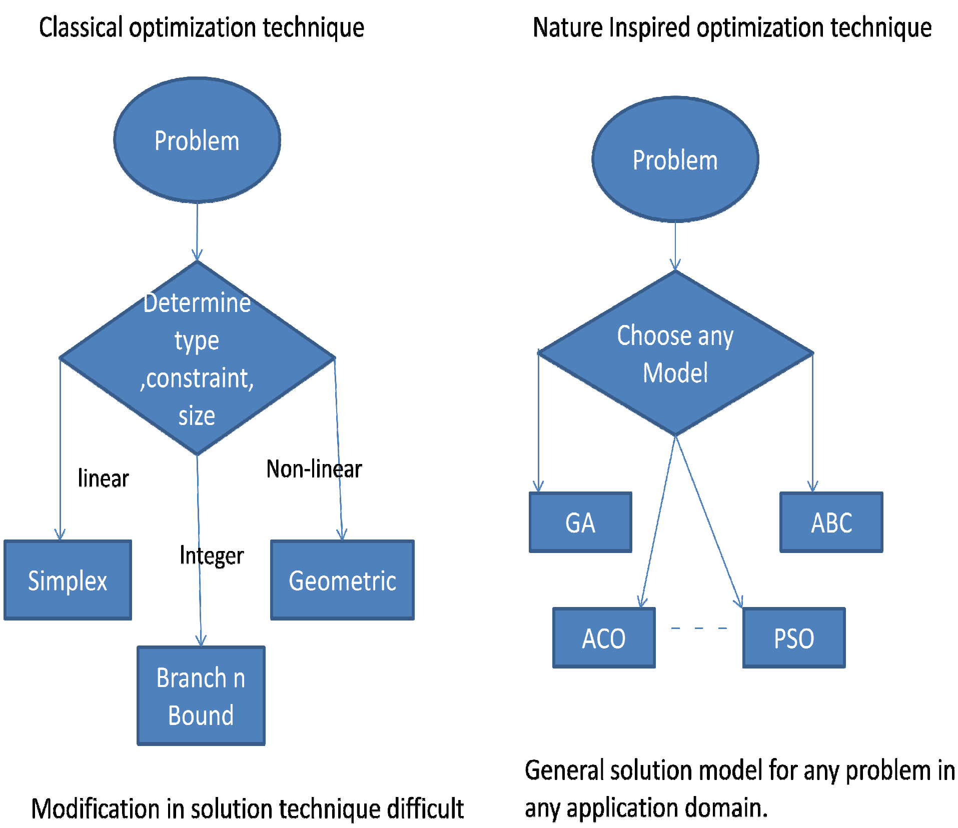


Figure 3.3 Comparison of Nature inspired and Traditional optimization algorithms

Nature-Inspired Algorithms are motivated by a variety of biological and natural processes. They are capable of modelling natural phenomenon in a proficient manner; the processes of nature have been widely understood and sculpt into computational paradigms. Natural systems, especially social insect's colonies, like ant, bee, termites etc and some natural processes like bird migration, osmosis, chromosomal genetics and plate tectonics etc, inspired scientist for discovering new algorithms, which help in the study of artificial intelligence.

The popularity of the Nature-Inspired Algorithms is primarily caused by the ability of biological systems to effectively adjust to frequently changeable environment. Evolutionary computation, neural networks, ant colony optimization, particle swarm optimization, artificial immune systems, and bacteria foraging algorithm, Biogeography based optimization, genetic algorithm and artificial bee colony algorithm are the significant algorithms and concepts that were motivated by nature. These algorithms are mostly optimization methods, and aim to find the global optimum of a real-valued function (fitness function) defined in a given space (search space).

A branch of nature inspired algorithms known as swarm intelligence is focused on insect behaviour in order to develop some meta-heuristics which can mimic insect's problem solution abilities. Ant colony optimization, particle swarm optimization, artificial bee colony, wasp nets etc. are some of the well known algorithms that mimic insect behaviour in problem modelling and solution.

3.2.1 Taxonomy of Nature Inspired Techniques

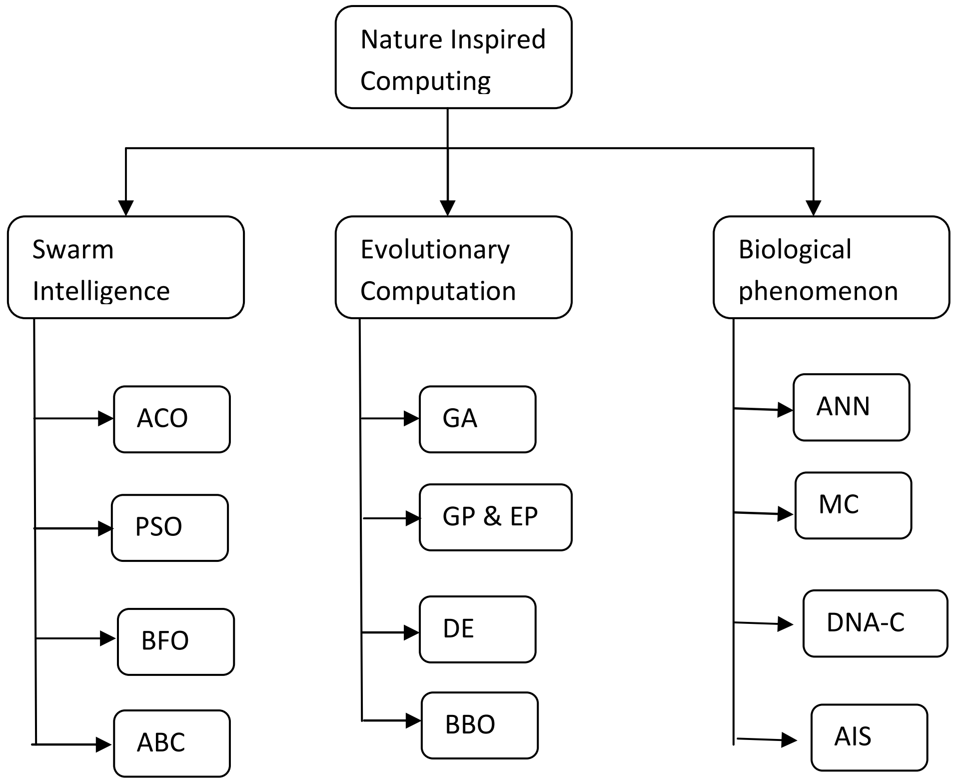


Fig 3.4 Taxonomy of Nature inspired Computing

Broadly we have divided the nature inspired computing in three categories namely:

- *Swarm Intelligent techniques:* they are focused on a large population of insect behaviour in order to develop some meta-heuristics which can mimic their problem solution abilities. The major algorithms that fall in this category are Ant Colony Optimization (ACO), inspired from intelligent social behaviour of ants [3], Particle Swarm Optimization (PSO) inspired from bird flocking [22], and Bacterial Foraging Optimization (BFO) from the social foraging behaviour of Escherichia coli [23] and Artificial Bee Colony (ABC) from the foraging behaviour of the honey bees [2].

- *Evolutionary Computation:* it is a paradigm that is concerned with the investigation of systems inspired by the neo-Darwinian theory of evolution by means of natural selection. Popular evolutionary algorithms include the Genetic Algorithm (GA), Genetic and Evolutionary Programming (GP &EP), Differential Evolution (DE) and Biogeography Based Optimization (BBO) which is based on the migratory behaviour of populations.

- *Biological phenomenon-* these algorithms simulate some of the core biological phenomenon like Artificial Neural Network (ANN) from the neurons in brain, Membrane Computing (MC) from passing objects through cellular membranes, DNA computing and Artificial Immune System (AIS)

3.2.2 Some Nature Inspired Classification Techniques

A. Biogeography Based Optimization (BBO)

Biogeography is the study of the geographical distribution of biological organisms. Biogeography based Optimization (BBO) is an application of biogeography to optimization problems. It is modelled on the immigration and emigration of species between the islands. BBO was first presented in December 2008 by D. Simon [24]. A supervised classification of remote sensing image based on BBO is proposed in [25]. In this first random clusters of the image are formed using rough set theory. Then each cluster is put in the universal habitat, considering each cluster as a species of universal habitat. The feature habitats initially contain training pixels and Habitat Suitability Index (HSI) is calculated on these training pixels. Then each of the species is migrated to feature habitat one by one and HSI of the habitat is recalculated after migration. A given species is absorbed (and hence it belongs) to that feature habitat where it makes minimum variation in HSI after migration.

B. cAnt Miner

Fernando, Freitas, and Johnson proposed an extension to Ant-Miner (based on ACO), named cAntMiner [42], which was able to cope with the continues values rather than only the discrete values. The goal of this algorithm is to extract the classification rules from the given dataset. This is a cluster based approach. The classification rules were generated using the Myra tool and then applied the extracted classification rules on each of the clusters of the image. Unlike Ant Miner cAntMiner doesn't requires the discretization as a pre-processing method. So it is suitable for all the types of attributes.

C. Hybrid of PSO/ACO

The "standard" PSO algorithm [43] does not deal with categorical values in a natural fashion as done by ACO. To cope with multi valued categorical attributes each of the particles is represented by bit string, where each bit can take value such as true or false which is encoded as 1 or 0. For image classification application this was being implemented by using an open source tool for hybridized PSO/ACO [44].

D. Hybrid of PSO/ACO2

PSO/ACO2 uses the sequential covering approach to discover one classification-rule-at-a-time. This new version was given by Nicholas and Freitas [45]. Underlying steps are:

- Initially Rule Set is empty.
- Loop for each class of cases Trs i.e. contains all training cases.
- Loop until amount of uncovered training cases of class A > Maximum uncovered cases per class.
- Run the PSO/ACO algorithm for finding best nominal rule
- Standard PSO algorithm is used to add continuous terms to Rule, and return the best discovered rule.
- Do Pruning of the discovered Best Rule

- RuleSet = Rule Set Union Best Rule
- Trs = Trs – {training cases that are correctly covered by the discovered rule}
- End of   inner loop
- End of outer lop
- Do Ordering of the rules in RuleSet by descending Quality

E. Membrane Computing

This natural technique was proposed by Gheorghe Păun and Grzegorz Rozenberg [26]. Membrane computing is basically a model of computation that is motivated by the structure and functioning of living cells- more specifically, by the role of membranes in compartmentalization of living cells into ―protected reactors [26]. It aims at abstracting computing models from the structure and the functioning of living cells, as well as the way the cells are arranged and organized in tissues. In MC all classes of computing devices are generally called *P systems* [27]. The main components of a P system are (i) the membrane structure, (ii) the Multi-sets of objects placed in the compartments of the membrane structure, and (iii) the evaluation rules for processing the objects and the membranes. Membrane Computing has been implemented for satellite image classification [28] and had good results.

F. Hybrid ACO2/BBO

Lavika, Panchal [29] have proposed a hybrid classification technique using ACO and BBO. The technique can be adapted with the database of expert knowledge for a more focused satellite image classification. BBO refines its solutions probabilistically after each iteration unlike ACO which produces new solutions. Hence BBO has been used for the refinement of the output of ACO Classifier image (cAntMiner) Algorithm.

Hybrid of ACO2/BBO/PSO

In order to further improve the performance hybridization of ACO, BBO and PSO was performed. In this approach [46] the firstly the clusters are formed by using fuzzy c means or any other unsupervised clustering approach then in the next step BBO is applied onto these clusters. The result from BBO is further refined by applying ACO2/PSO classifier.

4. THE NEW SWARM TECHNIQUE OF ARTIFICIAL BEE COLONY ALGORITHM

Intelligent social behaviour of living beings in the nature has lead researchers to develop new optimization method. Karaboga developed Artificial Bee Colony (ABC) Algorithm by modelling foraging behaviours of honey bees [2]. ABC was developed under the inspiration of collective behaviour of honey bees with better performance in function optimization problem compared with GA, differential evolution (DE), and particle swarm optimization (PSO) [30].

ABC tries to model natural behaviour of real honey bees in food foraging. Honey bees use several mechanisms like waggle dance to optimally locate food sources and to search new ones. This makes them a good candidate for developing new intelligent search algorithms. Since ABC algorithm is simple in concept, easy to implement, and has fewer control parameters, it has been widely used in many fields [31]. Until now, ABC has been applied successfully to some engineering problems, such as constrained optimization problems [32], neural networks [33], and clustering [34].

4.1 BIOLOGICAL INSPIRATION

Social insect colonies can be considered as dynamical system gathering information from environment and adjusting its behaviour in accordance to it. While gathering information and adjustment processes, individual insects don't perform all the tasks because of their specializations.

In the nature honey bees explore the locality of their hive in search of better nectar sources [35]. There are 3 types of bees depending on their work, employed bees, onlooker bees and the scout bees. The employed bees first leave the hive and search in particular direction for sources of pollen and nectar. After finding a suitable nectar source they bees go back to the hive and share their information to onlooker bees about the locations, quantity and quality of existing sources of pollen, nectar in the areas they have explored. Information exchange in the bees swarm is the most important occurrence in developing the collective knowledge. The area in the hive for this process is the dancing area. Bees exchange information about the food sources by a dancing ritual called a ‘waggle dance’ to promote their food locations and persuade the members of the colony to trail their lead. The onlooker bee after watching the dances chooses the most profitable source and follows one of the employed bees to the discovered source of food. Upon arrival to the food source, the foraging bee takes a load of nectar and returns to the hive relinquishing the nectar to a food storer bee. The options available for the bee now are [36]:

(1) Discard the nectar source location and become again an uncommitted follower;

(2) continue the foraging at the discovered nectar source, without recruiting any other bee;

(3) Recruit other bees with the dance ritual prior to return to the food location.

The bee decides for one of the possibilities using the mechanism based on the characteristics of the food source (quality, quantity, and distance from the hive) [36]. The described process continues constantly, while the bees from a hive collect nectar and investigate new areas with possible food sources.

There are three essential components of ABC optimization model of food source selection that leads to the emergence of collective intelligence of honey bee swarms: food sources, employed foragers and unemployed foragers. There are two basic behaviours: recruitment to a food source and the abandonment of a food source.

1) *Food sources:* In order to select a food source, a forager bee evaluates several properties related with the food source such as its closeness to the hive, richness of the energy, taste of its nectar

2) *Unemployed foragers:* these are of two types, scouts and onlookers. Their responsibility is exploring and exploiting food source. At the beginning, there are two choices for the unemployed foragers: (i). it becomes a scout – randomly search new food sources around the nest; (ii). It becomes an onlooker – determine

the nectar amount of food source after watching the waggle dances of employed bee, and select food source according to profitability.

3) *Employed foragers:* these search for and are equal to the number of food sources. The employed bees store the food source information and share with others according to a certain probability. The employed bee will become a scout when food source has been exhausted.

All the information about the currently rich food sources is available on the dance area and the onlooker watches numerous dances performed by the employed bees and chooses the profitable source. The onlooker bee decides the profit using the probability values of the food sources. The recruitment is thus proportional to the profitability of a food source.

4.2 ABC DESCRIPTION

In this way, artificial bee colony mimics the behaviour of natural bees to construct a relative good solution of realistic optimization problems.

The colony of artificial bees contains three groups of bees: employed bees, onlookers and scouts. First half of the colony consists of the employed artificial bees and the second half includes the onlookers. For every food source, there is only one employed bee, the number of employed bees is equal to the number of food sources.

The employed bee of an abandoned food source becomes a scout. The search carried out by the artificial bees can be summarized as follows four steps [37]:

1) Employed bees determine a food source within the neighbourhood of the food source in their memory. The foods sources are simulated by the position of solution of optimization problem, the profitability of food source are expressed as fitness of the solution.

2) Employed bees share their information with onlookers within the hive and then the onlookers select one of the food sources;

3) Onlookers select a food source within the neighbourhood of the food sources chosen by them;

4) An employed bee of which the source has been abandoned becomes a scout and starts to search a new food source randomly.

Basically, there are two important function supports the ABC algorithm [38]:

$$Pi = \frac{fiti}{\sum_{n=1}^{SN} fitn} \qquad \text{-(a)}$$

$$V_{ij=} X_{ij} + \Phi_{ij} (X_{ij} - X_{kj}) \qquad \text{-(b)}$$

Where P_i is the probability value associated with ith food source that calculated by the Equation (a). An onlooker bee selects a food source relying on P_i . In this equation, fit_i represents ith food source's nectar amounts, which is measured by employed bees and SN is the number of food source which is equal to the number of employed bees. Fitness is calculated using following equation:

$$fitness\ (i) = \begin{cases} 1/(1 + fnci, & fnci \geq 0 \\ 1 + abs(fnci), & fnci < 0 \end{cases}$$

fnc_i in the equation is the cost function of the quality of source. Greedy selection is applied to select the best source. In the real-world problems, X_{ij} and X_{kj} represent the different old food source positions. The difference between these two positions is the distance from one food source to the other one. D is the number of optimization parameters. Φ_{ij} is a random number between [-1, 1] and controls the distance of a neighbour food source position around X_{ij}. [38]

4.3 ABC ALGORITHM

In this section we discuss the artificial bee colony algorithm as proposed by Karaboga.

4.3.1 Original Algorithm

1. Initialise food source positions

Iterate till maximum cycle or desired error

2. Sending employed bees to food source positions
3. Calculating probability values used in probabilistic selection
4. Selection of food source positions by onlooker bees considering probability values
5. Onlooker bee following to that food source.
7. Abandoning sources with less probability and producing new food source in neighbourhood of old source.

4.3.1.1 *Producing Food Source Positions*

Algorithm starts the procedure by producing random food source positions corresponding to solutions in the search space. Producing source positions occur by producing random value after determining lower and upper limits [39].

$$X_{ij} = X_j^{min} + rand(0,1) * (X_j^{max} - X_j^{min})$$

Here, i equals to1…,SN and j equals to 1,…,D. SN stands for food number and D for parameter number that will be optimized. X_j^{max} , X_j^{min} are the lower and upper limits respectively.

4.3.1.2 *Sending Employed Bees to Food Source Positions*

Employed bee determines a new food source neighbour to food source it is responsible for and evaluates its quality. If the new source is higher in quality, employed bee keeps this new source in memory.

$$V_{ij=} X_{ij} + \Phi_{ij} (X_{ij} - X_{kj}) \qquad \text{-(b)}$$

X_{ij} and X_{kj} represent the different old food source positions or solution positions. The difference between these two positions is the distance from one food source to the other one. D is the number of optimization parameters. Φ_{ij} is a random number between [-1, 1] and controls the distance of a neighbour food source position around X_{ij}. [38]

4.3.1.3 *Calculating probability values used in probabilistic selection*

Employed bees completing their search process their source nectar information to onlooker bees. Onlooker bee selects a source that is proportional in probability to food nectar information by making use of this shared information.

$$Pi = \frac{fiti}{\sum_{n=1}^{SN} fitn}$$

Where P_i is the probability value associated with ith food source that calculated by the Eq. a. An onlooker bee selects a food source relying on P_i . In this equation, fit_i represents ith food source's nectar amounts, which is measured by employed bees and SN is the number of food source which is equal to the number of employed bees. Fitness is calculated using following equation:

$$fitness\ (i) = \begin{cases} 1/(1 + fnci, & fnci \geq 0 \\ 1 + abs(fnci), & fnci < 0 \end{cases}$$

fnc_i in the equation is the cost function of the quality of source. Greedy selection is applied to select the best source.

4.3.1.4 *Selection of Food Source Positions by Onlooker Bees*

P_i values calculated according to the above equation and are compared with the other probability values generated for each source in the neighbourhood of the solution. The solution with the highest probability gets selected by the onlooker.

4.3.1.5 *Abandoning Depleted Source and Producing new sources*

After employed and onlooker bees complete searching process (at the end of a cycle), new solutions are searched in the neighbourhood. The source with less probability is abandoned.

4.3.2 STRENGTHS OF ABC

In ABC algorithm the optimization process begins with a population of randomly generated solutions (food source positions) and then iteratively improves this population to discover better solutions (more profitable sources). The improvement process consists of three phases: sending employed bees to their solutions and recruiting onlooker bees to the solutions probabilistically depending on their fitness values for improving them; and after determining the exhausted sources, converting their employed bees into scouts discovering new solutions randomly for the exhausted ones.

In [30] the performance of ABC algorithm was compared with those of GA, PSO, DE and ES optimization algorithms. It showed that the performance of the ABC is better than or similar to those of other population-based algorithms with the advantage of employing fewer control parameters. It was also observed that the standard ABC can be efficiently used for solving multimodal and multidimensional optimization problems.

In [4] paper, the most popular AI optimization algorithms, PSO, GA, DE and ABC are used for the unsupervised classification of remotely sensed image data. The results show the superiority of the ABC over other AI algorithms. With the increase in computational complexity the requirement of computer processing power also increases, but determination of parameters affects the success of the algorithms significantly. This dependence to the values of parameters means heuristics like PSO, GA, DE are not robust to the values of parameters. The performance of ABC was found to be stable to the changes of the initial conditions [this]. If parameter values are properly selected, heuristic methods can produce comparable results.

5. ARTIFICIAL BEE COLONY CLASSIFICATION ALGORITHM

Swarm intelligence is a branch of evolutionary computation, which models the behaviours of the self-organizing individuals performing collective intelligence. Honey bees are a very good example of self-organizing individuals which carry out global tasks collectively without supervision. A global food search is performed based on self organization and division of labour. It is one of the crucial tasks of a bee colony. Karaboga described an algorithm called artificial bee colony-ABC which simulates the foraging behaviour of honey bees and used this algorithm to solve numeric optimization problems. In ABC approach as discussed in chapter 4, the position of a food source represents a possible solution to the optimization problem addressed and the profitability of the source corresponds to the quality of the solution. Land cover mapping is an important application area of remote sensing discipline and Classification is the preferential step for producing thematic spatial information from satellite image data. We used the strength of ABC to most optimally classify the images into respective land-usage.

In this chapter we propose an implementation of ABC for satellite image classification. The objective here is to adapt the bee communication and food search method of information exchange and hence achieve maximum classification accuracy.

5.1 ARCHITECTURE OF OUR PROPOSED CLASSIFIER

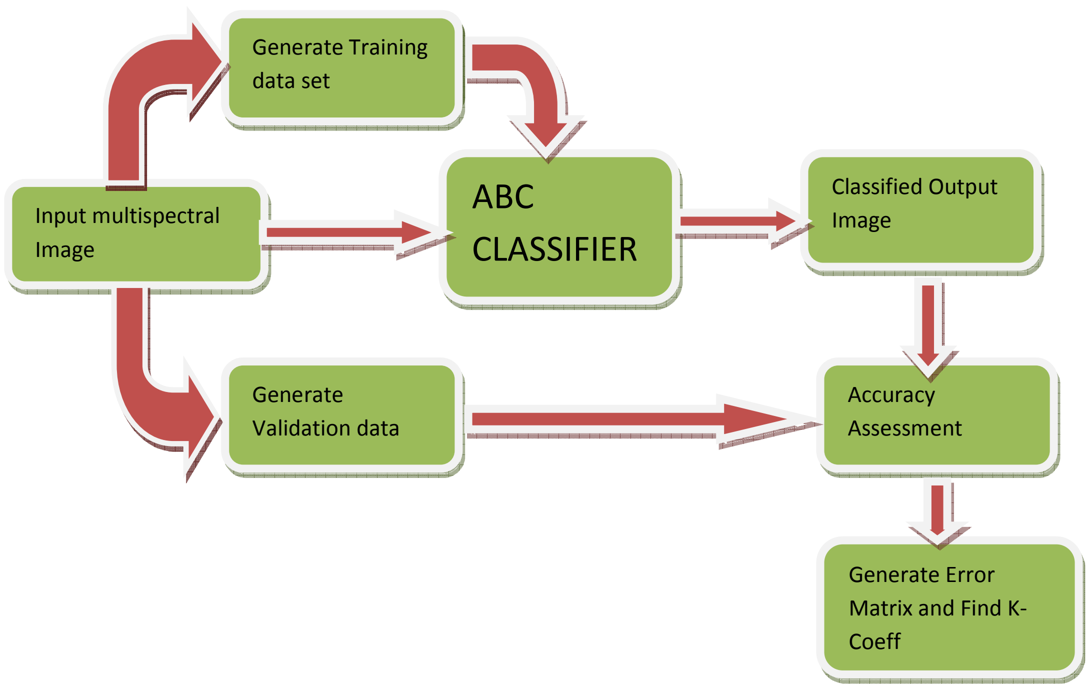


Figure 5.1 Architecture of Proposed Classifier

Stage 1: *Input Multispectral Image*

This is the very first stage where the input is fed to the system. For classification problem we have considered remote sensing multispectral image, like the 7 band image of Alwar. The images are input as '.tif' format and are read using function imread() in matlab.

Stage 2: *A- Generate training data set*

*O*ur classifier is a supervised one where we need some already classified expert's data to train our ABC classifier. This data is generated by some experts, in our case by scientists at DTRL laboratory at DRDO. This is also the second major input in our classifier. The training dataset is in form of excel sheets with some random pixel information of each band and the feature they belong.

B- Generate Validation dataset

A validation set is required to validate our classification results. This is also generated from the multispectral image by the experts. It is very much alike the training set just that is not input into the system but used for verification.

Stage 3- *ABC Classifier*

This is the core stage of our entire process where the input image is processed using the training dataset. Our proposed algorithm is implemented here for the classification of the image. We get the classified output image from the classifier.

Stage 4- *Classified Output Image*

This is the output of our classifier. The land cover features of the input multispectral image are extracted and given different colour codes for distinguish-ability. All the pixels of the image hence get unique colour for representation. This new classified image is then displayed in matlab output.

Stage 5- *Accuracy Assessment*

This is an important stage where the validation and accuracy checking of our classified output is done. The validation dataset is the other input for this stage. Here all the pixels contained in the validation set are checked with our classifier. It is seen whether the classifier correctly classifies those pixels or not. Since these pixels are marked by experts we can assess the accuracy of the classifier.

Stage 6- *Construction of error matrix and calculating kappa coefficient*

This can be called as a part of accuracy assessment, but shown is done after validation data set is acted upon by our classifier. We generate an Error matrix which determines quantitatively that how pixels are correctly absorbed in their required features. 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 also find the kappa coefficient. It serves as an indicator of the extent to which the percentage correct values of an error matrix are due to "true" agreement versus "chance" agreement.

5.2 OUR ARTIFICIAL BEE COLONY CLASSIFICATION ALGORITHM

In ABC system, bees search around in a multidimensional search space and some (employed and onlooker bees) choose food sources depending on the experience of themselves and their hive mates, and adjust their positions. If the nectar amount of a new source is higher than that of the previous one in their memory, they memorize the new position and forget the previous one.

Assumptions:

- Bees are represented by pixels of the image;
- Food sources are the land cover features- water, vegetation, urban, rocky and barren;
- Neighbourhood solutions are the neighbouring pixels of the already classified dataset provided by experts.
- Employed bees are simulated by pixels belonging to classified dataset which contain the pure values of the solution.
- Here the function values (nectar quality) are calculated using Euclidean distance, other distances like Bhattacharya and Manhattan can also be used.

The Algorithm:

Input: Multi-spectral satellite image

Output: Classified Image

Step 1: *Initialization*

- Initialize the population of solutions (food sources).
- Place the employed bees in the food sources.

Step 2: *Quality assessment of food source*:

- Evaluate the nectar quality of the source, here represented by the Euclidean distance between food source and particular bee.

Step 3: *Greedy selection by Onlooker bees*:

Repeat till all pixels classified

Repeat till all neighbourhood food sources exhaust

- Find function value (fnc_i) using Euclidean distance formula.
- Calculate the Fitness values using:

$$fitness\ (i) = \begin{cases} 1/(1 + fnci, & fnci \geq 0 \\ 1 + abs(fnci), & fnci < 0 \end{cases}$$

- Calculate the probability values P_i for each solution (food source)

$$Pi = \frac{fiti}{\sum_{n=1}^{SN} fitn}$$

- Find new food source in the neighbourhood of old source

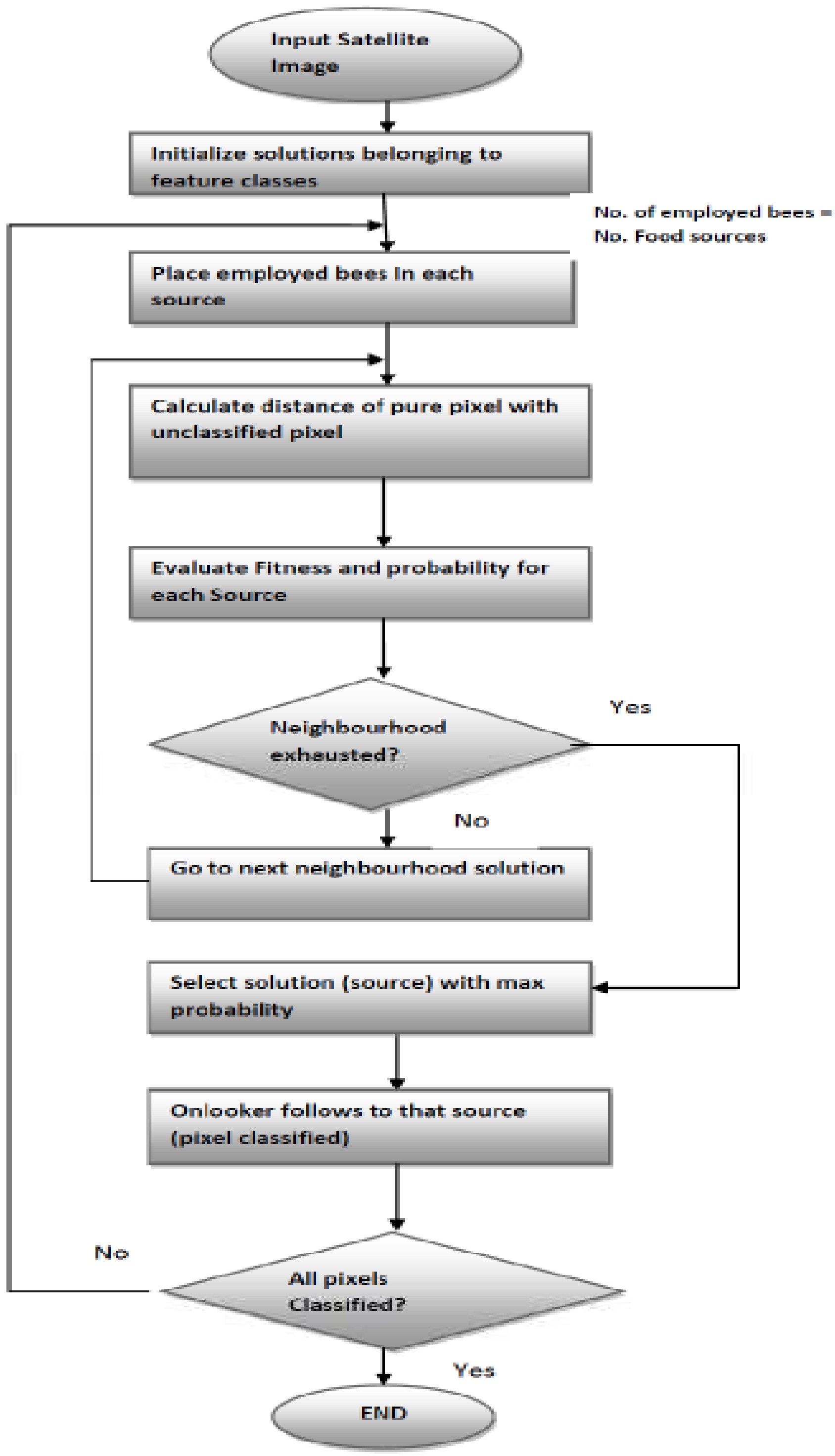
Inside loop end – (Neighbourhood exhaust)

- Selection of food source by the onlooker for maximum Pi value.
- Onlooker following the employed bee to that source.
- Another bee is now made recruiter to food source.

End loop when all pixels classified

Display the output classified image.

5.3 FLOWCHART REPRESENTATION OF ALGORITHM



5.4 DETAILED DESCRIPTION OF OUR PROPOSED CLASSIFICATION TECHNIQUE

Input – the input to the classifier is the multispectral remote sensing image.

For the two datasets we have used in our experiments the inputs are:

- 7 band Alwar region images
- 6 band Saharanpur region images

Output- the output after executing our technique is the classified image with all required land cover features extracted and represented with different colour codes.

To start with, for simulation of the honey bees foraging behaviour we have to correlate our image classification scenario with their food source finding problem.

Some issues that need to be resolved first are:

- What would be the food sources or the solutions in our case
- How do we simulate the bees of the hive
- How do we represent the employed and onlooker bees
- How is the neighbourhood of a solution found and defined
- How would the nectar qualities of any source/solution be evaluated

To resolve these issues we came by some assumptions and way to simulate the bee colony foraging. These assumptions are –

- *Food Sources:* The food sources or the solutions shall belong to the different training feature dataset provided by the experts. Hence if we know an image contains 5 feature classes and the experts provide us different data classes belonging to each feature, then each pixel of a class is a probable solution.
- *Hive and the bee*- the hive area is the entire image and the bees in general are the pixels of the image.

- *Employed bees*- Employed bees are simulated by pixels belonging to classified dataset which contain the pure values of the solution.
- *Onlooker bees*- they are the unclassified pixels remaining in the image. As these bees/pixels find their suitable solution and belong to that solution class, hence get classified.
- *Defining Neighbourhood*- the neighbourhood of a solution is represented by the class boundary of the classified training dataset. Neighbourhood solutions are the neighbouring pixels of the already this dataset provided by experts.
- *Nectar Quality* – the nectar quality are the function values used in the algorithm. The function values are calculated using Euclidean distance between the pure value of the solution and the raw value of unclassified pixel.

Step 1- Generating training dataset from the image

The classification here is supervised and hence we require training set for our classifier. This is generated by experts in geosciences and remote sensing field that provide excel sheets with some x, y coordinates and their band reflectance values and the feature class they belong.

Step 2- Initialization

Each pixel of the training set is considered as a solution with neighbourhood as the feature class boundary. The first pixels in each class are initial solutions after which the neighbouring pixels are considered for solution. The employed bee initially stores the pure reflectance value of the solution (food source)

Step 3- Quality assessment of solution

- The nectar quality of the source is evaluated, here represented by the Euclidean distance between food source and particular bee.

  Formula $\sqrt{(a1 - b1)^2 + (a2 - b2)^2}$

  a1, a2 and so on are the pure pixel values and b1, b2 are the unclassified pixel values.

Step 4- greedy selection of solution by the onlooker bee

The greedy selection of nectar sources is a characteristic of bees where they follow the employed bee whose nectar quality and probability is the highest.

Here we will loop till all the pixels in the image are classified or we can say all the onlooker bees find their sources.

Outer loop- for all pixels of the image

Inner loop- for exploring the neighbourhood solutions of each class.

- Now we take each pixel find the function values using the Euclidean distance
- After finding the function values we calculate the fitness of the solution using the standard ABC equation :

  $$fitness\ (i) = \begin{cases} 1/(1+fnci, & fnci \geq 0 \\ 1+abs(fnci), & fnci < 0 \end{cases}$$

  From the equation we see that more the function value (fnci), lower will be its fitness value, they are inversely proportional.
- From the fitness we evaluate the probability of solution using: $Pi = \frac{fiti}{\sum_{n=1}^{SN} fitn}$
- After one solution is being evaluated we go to the next neighbourhood solution and do the same procedure to find fitness and probability.
- The inner loop ends when all the neighbourhood of each class is considered, we can say that the neighbourhood has exhausted.
- The solution with maximum probability value is then chosen by the onlooker pixel and it gets classified in that class.
- This process continues for each pixel of the unclassified image and stops when the whole image is classified.

Step 5- Color codes assignment

As each pixel is getting classified we assign a unique color code to it and create a new output image where it shall be displayed. This output image will show our classification result once all pixels in input are processed.

6. EXPERIMENTS AND RESULTS

We have applied our classification process on different dataset and it had proved that it is a universal classifier. The implementation is done in MATLAB, using the image processing toolbox whose functions details are given in appendix B. First we have shown our result on Alwar region and then on Saharanpur region. In this section we prove that our classifier is able to correctly map the different terrain features of the two experimental datasets we take by showing the classified outputs in each case. We also study the dataset on which work has been done. For our supervised classification the training data is generated and studied. Figure 6.6 and 6.7 shows the original false color images and classified image of Alwar region. Figure 6.8 and 6.9 shows the original false color image and classified image of Saharanpur area respectively. Our objective is to use the proposed nature inspired computation algorithm as an efficient land cover classifier for satellite image. Based on the results obtained on applying artificial bee colony to the 7-band image of Alwar region and 6-band image of Saharanpur region for image classification, we conclude that out classifier is successful in its optimization problem.

6.1 EXPERIMENTAL SETUP USED IN THIS WORK

Our objective is to use our universal classifier as an efficient land cover classifier for satellite image. As we have applied our classifier in two different dataset thus we have described both dataset here. We have taken multi-spectral, multi resolution and multi sensor image of size 472 X 546 of Alwar area in Rajasthan and 641 X 641 resolution image of Saharanpur region in Uttar Pradesh. For the first case 7 different bands and for the second one 6 different bands are taken.

6.1.1 The 7 Band multispectral image of Alwar Region

We have taken a multi-spectral, multi resolution and multi-sensor image of size 472 X 546 of Alwar area in Rajasthan. The satellite image for 7different bands is taken. These bands are Red, Green, Near Infra Red(NIR), Middle Infra Red (MIR), Radarsat-1 (RS1), Radarsat-2(RS2), and Digital Elevation Model) DEM. The ground resolution of these images is 23.5m and is taken from LISS-III, sensor and is provided to us courtesy of DTRL (Remote sensing) Lab, DRDO, Delhi.

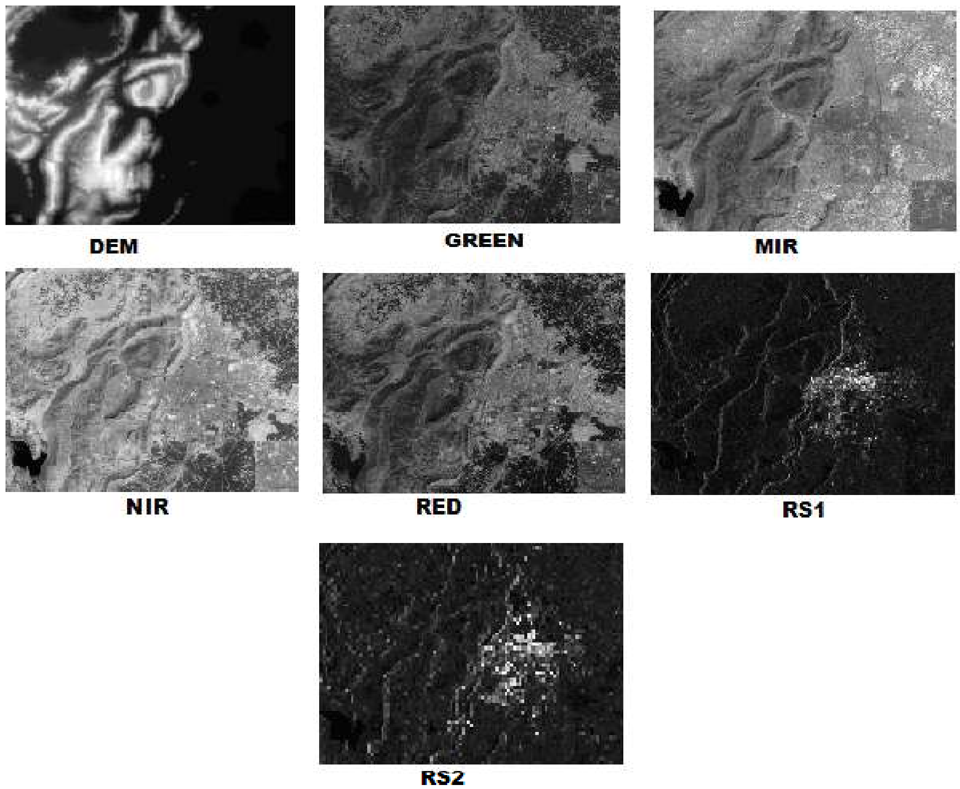


Fig 6.1 The 7 Band Multispectral Images of Alwar Dataset

Alwar region has mainly the following features:

- Water
- Urban
- Vegetation
- Rocky
- Barren

Our geosciences experts recognized all these features in Alwar region. Thus our aim is to extract all these features from Alwar image.

6.1.2 The 6 Band Saharanpur Region Image

We have taken a multi-spectral, multi resolution and multi-sensor image of size 641 X 641 of Saharanpur area in Uttar- Pradesh. The satellite image for 6 different bands is taken. Figure 6.2 below shows all the bands.

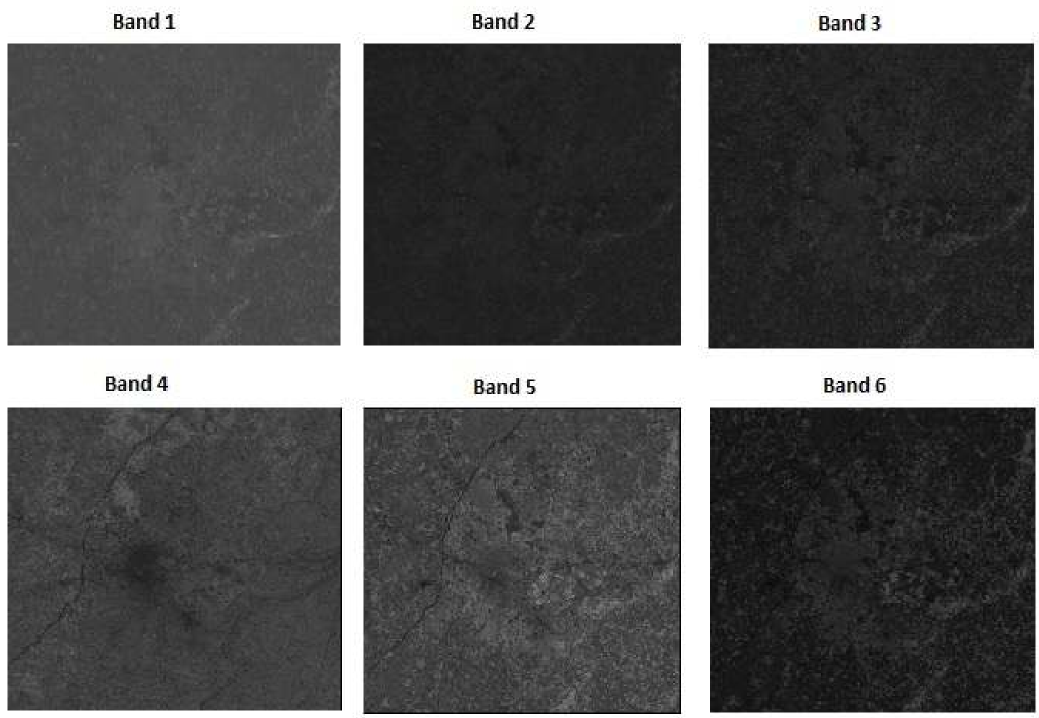


Fig 6.2 The 6 band Multispectral Satellite Images of Saharanpur Dataset

Saharanpur region has mainly the following features:

- Barren
- Dense Vegetation
- Medium Vegetation
- Sparse Vegetation
- Urban region

- Water

Our geosciences experts recognized all these features in Saharanpur region. Thus our aim is to extract all these features from the dataset.

6.2 TRAINING DATASET

A suitable classification system and a sufficient number of training samples are prerequisites for a successful classification. A sufficient number of training samples and their representativeness are critical for supervised image classifications. Training samples are usually collected from fieldwork, or from fine spatial resolution aerial photographs and satellite images. When the landscape of a study area is complex and heterogeneous, selecting sufficient training samples becomes difficult. This problem would be complicated if medium or coarse spatial resolution data are used for classification, because a large volume of mixed pixels may occur. Therefore, selection of training samples must consider the spatial resolution of the remote-sensing data being used, availability of ground reference data, and the complexity of landscapes in the study area.

The image comprises of several land covers like rocky, water, vegetation, open land, barren etc. In a satellite image the regions are not displayed appropriately. So, we need some system expert, to differentiate between these land covers. ERDAS software is used for expert knowledge. In order to classify the image by different techniques and check their accuracy assessment each classified image has been compared with the image generated by expert system. Training set has been generated by this software manually fig 6.3. The individual classes' pixel has been picked from an image and their proper class has been assigned.

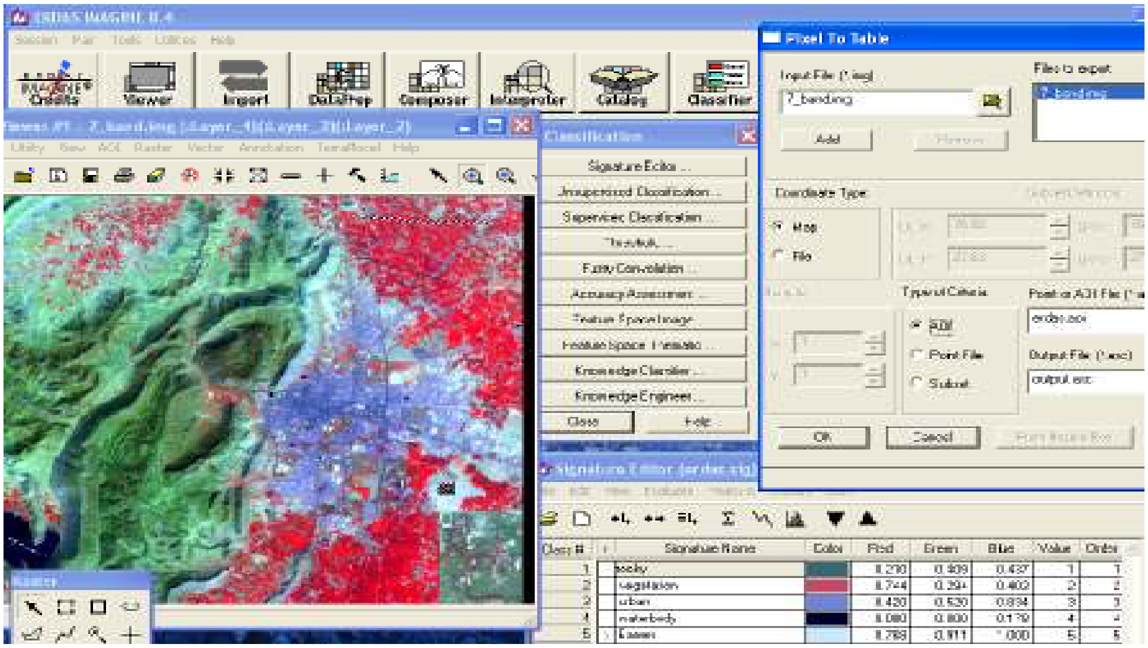


Figure 6.3 Screen shot of ERDAS Image Training set generation

In ERDAS first we have to open the 7 band image in View. Then we will select type of Classification i.e. Supervised Classification from classifier menu. We will open thetool box for Raster Image, it is shown in left bottom corner. Then we will select appropriate tool for the selection of the pixel/ pixels and merge all similar pixels into one class. Class creation has been done in Signature Editor where we merge the selected pixel of one class into one value and correspondingly one ASCII file has been created for it.

6.2.1 Part of Training set of Feature Classes of Alwar Region:

The first two columns show the pixel coordinates followed by the 7 band reflectance values and last is the feature decision.

Table 6.1 Training set for Water Class

27	372	21	28	12	10	2	1	30	WATER
35	372	21	25	10	10	1	0	30	WATER
28	374	23	27	10	12	3	4	30	WATER
30	374	23	28	10	10	4	4	30	WATER
36	374	21	24	10	5	0	2	30	WATER
32	375	23	25	10	7	3	2	30	WATER
40	375	23	27	9	7	2	4	30	WATER
36	376	19	25	9	5	3	2	30	WATER
42	377	17	25	9	9	5	1	40	WATER
46	377	23	25	9	9	2	1	60	WATER

Table 6.2: Training set for Vegetation Class

512	127	17	32	69	246	25	37	10	VEGETATION
479	128	29	40	76	186	23	21	14	VEGETATION
496	129	15	25	60	237	13	26	12	VEGETATION
496	130	15	28	64	219	20	25	12	VEGETATION
501	130	13	30	64	228	28	24	12	VEGETATION
479	133	11	28	60	245	14	23	13	VEGETATION
482	135	15	33	64	245	13	30	13	VEGETATION
506	136	15	33	82	255	15	20	11	VEGETATION
446	139	29	40	74	186	17	32	15	VEGETATION
449	139	21	38	78	188	22	25	15	VEGETATION

Table 6.3: Training set for Urban Class

395	116	115	101	164	137	17	35	15	URBAN
376	117	101	88	149	115	15	20	15	URBAN
387	119	121	103	181	144	16	28	15	URBAN
390	119	146	130	186	146	31	34	15	URBAN
381	120	142	127	166	124	24	24	15	URBAN
383	120	123	112	153	124	24	24	15	URBAN
393	120	125	111	168	126	15	23	15	URBAN
393	124	115	107	159	126	33	38	15	URBAN
383	144	113	106	140	113	44	40	15	URBAN
385	144	113	103	140	111	11	56	15	URBAN

Table 6.4: Training set for Rocky Class

19	15	84	67	171	95	5	57	118	ROCKY
6	17	74	61	140	84	51	20	216	ROCKY
15	18	85	67	168	102	14	88	137	ROCKY
293	18	64	56	115	78	14	22	122	ROCKY
301	18	93	69	138	85	13	36	180	ROCKY
293	20	56	46	104	69	21	26	117	ROCKY
6	21	76	53	153	98	12	22	206	ROCKY
12	22	64	49	153	95	15	80	143	ROCKY
300	22	74	61	124	82	18	22	183	ROCKY
293	25	50	41	100	73	14	24	125	ROCKY

Table 6.5: Training set for Barren Class

103	2	113	88	199	113	16	38	36	BARREN
105	2	117	88	201	118	29	24	30	BARREN
255	2	115	94	175	131	35	36	15	BARREN
74	3	166	130	223	151	32	45	112	BARREN
73	4	168	130	226	149	26	29	114	BARREN
93	4	109	82	179	113	12	21	51	BARREN
99	4	130	99	199	124	37	16	40	BARREN
122	4	136	107	195	151	24	44	21	BARREN
72	5	179	138	232	159	20	36	106	BARREN
74	5	160	133	210	140	49	41	106	BARREN

6.2.2 Part of Training set of Feature Classes of Saharanpur Region:

The first two columns show the pixel coordinates followed by the 7 band reflectance values and last is the feature decision.

Table 6.6 Training set for Barren Class

176505	3319844	117	66	85	79	133	85
176535	3319844	120	67	84	80	137	86
176835	3319784	111	65	83	76	130	83
176865	3319784	111	63	81	75	138	85
176835	3319694	107	61	76	75	128	79
176805	3319664	109	61	77	73	125	79
176295	3319514	112	63	79	76	129	84
176925	3316544	81	42	52	76	102	49
176865	3315584	84	41	45	75	70	28
176655	3315374	83	44	48	82	89	39

Table 6.7 Training set for Dense vegetation Class

168315	3327044	69	29	28	56	49	21
167835	3326654	69	29	27	54	59	21
167865	3326654	68	28	27	54	58	21
167835	3326624	68	28	28	57	58	20
167865	3326624	67	29	27	56	56	19
169065	3326594	66	28	25	59	45	16
169065	3326564	67	27	25	61	51	17
166575	3325064	67	27	25	55	47	17
166605	3325064	69	27	25	52	47	18
166575	3325034	66	27	24	55	45	16

Table 6.8 Training set for Medium vegetation Class

174465	3321314	69	30	26	94	74	24
174435	3321284	69	29	25	93	66	20
168525	3321104	70	29	27	75	59	21
164325	3321074	71	30	28	80	66	24
168495	3321044	70	30	27	75	54	18
168465	3321014	70	29	27	65	55	22
168465	3320984	69	29	27	73	54	19
168465	3320954	69	29	27	74	56	20
168075	3320594	69	28	25	80	62	21
168015	3320534	68	28	25	77	60	19

Table 6.9 Training set for Sparse vegetation Class

165585	3330014	75	38	36	92	70	26
165405	3329984	74	37	35	91	69	25
165135	3329954	76	38	37	83	71	26
165585	3329954	76	38	37	90	69	25
165525	3329894	76	40	36	98	69	25
165555	3329894	76	40	36	98	69	25
166275	3329894	74	37	33	93	65	23
166305	3329894	73	36	33	88	63	23
165165	3329864	74	36	33	91	70	26
165195	3329864	75	37	34	93	70	25

Table 6.10 Training set for Urban Class

167685	3324824	76	34	39	56	83	43
167715	3324824	76	34	38	56	82	42
167745	3324824	74	32	34	57	72	36
167655	3324794	75	34	36	59	74	33
167685	3324794	74	35	37	56	75	36
167715	3324794	76	34	38	55	72	36
167745	3324794	76	35	36	60	78	38
168375	3324614	77	33	38	40	65	41
168405	3324614	75	32	36	48	65	36
168435	3324614	73	32	33	58	62	28

Table 6.11 Training set for Water Class

167955	3329684	74	33	29	26	22	10
167925	3329564	74	32	29	34	35	16
167895	3329534	74	33	29	28	28	10
167835	3329414	76	34	31	36	32	18
167805	3329384	74	34	29	29	24	11
167655	3329204	74	34	30	28	24	12
167085	3328604	72	33	30	35	35	12
166965	3328514	75	34	30	32	28	11
166995	3328514	75	33	29	23	19	9
166815	3328394	75	34	29	24	19	8

6.3 DATA DISTRIBUTION

The distribution of data in our case the reflectance values in different bands is important to consider. The data distribution is very well represented by the standard deviation of values in each band for different feature. Standard deviation gives the seperability of values over a range. We cannot calculate the deviations for whole image as we process the values on the fly and do not store the values. Since the training set are a part of the image and they consider pixels randomly from the image, we find the data distribution of the training set of each alwar and Saharanpur region. We read the excel files of each feature of a region and then apply malab inbuilt function 'std(x)' for calculation. If X is a matrix, std(X) returns a row vector containing the standard deviation of the elements of each column of X.

We got the following tables of standard deviations for each region:

Table 6.12 – Data distribution of Alwar Training set

Feature	Red	Green	NIR	MIR	RS1	RS2	DEM
Water	2.3389	3.8889	5.0069	3.5688	20.1005	9.1508	11.965
Vegetation	5.7831	4.9832	12.4433	22.9663	8.4227	7.4006	13.4375
Urban	13.7826	11.4664	16.2423	14.1736	48.4533	56.2948	2.3595
Rocky	16.8211	13.4537	23.6914	15.3246	11.5498	14.9548	46.2775
Barren	18.1582	16.3737	15.9498	15.1567	14.163	16.1108	37.8196

Table 6.13 – Data Distribution of Saharanpur Training set

Feature	Band1	Band2	Band3	Band4	Band5	Band6
Barren	14.4866	10.6045	15.1123	3.5442	21.3991	19.183
Dense_veg	1.1893	1.0208	1.8792	2.5942	4.2743	2.0131
Medium_veg	1.7803	1.4733	1.7768	6.0037	5.9216	2.4058
Spar_veg	1.1153	1.3442	1.8439	4.4763	3.0952	1.974
Urban	3.1755	1.7276	3.0993	7.5572	8.255	4.0494
Water	1.3988	0.5	0.6272	6.2295	6.6375	2.8431

Graphical representation of Data distribution

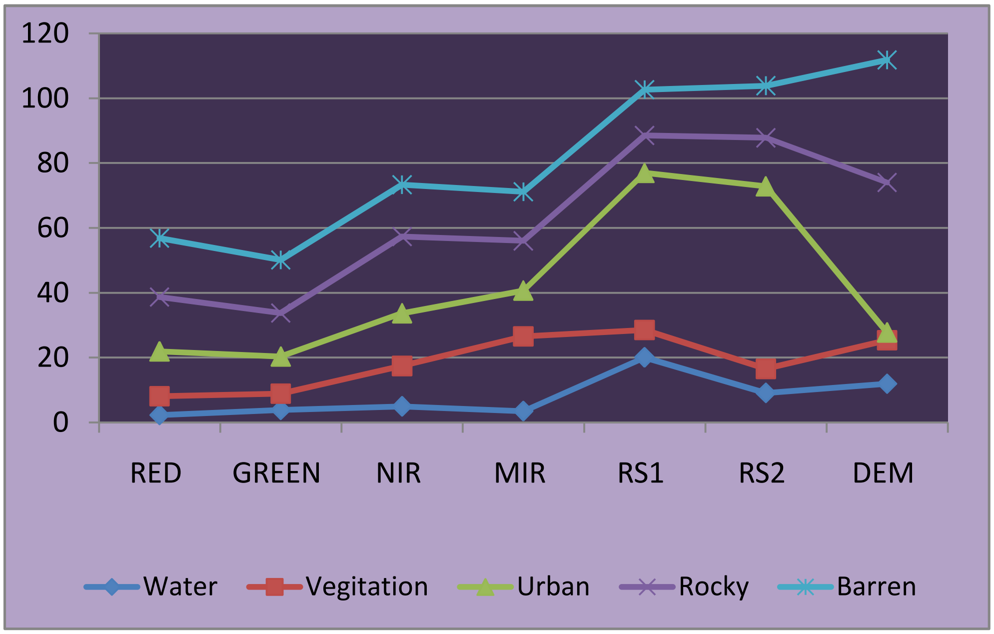


Figure 6.4 Standard Deviation of different Classes in Alwar Dataset

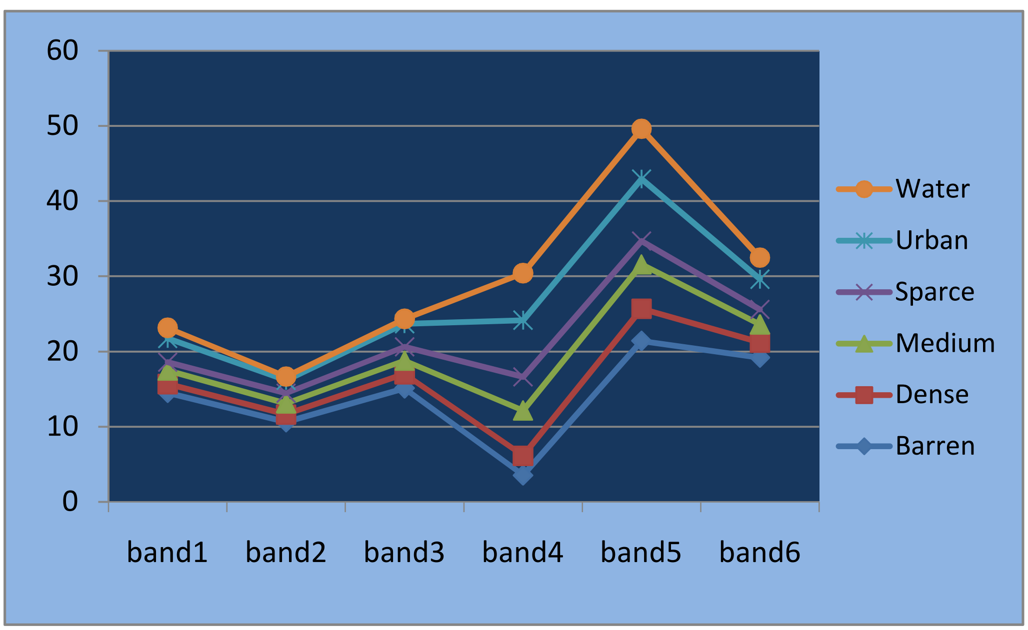


Figure 6.5 Standard Deviation of different Classes in Saharanpur Dataset

6.4 CLASSIFICATION RESULTS

We have applied our classification technique on two different dataset and it has shown good results with both dataset, confirming its applicability for different features. First we have shown our result on Alwar region and then on Saharanpur region. Our objective is to use the proposed ABC based algorithm as an efficient Land cover classifier for satellite image and extract these features optimally.

6.4.1 Classification Results of Alwar Image

After applying the proposed algorithm to the Image, the classified image is obtained .We have different features with colors for distinguish ability. For the Alwar Classified output image colors, blue, green, pink, red and black color represents water, vegetation, urban, rocky and barren region respectively. Figure 6.6 shows the satellite false color image and Figure 6.7 classified image of Alwar region after applying ABC.

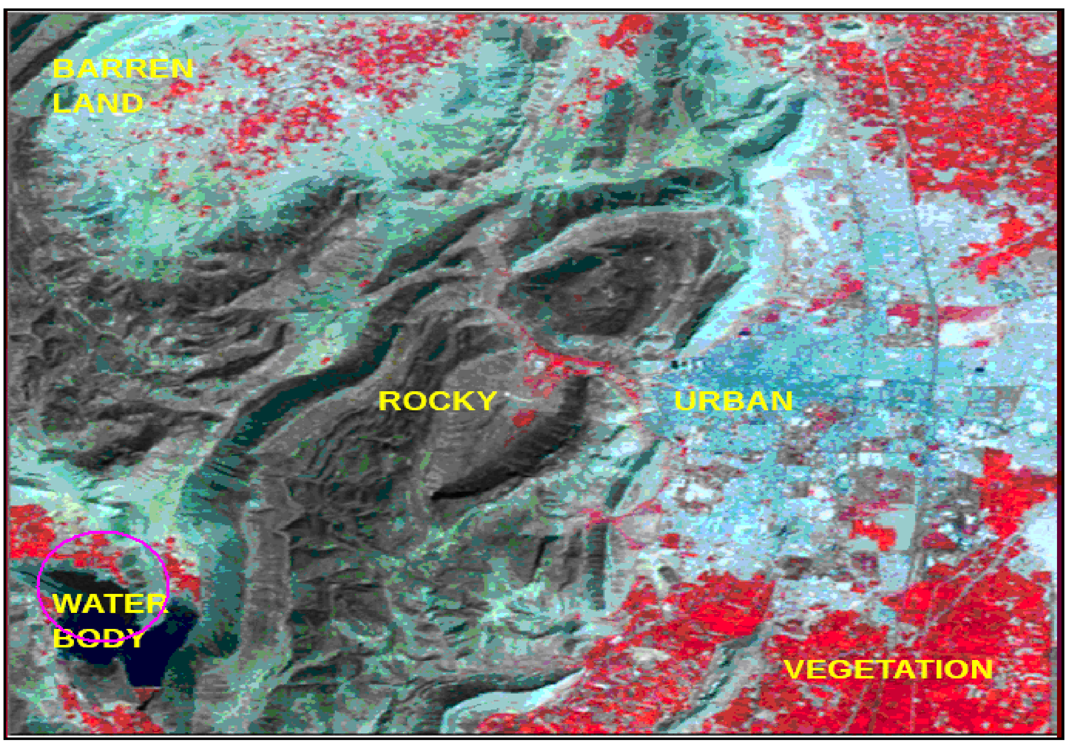


Figure 6.6 Original false Color Image of Alwar

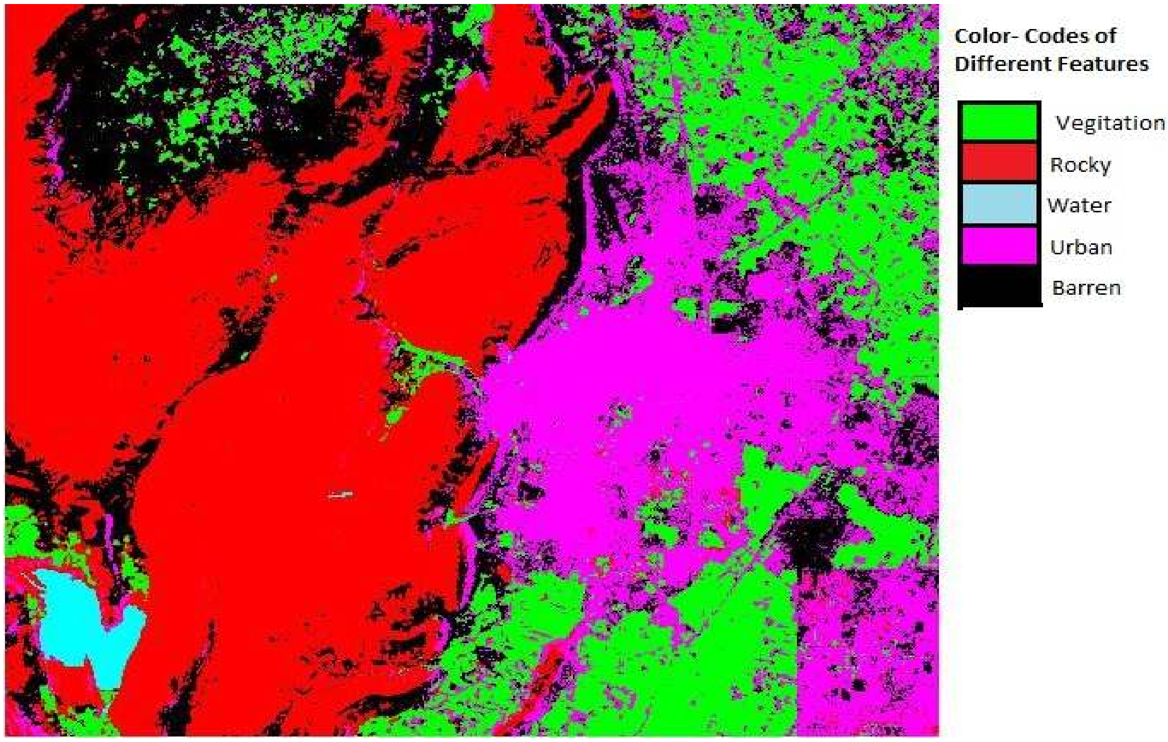


Figure 6.7 Classified output of Alwar

Comparing the original image and our classified output we can clearly see the good correlation between the two. All the features namely water, vegetation, urban, rocky and barren are well classified.

6.4.2 Classification results of Saharanpur Region

Saharanpur dataset had 6 band images and we had to identify 6 land features namely barren, dense vegetation, medium vegetation, sparse vegetation, urban and water. Saharanpur district is in Uttar-Pradesh state of India which is mainly famous for its sugarcane and other crop harvesting. It is a well habited area with population focussed on some regions while crops on the outskirts. It didn‘t have any rocky region as in the case of Alwar whereas a different types of vegetation densities exits. The colour codes for Saharanpur region are different. The red, black, green, yellow, pink and blue represents barren, dense vegetation, medium vegetation , sparse vegetation, urban and water region. Figure 7.8 shows the original false color image generated by the experts and figure 7.9 shows our classification output.

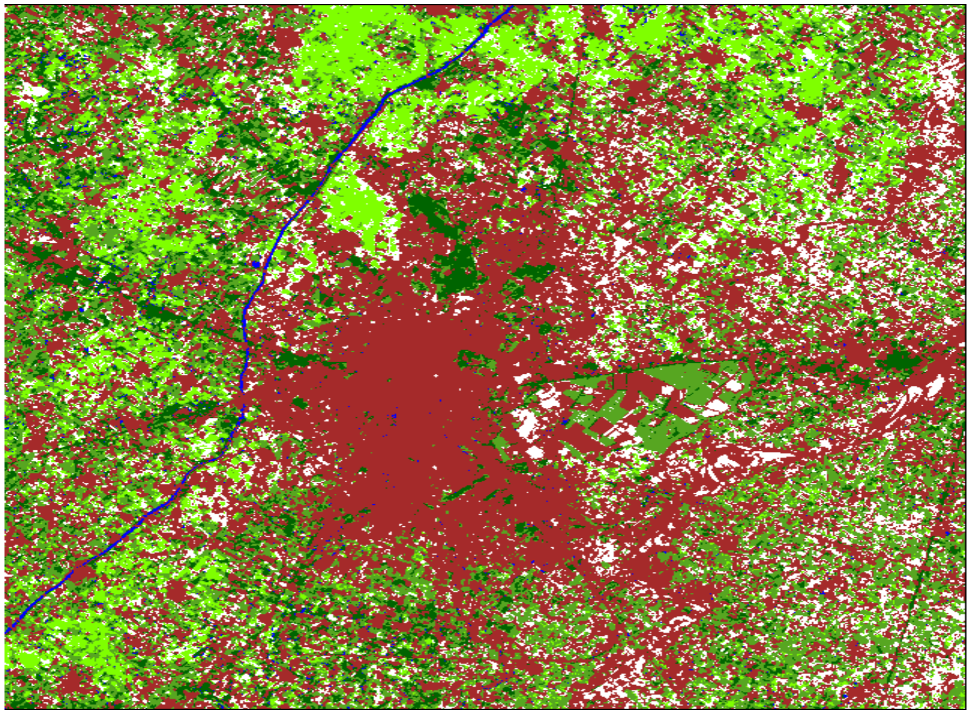


Figure 6.8 Original false color image of Saharanpur

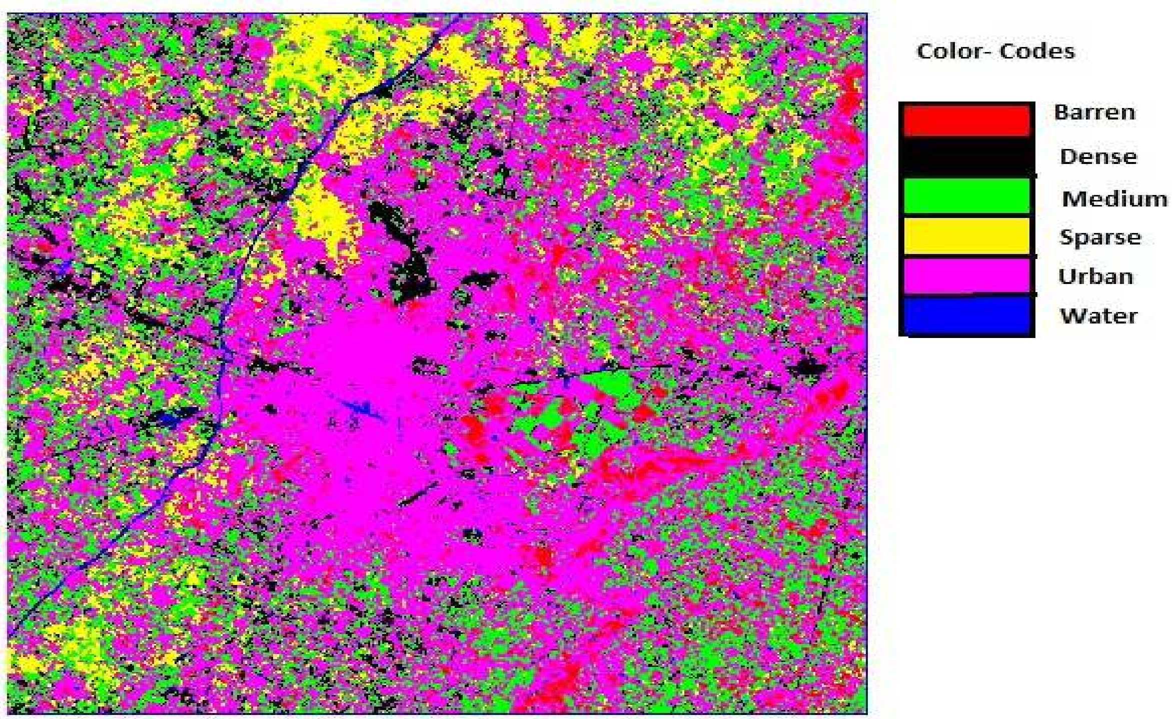


Figure 6.9 Classified Output of Saharanpur

From the above two figures we observe high correlation and notice that all the 6 features have been very well extracted.

6.5 FEATURE EXTRACTION FROM HETEROGENEOUS REGIONS

An image is a collection of pixels. Remote sensing image which captures the land covers are a major study for analysing different terrain features. Land is composed of so many features and hence it becomes very important that when we do land cover mapping of these features we are able to identify each feature correctly. Natural features on the land do not have clear distinct boundaries. Many of these features are overlapping or occurring together in some areas. Heterogeneous regions in an image are small areas in the image where a combination of many features is co-existing together. Such regions are characterised by rapid fluctuations in reflectance values in a very small range. These are often called conflicting regions [47] since a clear identification of each feature individually becomes cumbersome. The separating of features from heterogeneous regions is important to increase the classification accuracy. Our proposed ABC technique is able to extract the features from both the input images of Alwar and Saharanpur. We have marked two heterogeneous regions in each of them and then shown the zoomed output image of those regions.

6.5.1 Feature Extraction In Alwar Region

Figure 6.10 shows our classified Alwar output with two heterogeneous marked and also zoomed to show that the features have been well distinguished. The region on the right has mixture of all the 5 features of water, vegetation, rocky, urban and barren. The other has all except water. Hence we can conclude that our classification for Alwar is complete and we have been able to successfully mark all the features even from the conflicting ones.

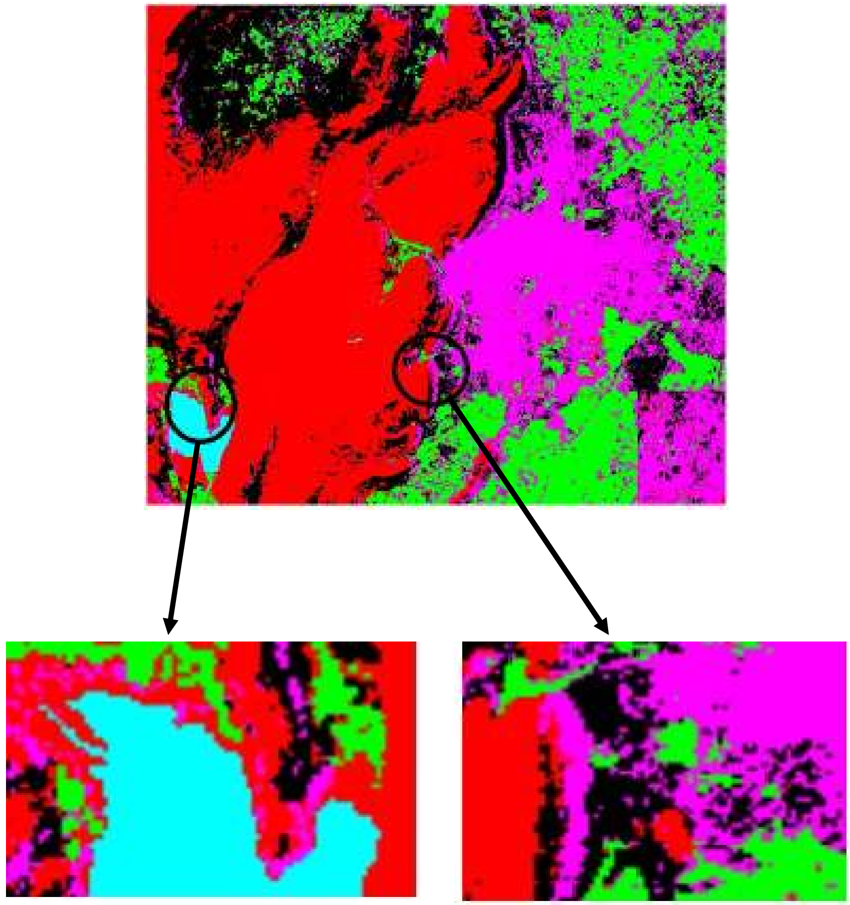


Figure 6.10 Alwar image showing heterogeneous regions identified

As we can see from above that these two small regions have many features together and from the zoomed portions we can infer that all the features were identified uniquely. Hence we can conclude that our classification for Alwar is complete and we have been able to successfully mark all the features even from the conflicting ones.

6.5.2 Feature Extraction in Saharanpur Region

Figure 6.11 deals with the heterogeneous regions of the Saharanpur image. Two regions have been marked and each has a combination of features to be identified. In Saharanpur the problem comes in separating three types of vegetation densities dense, medium and sparse. But as we can see below all features have been extracted. and each pixel has been identified to the class it belongs.

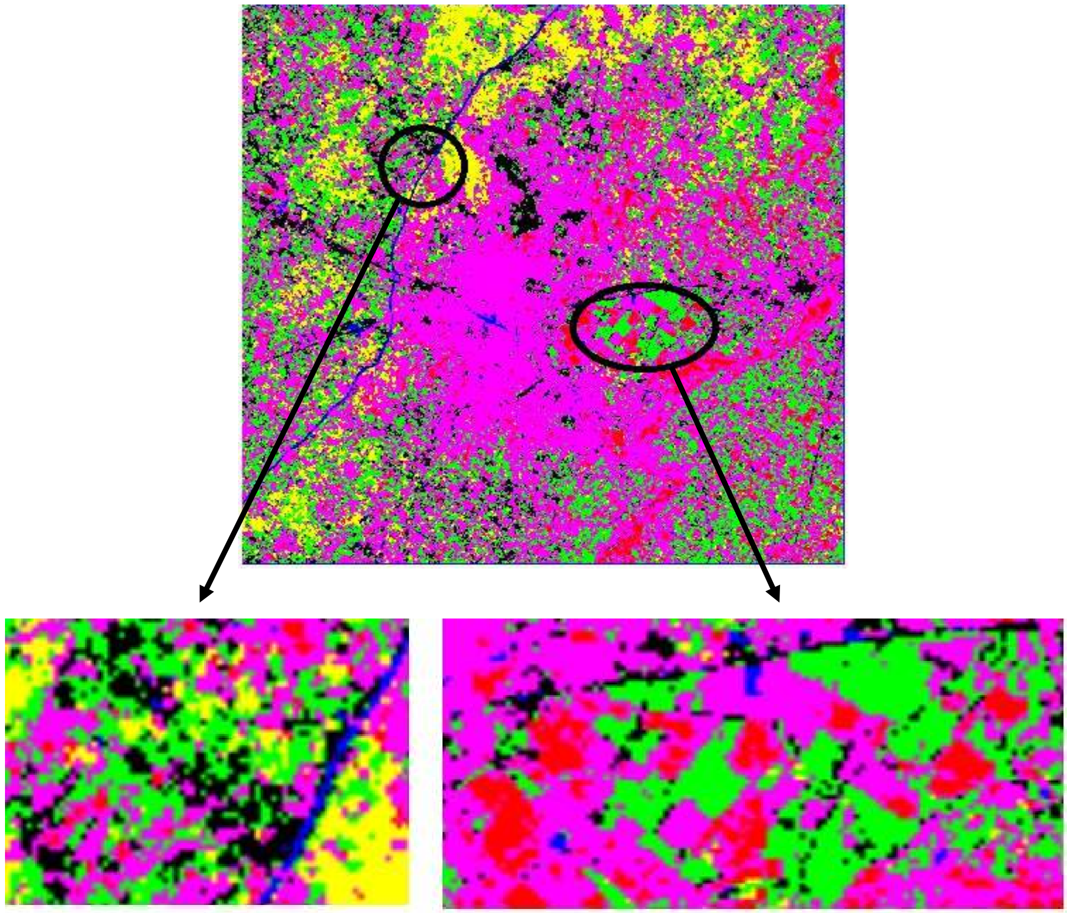


Figure 6.11 Saharanpur image showing heterogeneous regions identified

From both the above observations we can prove that the ABC classifier not just classifies the homogeneous regions but also the heterogeneous. All features have been extracted and each pixel has been identified to the class it belongs.

7. EVALUATION AND ANALYSIS

OVERVIEW

Evaluation of classification results is a significant step in the classification procedure. Many different approaches may be used, ranging from a qualitative evaluation based on expert knowledge to a quantitative accuracy assessment based on sampling strategies. To evaluate the performance of a classification method, Cihlar et al. (1998) proposed six criteria, accuracy, reproducibility, robustness, ability to fully use the information content of the data, uniform applicability, and objectiveness[40]. But practically due to the environmental setting and the datasets used, no classification algorithm can satisfy all these requirements nor be applicable to all studies. Classification accuracy assessment is, however, the most common and accepted approach for an evaluation of classification performance.

7.1 ACCURACY ASSESSMENT

Before we implement the classification accuracy measurement we should know the sources of probable errors. They are:

- Poor Classification scheme

- Interpretation errors

- Poor training dataset

It is commonly assumed that the difference between an image classification result and the reference data is due to the classification errors. From reference truth data we get assumed data. It is not feasible to validate every pixel of a classified image hence a set of reference pixels is used. Reference pixels are points on the classified image for which actual features are known. These are also known as validation dataset generated by the experts. The reference pixels are randomly selected. The goal of accuracy assessment to quantitatively determine how effectively pixels were grouped into the correct feature classes in the area under investigation.

7.1.1 Error Matrix

The error matrix approach is the one most widely used in accuracy assessment. In order to properly generate an error matrix, we must consider the following factors: (1) reference data collection, (2) classification scheme, (3) sampling scheme, (4) spatial autocorrelation, and (5) sample size and sample unit. Error matrices compare, on a category by category basis, the relationship between known reference data (ground truth) and the corresponding results of an automated classification. The columns usually represent the reference data while the rows indicate the classification generated from the remotely sensed data. An error matrix is a very effective way to represent accuracy in that the accuracies of each category are plainly described along with both the errors of inclusion (commission errors) and errors of exclusion (omission errors) present in the classification.

For the validation process of the Alwar classification we have taken the validation dataset which has following number of pixels:

- 68 water pixels
- 109 vegetation pixels
- 139 urban pixels
- 96 rocky pixels
- 63 barren pixels

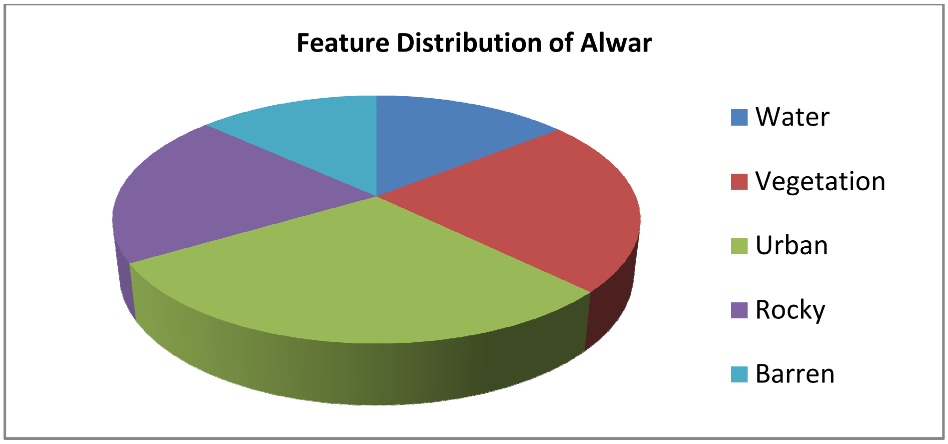


Figure 7.1- Feature distribution of Alwar validation dataset

Table 7.1 Error Matrix of Alwar

Feature	water	Veg	Urban	Rocky	Barren	Total
Water	68	0	0	0	0	68
Veg	0	109	1	0	0	110
Urban	0	0	112	0	4	116
Rocky	0	0	0	96	0	96
Barren	0	0	26	0	59	85
Total	68	109	139	96	63	475

Analysis from Error Matrix of Alwar:

From the training set and the error matrix which is obtained after implementing our classifier ABC is shown in Table 7.1. The error matrix‘s interpretation along column suggests how many pixels are classified correctly by our algorithm. The diagonal elements depict number of properly classified pixels in that category. We can infer that for the given dataset we have been able to classify water, vegetation and rocky areas perfectly. It can however be improved for barren and urban region.

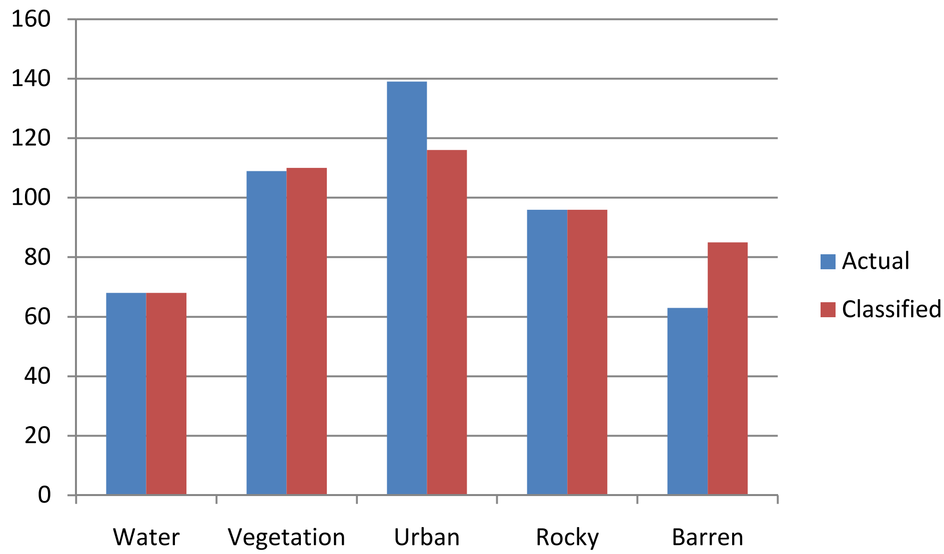


Figure 7.2 Comparison between data given by expert and result obtained

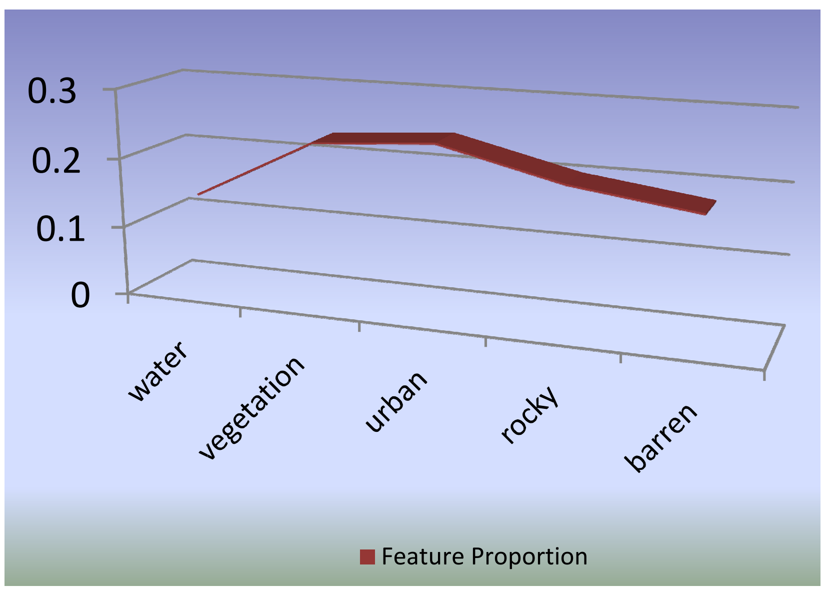


Figure 7.3 Feature Proportion of Alwar extracted

For the validation process of the Saharanpur classification we have taken the validation dataset which has following number of pixels:

- 15 Barren pixels
- 24 Dense vegetation pixels

- 35 medium vegetation pixels
- 64 sparse vegetation pixels
- 70 Urban pixels
- 13 Water pixels

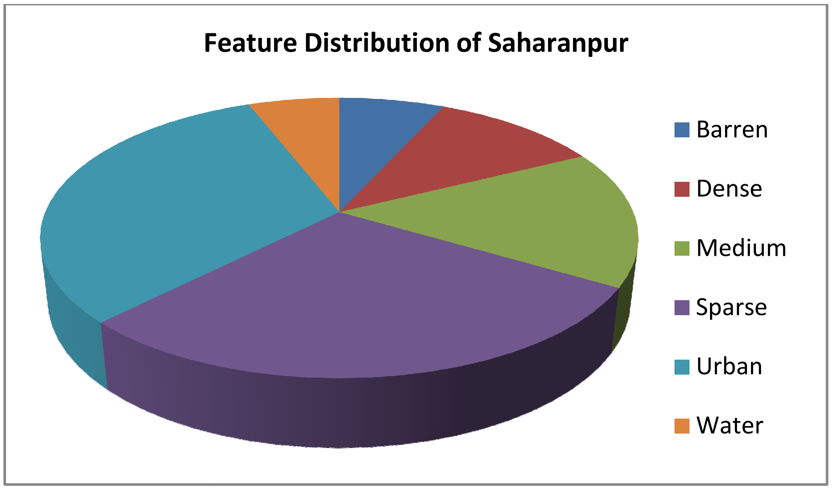


Figure 7.4- Feature distribution of Saharanpur validation dataset

Table 7.2 Error Matrix of Saharanpur

Feature	Barren	Dense	Medium	Sparse	Urban	Water	Total
Barren	11	0	0	0	0	0	11
Dense	0	24	0	0	0	0	24
Medium	0	0	34	0	0	0	34
Sparse	0	0	1	64	0	0	65
Urban	4	0	0	0	70	1	75
Water	0	0	0	0	0	12	12
Total	15	24	35	64	70	13	221

Analysis from Error Matrix of Saharanpur:

From the training set and the error matrix which is obtained after implementing our classifier ABC is shown in Table 7.2 we can infer that most of the pixels have been successfully extracted. Water region has been perfectly classified. Out of total 221 validation pixels taken, 214 of them have been classified in the right category.

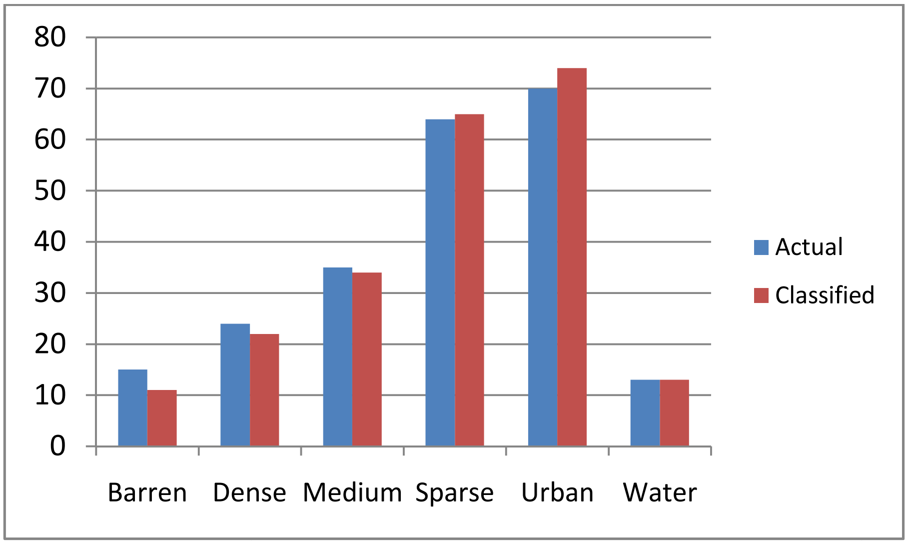


Fig. 7.5 Comparison between data given by expert and result obtained

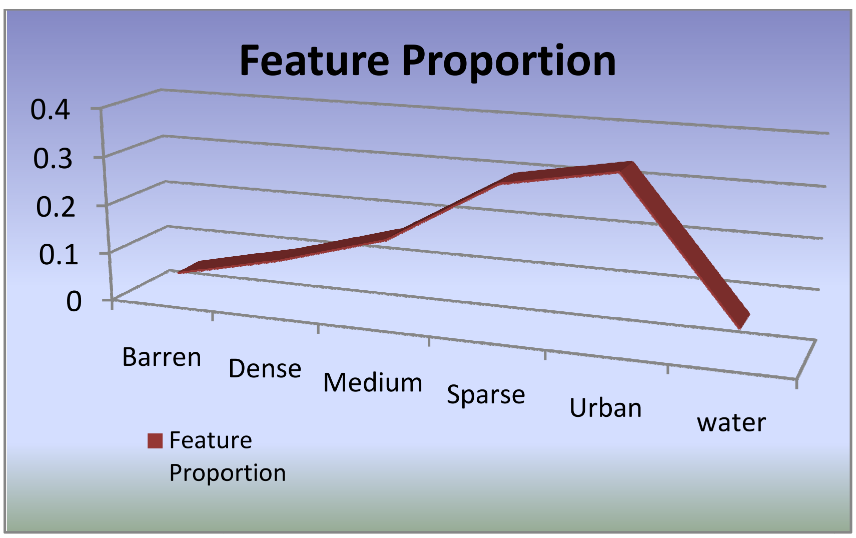


Figure 7.6 Feature Proportion of Saharanpur extracted

7.1.2 *Omission and Commission Errors:*

An omission error means that we omitted or missed some feature pixels. A commission error means that we committed (predicted) pixels erroneously to a specific class. Both omission and commission and overall error are important to evaluate the error in classification.

Omission error = 1- (total correct pixels in the class / total pixels in that category in reference data)

Commission error = 1- (total correct pixels in the class / total pixels that got classified in that class)

Table 7.3 Omission and Commission error of Alwar Error Matrix

Feature	Omission error	Commission error
Water	0	0
Vegetation	0	0.00909
Urban	0.1942	0.0344
Rocky	0	0
Barren	0.06349	0.3058

Table 7.4 Omission and Commission error of Saharanpur Error matrix

Feature	Omission error	Commission error
Barren	0.266	0
Dense	0.0833	0
Medium	0.02857	0.0555
Sparse	0	0.0153
Urban	0	0.0541
Water	0	0

7.1.3 Producer's Accuracy

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. It is called so because the producer of the classification is interested in how well a certain area can be classified. Producer‘s accuracies(as shown in table 7.5) 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). The error counterpart of PA is the omission error.

Table 7.5 Producer's Accuracy for Alwar Region

Feature	PA calculation	Producer Accuracy
Water	68/68*100	100%
Vegetation	109/109*100	100%
Urban	112/139*100	80.57%
Rocky	96/96*100	100%
Barren	59/63*100	93.65%

Table 7.6 Producer's Accuracy for Saharanpur Region

Feature	PA Calculation	Producer's Accuracy
Barren	11/15*100	73.33%
Dense	22/24*100	91.66%
Medium	34/35*100	97.14%
Sparse	64/64*100	100%
Urban	70/70*100	100%
Water	13/13*100	100%

7.1.4 User's Accuracy

User‘s Accuracies 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 shown in previous section and indicates the probability that a pixel classified into a given category actually represents that category on the ground.

Table 7.7 User's Accuracy for Alwar region

Feature	UA Calculation	User's Accuracy
Water	68/68*100	100%
Vegetation	109/110*100	99.09%
Urban	112/116*100	96.55%
Rocky	96/96*100	100%
Barren	59/85*100	69.94%

From the above table we can infer that all regions have been very well classified, only barren losses out some accuracy.

Table 7.8 User's Accuracy for Saharanpur region

Feature	UA Calculation	User's accuracy
Barren	11/11*100	100%
Dense	22/22*100	100%
Medium	34/36*100	94.44%
Sparse	64/65*100	98.46%
Urban	70/74*100	94.59%
Water	13/13*100	100%

7.2 KHAT Statistics

Another discrete multivariate technique of use in accuracy assessment is called KAPPA (Cohen, 1960). The result of performing a KAPPA analysis is a KHAT statistic (an estimate of KAPPA), which is another measure of agreement or accuracy. The ˆk ("KHAT") statistic is a measure of the difference between the actual agreement between reference data and an automated classifier and the chance agreement between the reference data and a classifier. The KHAT statistic is computed as The Kappa statistic is a very useful statistic. It describes or estimates how much better or worse than our classification is, relative to chance alone.

Conceptually, ^*k* can be defined as:

ˆk= (observed accuracy- chance agreement) / (1- chance agreement)

This statistic serves as an indicator of the extent to which the percentage correct values of an error matrix are due to "true" agreement versus "chance" agreement. As true agreement (observed) approaches 1 and chance agreement approaches 0, its value approaches 1. This is the ideal case. In reality, its value usually ranges between 0 and 1. For example, a value of 0.98 of KHAT statistics can be thought of as an indication that an observed classification is 98 per cent better than one resulting from chance. A KHAT coefficient with the value of 0 suggests that a given classification is no better than a random assignment of pixels. In cases where chance agreement is large enough, KHAT coefficient can take on negative values-an indication of very poor classification performance. (Because the possible range of negative values depends on specific matrix, the magnitude of negative values should not be interpreted as an indication of relative classification performance). The principle advantage of computing KHAT coefficient is the ability to use this value as a basis for determining the statistical significance of any matrix or the differences among matrices. The KHAT coefficient incorporates the non-diagonal elements of error matrix (and hence error of omission and commission) as a product of the row and column marginal. The khat coefficient can be computed as:

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r}(x_{i+} \cdot x_{+i})}$$

r = number of rows in the error matrix

$\boldsymbol{x}_{ii}$ = the number of observations in row i and column i (on the major diagonal)

$\boldsymbol{x}_{i+}$ = total of observations in row i (shown as marginal total to right of the matrix)

$\boldsymbol{x}_{+i}$ = total of observations in column i (shown as marginal total at bottom of the matrix)

N = total number of observations included in matrix.

KAPPA Computations for Alwar Region:

N= 475

$\sum x_{ii}$ =68+109+112+96+59= 444

N($\sum x_{ii}$) = 475*444=210900

$\sum(x_{i+} * x_{+i})$ =(68*68)+(109*110)+(139*116)+(96*96)+(63*85) = 47309

N^2 = 475*475 = 225625

ˆk(KAPPA) =(210900-47309) / (225625-47309) = 0.9174

KAPPA Computations for Saharanpur region:

N= 221

$\sum x_{ii}$ =11+22+34+64+70+13=214

N($\sum x_{ii}$) = 221*214=47294

$\sum(x_{i+} * x_{+i})$ =(11*15)+(22*24)+(34*35)+(65*64)+(74*70)+(13*13) = 11426

N^2 = 221*221 = 48841

ˆk(KAPPA) =(47294-11426) / (48841-11426) = 0.9586

Hence the KHAT statistics of the proposed ABC computing based land cover feature extractor is 0.9174 for Alwar dataset and 0.9586 for Saharanpur dataset. This value is a substantial agreement between the ideal classifier and the proposed algorithm. This indicates that an observed classification is 91.74% better for Alwar and 95.86% better for Saharanpur than one resulting from chance.

7.3 COMPARATIVE ANALYSIS OF PERFORMANCE INCLUDING HETEROGENEITY

With the advent of more advanced digital satellite remote sensing techniques, it is becoming more necessary to perform accuracy assessment of these techniques. Accuracy and performance assessment has hence received a great deal of renewed interest. This is not to say that accuracy assessment is unimportant for the more traditional remote sensing techniques. However, given the complexity of digital classification, there is more of a need to assess the reliability of the results of the different techniques.

We have seen in previous sections the problem of extracting features from heterogeneous regions. In the judgement of performance this factor is also taken into consideration that how well a particular technique deals with such conflicting regions. A comparative review is necessary to strengthen our belief on the proposed algorithm. We consider for comparison only one dataset that is of Alwar. So that it gives uniform criteria of evaluation.

7.3.1 Overall Accuracy Based on KAPPA Coefficient

KAPPA coefficient is a very good performance evaluation parameter. This is because most of the classification techniques use it as a standard and give the overall classification accuracy. We compare the results of our proposed artificial bee colony with some other implemented classification techniques, we have considered kappa coefficient as the parameter against which performance analysis is done. The techniques considered here are Minimum Distance Classifier (MDC), Maximum Likelihood Classifier (MLC), Biogeography Based Optimization (BBO), Membrane Computing (MC) , Hybrid Flower Pollination by Artificial Bees and Bacteria Foraging Optimization (FPAB/BFO) and Fuzzy classifier with Kappa of 0.7364, 0.7525, 0.6715, 0.68812, 0.6874 and 0.9134 respectively [13,14]. However, ABC with k= 0.917 shows our classifier has good classification accuracy compared to other soft computing techniques. Figure 7.7 shows the graphical representation of this comparison.

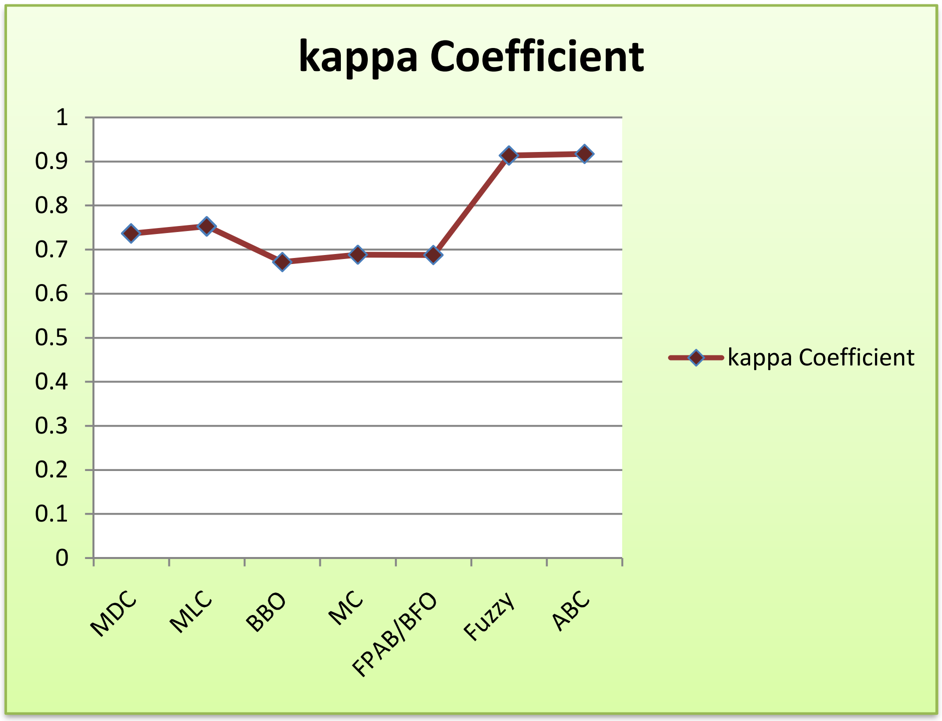


Figure 7.7 Comparison of kappa Coefficients

7.3.2 Evaluation Based on Heterogeneous region classification

Recognition of heterogeneous regions and its accurate classification is a challenge for many techniques. Most techniques focus on the global model and overlook these regions which have a great impact on the overall classification accuracy. In this section we analyse some soft computing techniques and their efficiency in extracting all features from the heterogeneous regions. Even though they might have high KAPPA coefficients yet they are not able to correctly extract the different features from these regions. A very important characteristic of our algorithm is that it works pixel wise and not cluster wise, like BBO uses fuzzy-c-means cluster to group all the similar pixels together. So ABC classifier uniquely identifies each pixel and is able to separate the features from the heterogeneous areas. We have studied a number of classifiers to see how they have handled the heterogeneous regions and weather they are able to extract all the features of them.

Figure 7.8(a) below shows the output of Alwar image from our artificial bee colony classifier and two regions of conflict. We can clearly see that all the features namely water, urban, vegetation, rocky and barren have been successfully separated and uniquely identified.

We have considered many soft computing techniques to compare the performance of our ABC classifier for identifying features from heterogeneous regions. We first take two traditional techniques of classification namely Minimum Distance to Mean classifier (MDC) and Maximum Likelihood Classifier (MLC). Figure 7.8(b) shows their results of heterogeneity. MDC and MLC have not been able to identify all the features from the marked heterogeneous region. Both of the classifier has Kappa coefficient for Minimum distance it is 0.7364 and for Maximum likelihood it is 0.7525.

The fuzzy classification is a popular technique and when applied to the given dataset it had kappa coefficient of 0.9134. Figure 7.8(c) shows the result. The fuzzy was not able to properly extract all the features from the region. The result shows that it classified most pixels as vegetation, some as urban but none for rocky and very few barren pixels.

In Rough- Fuzzy Classification technique the fuzzy theory was combined with the rough set theory to achieve a Kappa Coefficient of 0.9700, which is higher than our overall accuracy of 0.9174. Still it is not able to accurately classify the heterogeneous region. Figure 7.8(d) gives the rough-fuzzy tie up result .It classified most of pixels as rocky, very few as barren and none as urban and vegetation.

BBO is an important classification technique. We studied its performance in analysis or kappa. It has kappa coefficient of 0.6715 which is less for an optimization classification. Figure 7.8(e) shows the result. BBO is implemented as a cluster based approach and its efficiency is very less. It has not been able to classify the heterogeneous regions properly.

Membrane Computing technique has been recently explored [28] and has a kappa coefficient of 0.68812. Though it is a pixel based approach yet some of the heterogeneous regions are not well classified. Fig 7.8(f) shows the results

Classification using hybrid approach of Flower Pollination of artificial bees and Bacterial Foraging optimization was done with the Kappa coefficient of 0.68745 [41]. Figure 7.8(g) shows the results. The features of two heterogeneous regions marked were not identified properly.

Fig 7.8(h) shows the result after applying hybrid ACO2/BBO/PSO with Kappa coefficient of 0.9812. Fig 7.8(i) applies the cAntMiner Algorithm on the Alwar Region which has a Kappa Coefficient of 0.964. Fig 7.8(j) shows the result of applying the hybrid ACO-BBO Technique on the Alwar Image which has a Kappa-Coefficient of 0.96699.

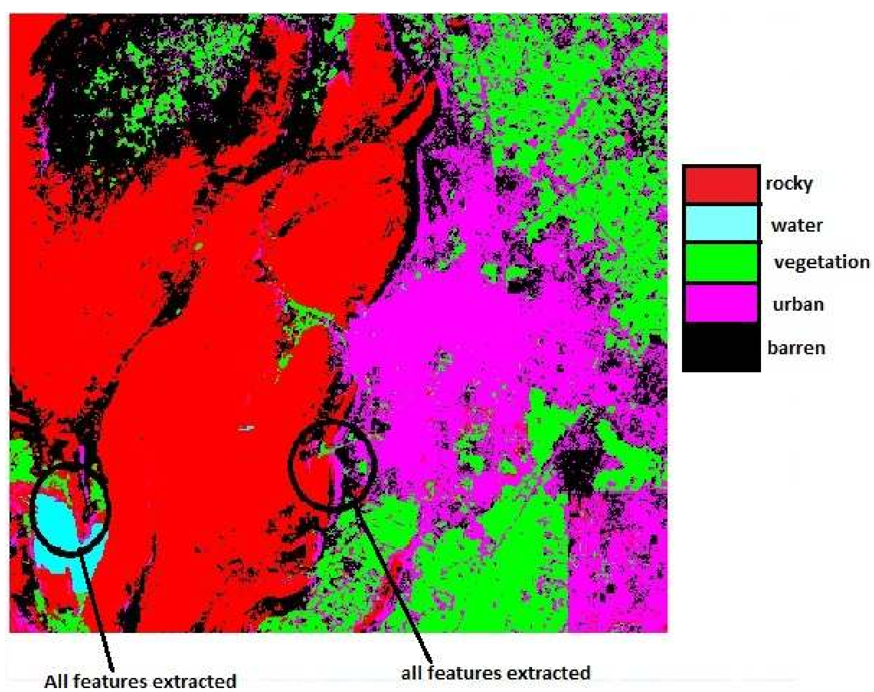


Figure 7.8(a) Classification results using ABC

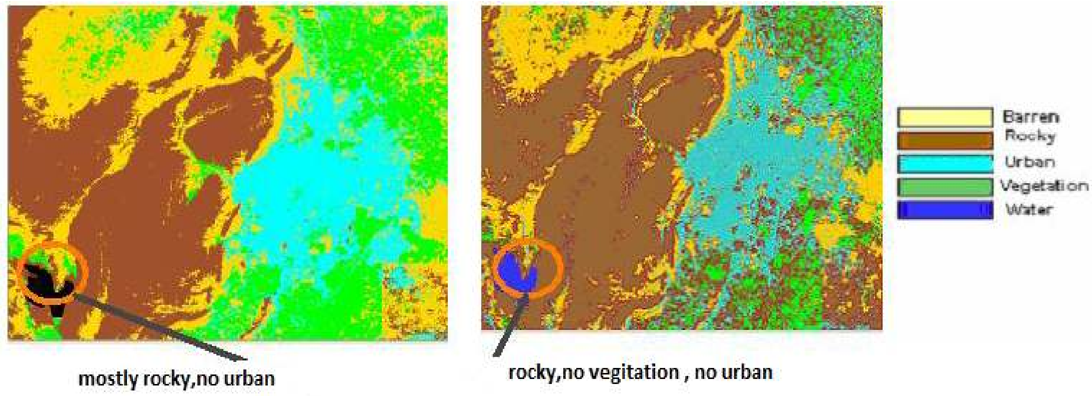


Figure 7.8(b) MDC classification (left) , MLC classification

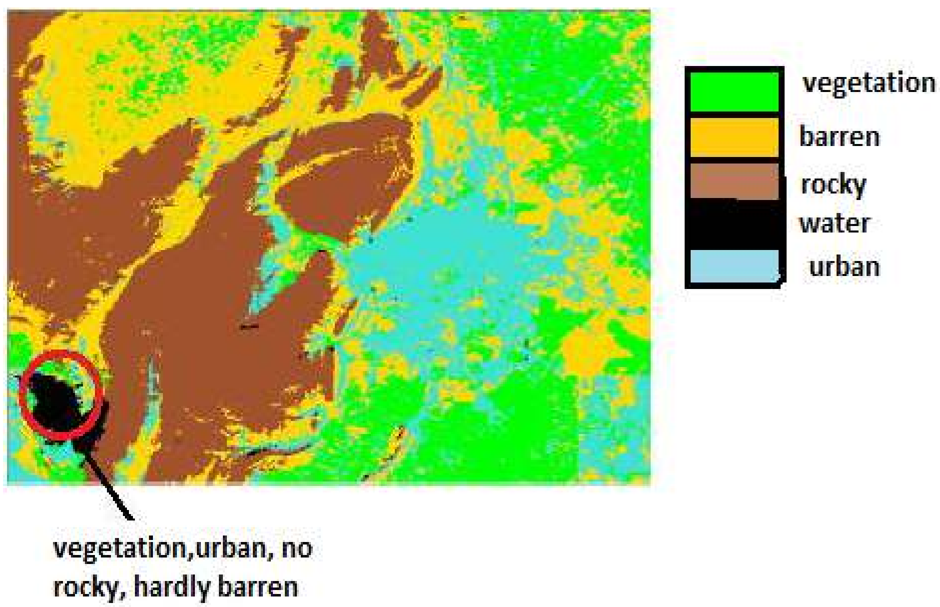


Figure 7.8(c) Fuzzy classification result

The above fuzzy classification is a popular technique, when applied to the given dataset it had kappa coefficient of 0.9134. It classified most pixels as vegetation, some as urban but none for rocky and very few barren pixels. In Rough- Fuzzy Classification technique the fuzzy theory was combined with the rough set theory to achieve a Kappa Coefficient of 0.9700, which is higher than our overall accuracy of 0.9174. But it is not able to accurately classify the heterogeneous region. It classified most of pixels as rocky, very few as barren and none as urban and vegetation. Fig. 7.8 (d) shows the result.

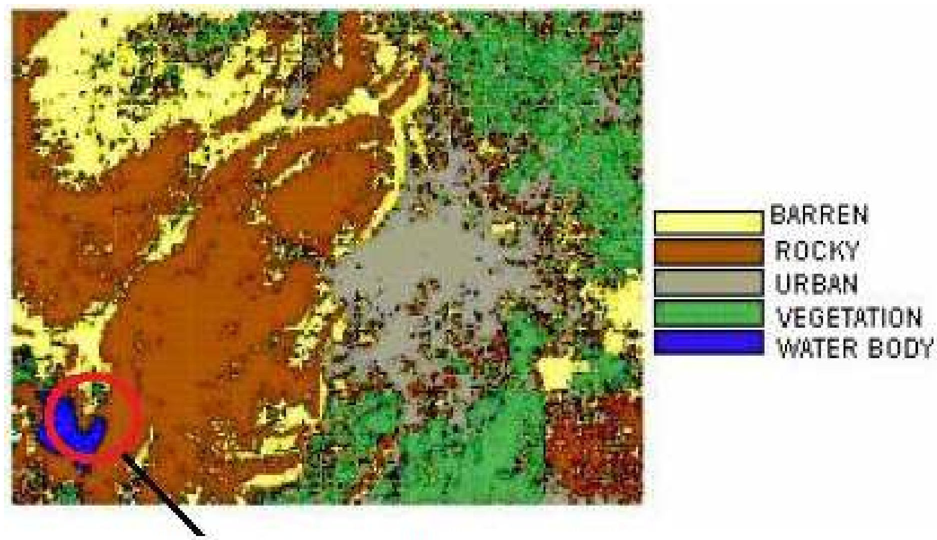


Figure 7.8(d) Classification result of Rough-Fuzzy

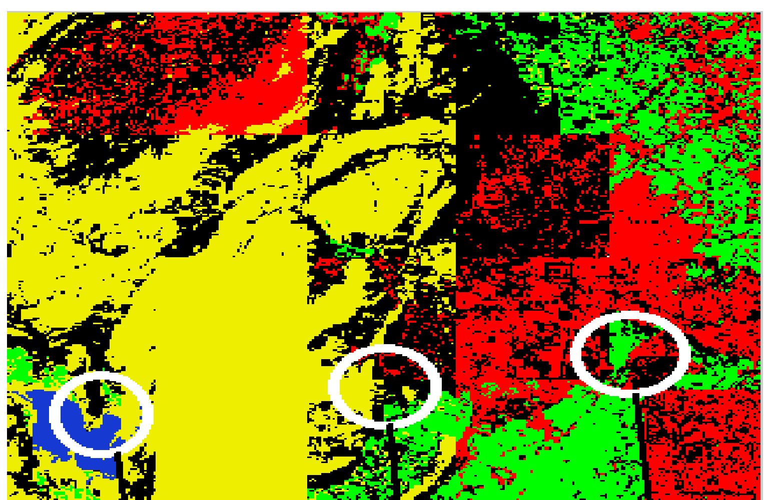


Figure 7.8(e) BBO classification result

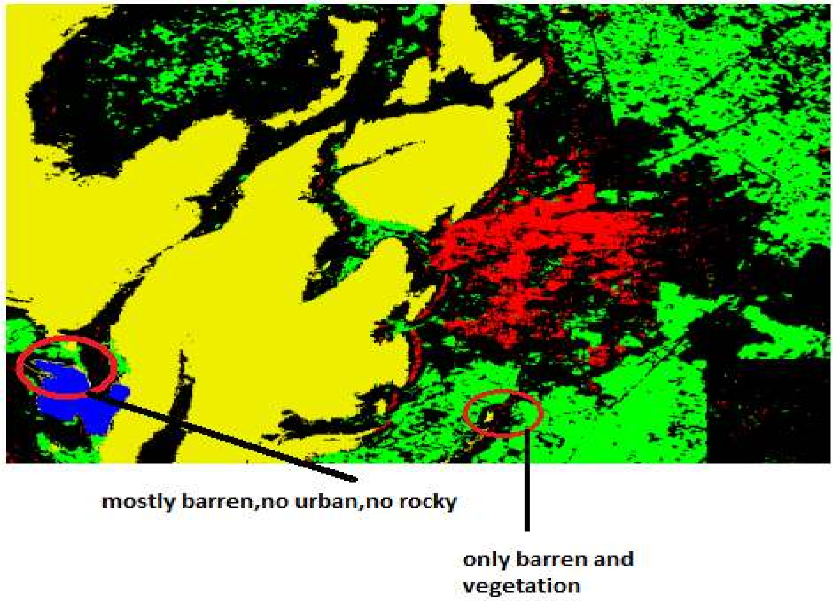


Figure 7.8(f) Classification result of Membrane computing

Membrane computing technique has been recently explored [7] and has a kappa coefficient of 0.68812. Though it is a pixel based approach yet some of the heterogeneous regions are not well classified. In the right region marked it was not able to identify any urban and rocky features and the other region only vegetation and barren got identified.

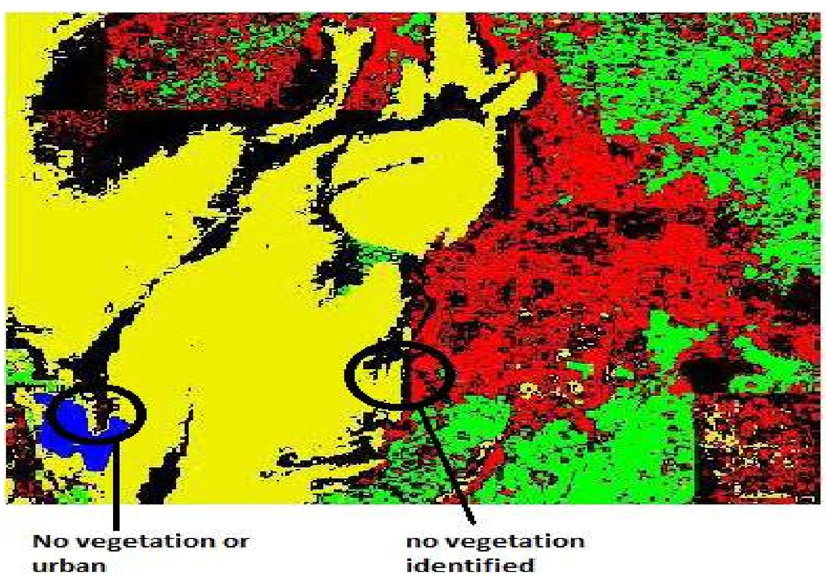


Figure 7.8(g) Classification result of FPAB/BFO

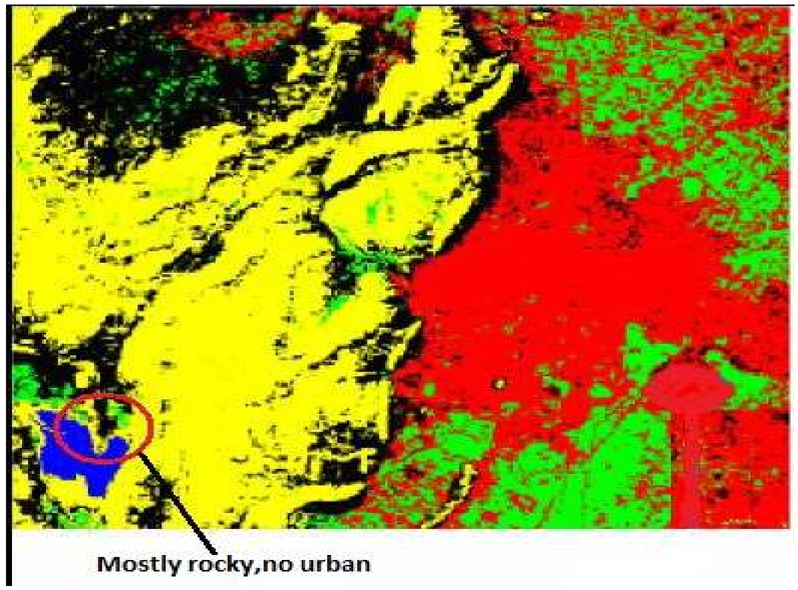


Figure 7.8(h) Classification result of Hybrid ACO/BBO/PSO

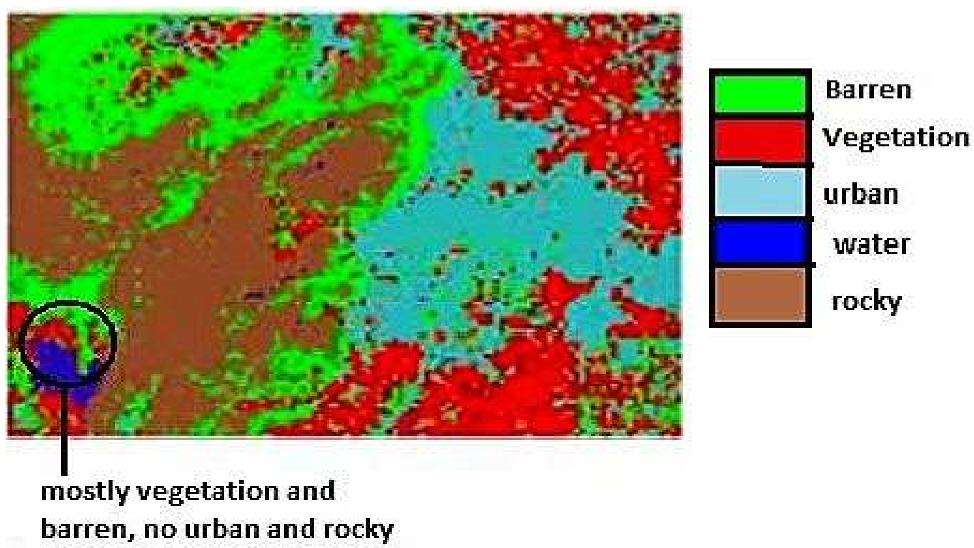


Fig 7.8 (i)Classification results of cAntMiner algorithm

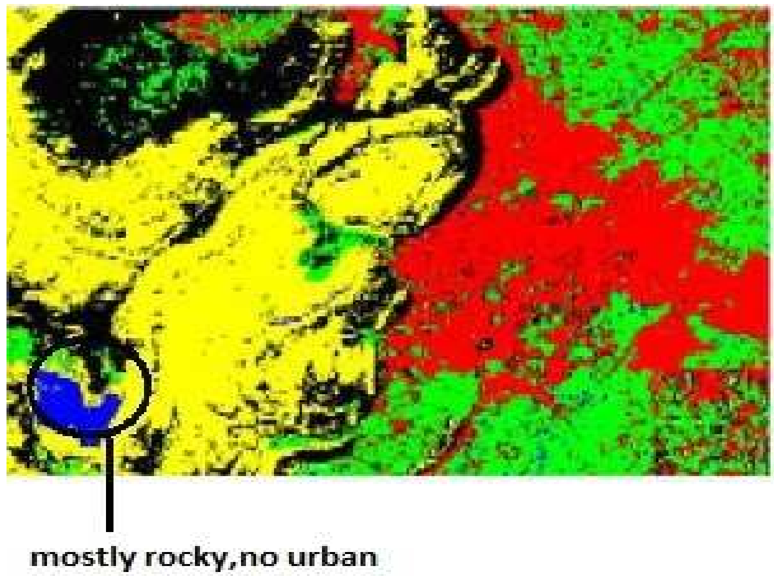


Fig 7.8(j) Classification results of Hybrid ACO/BBO

As we notice from above that both above techniques have high Kappa but yet they are not able to extract all the features from the heterogeneous regions.

Thus our ABC classifier has proved that it observed the heterogeneous regions of the image and even recognized the correct feature of that portion. Kappa coefficient is considered a universal way of evaluating the performance of any classification technique but from above results we infer that even though the overall accuracy of the classifier may be high but it does not mean that the heterogeneous regions in that image are well identified. It is because for kappa we only consider some random pixels from entire image and they may not fall under the heterogeneous regions. So to analyse the total efficiency of any technique we should also verify if the technique is able to extract all features from the conflict regions.

8. APPLICATION OF ABC ON HYPERSPECTRAL IMAGES

In this project we have worked on multispectral images, have used artificial bee colony for extracting land terrain features from remote sensing satellite image and have been successful in efficiently classifying the images. We can extend this work for hyperspectral images. The scope of ABC has not been explored much yet and we can utilize the bee optimization in reducing the dimension of hyperspectral images. In this section we discuss our survey work on various dimension reduction techniques and propose a theory of ABC for the same.

8.1 INTRODUCTION TO HYPERSPECTRAL IMAGES AND ITS DIMENTIONALITY PROBLEM

Hyperspectral images are spectrally over-determined. Their imagers measure the reflected radiance at a series of narrow contiguous wavelength bands. Such images contain enormous information to identify and distinguish between spectrally similar objects. The hyperspectral data can be represented as a cube with spatial information in X-Y plane and number of spectral band information along the Z axis (Fig.8.1). This image has potential of direct identification of surface materials and hence used in a wide variety of remote-sensing applications. Hyperspectral imagery possesses very rich spectral information and has the potential to identify subtle differences in land covers and thus finds huge application in remote sensing. The high dimensional images captured with the advance hyperspectral and ultraspectral airborne and landborne sensors provide huge spectral information for analysis. The applications range from satellite remote sensing and surveillance to medical imaging, geology and product inspection.

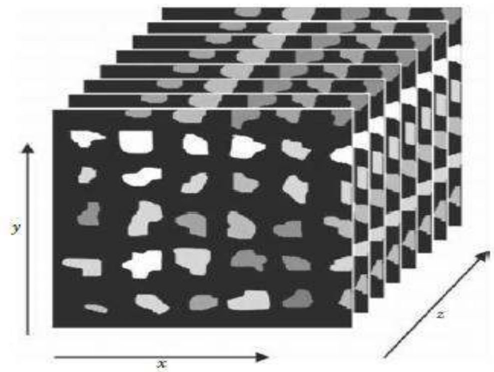


Fig. 8.1- A typical image cube generated by a hyperspectral imager, with two spatial dimensions x and y and one spectral dimension z [48]

But with vast amount of data comes difficulty of interpretation and analysis. The high dimension makes it difficult for transmission, storage and also for feature extraction and classification. It has been observed that there's a lot of redundancy and noise in the very close bands. Usually the information needed is contained within a range of wavelengths also relatively few bands can explain the vast majority of the information. This is where Dimension reduction comes into play as a very important pre-processing step for analysis of HSI. DR can be viewed as a transformation from high dimension into meaningful representation of reduced dimensionality.

Utilizing these powers of HSI is a challenge due to its high dimension which makes data interpretation, analysis and classification difficult. Hence reducing the dimension is the most critical pre-processing step for effective analysis of HSI. Due to the huge spectral information the standard reduction techniques cannot be applied. Many new techniques have been developed and few old ones are modified to optimally reduce the number of bands in image. There are the swarm Intelligence(SI) techniques using PSO, ACO and firefly, Genetic algorithm(GA) based, Simulated annealing(SA), the linear PCA, MNF, ICA and nonlinear like LLE, ISOMAPS, LE and wavelets.

8.2 THE DIMENSION REDUCTION TECHNIQUES

The Hyperspectral Dimension Reduction (HS-DR) significant techniques can be classified as below with two broad domains of artificial intelligence based and Transformation based. In AI the techniques under the AI domains are further classified as shown (Fig.8.2). The Transformation based reduces the dimension but also irreversibly transform the original dataset features. In each category the algorithm used is shown and later explained in the next section.

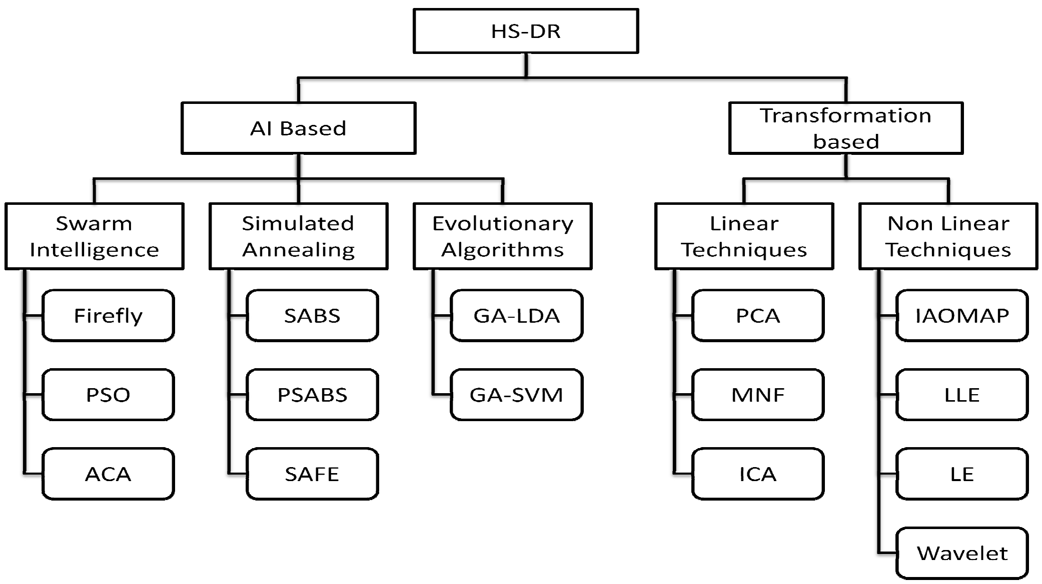


Fig. 8.2- Classification schema of HS-DR techniques

Artificial Intelligence Based Dimension Reduction

Swarm Intelligence Based

It aims to reduce the bands with least effecting the classification accuracy and loss of information. Meta heuristic optimization algorithms such as evolutionary algorithm and swarm Intelligence based methods have seen significant application for optimal feature selection.

Firefly Algorithm

Firefly algorithm (FA) [49] is inspired by biochemical and social aspects of real fireflies. Band selection can be implemented as an optimization procedure for searching an optimal subset of bands that better satisfy a desired measure [50]. F.Samadzadegan et.al, in their paper [51]] have evaluated the potential of FA and PSO as representatives of swarm intelligence-based methodologies in optimal band selection and dimensionality reduction of HSI. FA is used to select different subsets of bands from Hyperspectral data sets. The selected bands are evaluated using KNN classifier and the fittest of them used as optimum feature (band) for HSI classification.

Analysis— The experimental results [51] have shown that FA has very good classification accuracy. This technique still needs to be exploited and not much research is done yet.

Particle Swarm Optimization (PSO)

Particle Swarm optimization (PSO) is an meta-heuristic technique that was inspired by the bird flocking simulation by Kennedy and Eberhart [52] to guide particles to search for global optimum solution. PSO's simplicity, robustness and fast convergence are the reason for its increasing use as an efficient optimization technique. PSO visualizes the location of particles position in search space. For Hyperspectral images each spectral dimension corresponds to one band wavelength, the particle' location may be useful to identify useful correlation and characteristics between the bands. In [53] two swarms of particles are utilised for selecting the optimum bands, one continuous and other binary. He Yang and Qian Du in their paper [54], have proposed a PSO based DR approach to improve SVM based classification

Analysis— PSO can be applied even when the solution space is high dimensional and is proved to have excellent global searching capability. The particle swarms also possess the advantage of allowing the visualization of the selected features in contrast with their spectral locations, providing an appealing analysis tool for the field of remote sensing.

Ant colony Optimization (ACO)

ACO is a cooperative search and optimization technique that is motivated from the behaviour of real ant colonies. The ants rapidly establish the shortest route from their nest to the food source. Marco Dorigo inspired by this unique and intelligent behaviour proposed ACA[55]. Based on the hyperspectral image characteristics, such as spectral resolution, spatial correlation, spectrum correlation, data dimensionality, information capacity, the applications of ant colony algorithm to the processing are researched to find its suitability for hyperspectral image dimensionality reduction and classification. The process of ants searching optimal path is the process of forming the optimal band combination. The ant judges its path in terms of the correlation between the bands. High-dimensional hyperspectral data space is decomposed into several lower-dimensional data sub-space by ant optimal searching process. In the paper [56], the authors use ACO to decompose high dimensional data space into several low dimensional data subspace in terms of correlation between bands. Here every band of HSI was treated as ants with different features.

Analysis— This nature inspired technique is very simple to implement and also scalable for different applications. ACA has discrete and parallel characteristics which is beneficial for clustering of the data.

Simulated Annealing Based Dimension Reduction

Simulated Annealing (SA) is used for feature extraction of HSI as an optimization algorithm. SA is a probabilistic method proposed for finding the global minimum of a cost function that may possess several local minima. It works by emulating the physical process whereby a solid is slowly cooled so that when eventually its structure is 'frozen', this happens at a minimum energy configuration. Simulated Annealing [SAFE] method [57] provides an annealing schedule that starts at an effective high temperature and gradually decreases until it is slightly above zero. SABS [58] is designed to group the more correlated HSI bands into smaller subsets of modules regardless of original order in terms of wavelength. A parallel Simulated Annealing Based selection (PSABS) approach to band selection [59] has recently been proposed which is based on SA. The PSABS makes use of parallel computing technique and improves the performance.

Analysis— SAFE[57] is not trapped by the local minima at non zero temperature and provides quick distance measure of HS test samples as compared to conventional feature selection methods. It has lesser complexity. The PSABS overcomes the disadvantage of lengthy execution as in simple SAFE.

Evolutionary Algorithm (Genetic Algorithm)

GA has good ability to solve combinational problems and hence it is becoming an important tool for feature subset selection. This selection is useful when the analysis involves high dimensional feature space. Currently, there are two kinds of feature selection methods: filter methods and wrapper methods. The filter function behaves like a metric being optimised during GA search. The main advantage of filter method is the low computational cost and they require no feedback from classifiers and estimate classification performance using distance measures. The earlier techniques, forward selection and backward rejection evaluate the merits of each feature at a time but overlook the importance of relationship between features. GA takes into account this factor and is able to select the best features. The paper [60] has selected hyperspectral features by a GA-LDA approach where GA first recognises a smaller dimensional subset of features upon which Linear Discriminant Analysis(LDA) is applied for final dimensionality reduction. The wrapper methods are classifier dependent. The significance of the selected feature subsets are evaluated directly based on the classification accuracy. The paper [61] has discussed GA based wrapper feature selection using Support Vector Machine(SVM) for optimal classification.

Analysis— GA-LDA is seen as a good projection based dimensionally reduction technique [60]. The results showed that the GA-SVM method could significantly reduce the computation cost while improving the classification accuracy [61]. This method provides a robust intermediate step of pruning redundant and less useful features.

Transformation Based

Linear Methods

Principal Component Analysis (PCA)

PCA is the most widely used linear dimensionality reduction method based on the second order statistics. Its conceptual simplicity and existence of relatively efficient algorithms for computation are reasons for its popularity. PCA transforms the original features of a dataset with typically reduced number of uncorrelated ones termed principal components. Data is transformed in such a way that features with low variance are removed. This is done by finding the eigenvectors of the co-variance matrix of data points . Many modifications to original PCA algorithm have been suggested for optimal band selection and dimension reduction of hyperspectral images. One of the recent one uses hierarchical PCA [62] which reduces memory requirement for the heavy HSI by breaking the image into several parts and performing PCA on individual parts and then combining the results.

Analysis— PCA computes orthogonal projections that maximise the data variance and yields data set in a new uncorrelated coordinate system. Unfortunately information content in hyperspectral images does not always match such projections , PCA might fail to preserve all information needed for good classification. PCA transformations are crippled by its reliance on second order statistics. Though uncorrelated, the principal components can be highly statistically dependent. Under such conditions PCA fails to find the most compact description of the data. In that case PCA will require a larger dimensional representation than a non-linear technique [63]. This realization has prompted the development of non-linear alternatives of PCA and some hybrid approaches.

Minimum Noise Fraction (MNF)

The minimum (or maximum) noise fraction (MNF) is a second major algorithm belonging to the family of PCA techniques. It takes into account the sensor noise, unlike PCA which considers only the variances of each PC and assumes that noise is already isotropic. PCs do not always produce images of decreasing image quality, even though the variance is declining monotonically with PC number. It is not unusual to find, that in noisy images, local information can be represented in higher PC components. The MNF transform is effectively an algorithm consisting of two consecutive Principal Component's transformations. The first transformation, based on an estimated noise covariance matrix, results in transformed data for which the noise is uncorrelated with unit variance. The second transformation is a basic PCA on the noise-whitened data. Then the eigenvalues and images are analyzed. By discarding the components with small eigenvalues, the noise is separated from the data and the inherent dimensionality of the image is determined.

Analysis— Signal to noise ratio (SNR) is an effective measure of image quality and since MNF transformation orders the image in terms of this metric the ordering is based on image quality [64]. Therefore, it is invariant to scale changes, in any band, because it depends on the SNR instead of variance, like basic PCA.

Independent Component Analysis (ICA)

The DR problem in HSI is commonly dealt using PCA and MNF. The criteria both these techniques use are data variance and SNR which are designed to measure data second order statistics. Due to this some subtle, relatively small targets are not effectively uncovered. The ICA exceeds the second order statistics; it uses mutual information as a criterion to measure data statistical independency, therefore can capture information that cannot be retained by 2^{nd} order statistics. The key idea of the ICA assumes that data are linearly mixed by a set of separate independent sources and demix these signal sources according to their statistical independency measured by mutual information [65].

Analysis— ICA received little attention to perform DR in the past due to the fact that the ICA was not developed for this purpose. Wang and Chang [65] have implemented three algorithms of ICA and experimentally compared their results with PCA and MNF. ICA-based DR for dimension reduction is proposed particularly in applications such as endmember extraction, target detection and classification. The experimental results [65] demonstrate that the ICA algorithms generally outperformed second-order statistics-based transforms such as PCA, MNF to perform DR. However, when both second-order statistics and high-order statistics are required to be preserved during data compression, a mixed PCA/ICA algorithm was recently developed for this purpose in [66].

Non linear Methods

Isomaps

Isomap [67] is an extension of Multidimensional Scaling (MDS) in which embeddings are optimized to preserve geodesic distances between pairs of data points, estimated by calculating the shortest paths through large sub-lattices of data. The algorithm can discover nonlinear degrees of freedom as these geodesic distances represent the true low-dimensional geometry of the manifold. Constructing a neighbourhood graph G, weighted by shortest geodesic distances between all k -nearest neighbours. Isomap captures paths along a nonlinear manifold instead of the direct Euclidean distance. In the paper [68] manifold learning theory is used to solve the nonlinear dimensionality reduction problem of hyperspectral data.

Analysis— An important weakness of the Isomap algorithm is its topological instability. Isomap may construct erroneous connections in the neighbourhood graph G. Such short-circuiting can severely impair the performance of Isomap. A second weakness is that Isomap may suffer from ‘holes’ in the manifold. This problem can be dealt with by tearing manifolds with holes. A third weakness of Isomap is that it can fail if the manifold is nonconvex.

Locally Linear Embedding (LLE)

LLE [69] is a nonlinear dimensionality reduction method that maps high dimensional data into a low dimensional Euclidean space while preserving local topological structures. LLE achieves this by exploiting the local symmetries of linear reconstructions. It approximates images in high-dimensional space with small, flat patches and stitches them together in low-dimensional space to retain nonlinear structures. In LLE, the local properties of the data manifold are constructed by writing the high-dimensional data points as a linear combination of their nearest neighbours. In the low-dimensional representation of the data, LLE retains the reconstruction weights in the linear combinations.

Analysis— Unlike Isomap, LLE attempts to preserve solely local properties of the data. Hence, it is less sensitive to short-circuiting than Isomap, as only a small number of local properties are affected if short-circuiting occurs. It also tends to accumulate very sparse matrices whose structure can be exploited to save time and space. LLE is an unsupervised and non-iterative method, and it avoids the local minima problems faced by its competing methods.

Laplacian Eigenmaps (LE)

Belkin and Niyogi [70] developed the Laplacian Eigenmap for dimensionality reduction, and it preserves the relative distance between data points. LE is a geometrically motivated algorithm for constructing a representation for data sampled from a low dimensional manifold embedded in a higher dimensional space. Here the local properties are based upon the pairwise distances between the neighbours. In this first a weighted graph G is constructed in which every data point xi is connected to its k-nearest neighbours. Then, the distances between the low dimensional representations are minimized and nearby samples are highly weighted and thus brought closer together.

Analysis— LE suffers from many of the drawbacks as the LLE, such as the presence of a trivial solution that is prevented from being selected by a covariance constraint that can easily be cheated on. Despite this weakness it has been successfully been employed in many DR applications. The sampling technique of LE implicitly enforces natural clusters in the data.

Wavelets

PCA technique can be computationally expensive and is not able to remove the anomalies in an arbitrary band. Reduction of spectral data using wavelet decomposition is very helpful as it preserves the distinction between the spectral signatures of the high dimensional image. Automatic wavelet saves on computational complexity also filters the data anomalies. Its inherent property of preserving high and low frequency features thus preserving peaks and valleys in spectra is also an added advantage. The wavelet based reduction can be efficiently applied specifically for larger dimensions. The image is operated by firstly decomposing it using a filter (like Daubechies Wavelets).Inverse discrete wavelet transform (IDWT) is then used to obtain an approximation of each hyperspectral pixel. The number of decomposition levels is automatically computed as the lowest level after discarding outliers. Thus the reduced output data consists of now N/2l bands as that of N, which we originally had. In [71] dyadic discrete wavelet transform is proposed for DR and it shows higher accuracy for classification.

Analysis— Wavelet transform performs better for larger dimensions of data; this is due to its nature of compression where significant features of the image may be lost when the signal is undersampled. As the classifiers are pixel based they are well suited for wavelets which are also pixel based transformations. The lowpass and few highpass features which are lost in PCA are well retained by this method. For high accuracy of classification the number of spectral bands should be integral power of 2. The paper [72] has elaborated on the wavelet's higher efficiency in DR of HSI than PCA. The reasons described here are 1) the wavelet has a better time complexity; 2) the wavelet has better data locality characteristics, and 3) the wavelet suits better parallel processing. One issue with wavelets is that since it is being applied to each local pixel, might not be able to distinguish classes among neighbourhood pixel in the spatial domain.

8.3 OUR PROPOSAL: ARTIFICIAL BEE COLONY FOR DIMENSION REDUCTION

We go beyond the multispectral regime to the hyperspectral images and their dimensionality reduction problem. The scope of ABC has not been explored much yet but we can utilize the bee optimization in reducing the dimension of hyperspectral images.

Till now ABC has not been applied in any Hyperspectral imagery problem, be it classification or dimension reduction. After verifying its potential for multispectral images we have a vision that artificial bee colony has good prospects in dealing with the dimensionality problem of the high spectral images.

We have studied that in artificial bee colony algorithm the most promising food sources are selected by the onlooker bees and they follow the employed bee to that food source. If the quality and quantity of nectar in any source is not good then that source is rejected and new food source is searched in neighbourhood of old source.

For the dimensionality problem we can represent the food sources by the features (bands) of different dimensions. Every spectral band of hyperspectral imaging data which is in the high-dimension space may be looked as a feature. So the feature selection is band selection in the hyperspectral imaging. The selected subset will reduce the numbers of dimension of data and reserve the interesting information as much as possible. The process discards the redundancy features from a set of features in order to reduce the number of dimension

The profitability of any feature/band can be the quantity and quality of information content in that band. For hyperspectral image some contiguous bands do not contain unique information. The importance of any band is the information content. If band contains significant information of the image then it shall be selected by the onlookers. The bands with low content or overlapping information will be abandoned and thus removed from the solution set.

As the bees go on selecting the best food sources the bands get selected and some of the bands get abandoned. At the end the set of final food sources represent the final reduced dimension of the whole image.

ABC has better performance in terms of global optimization and hence it can be used to optimally select the bands from the hyperspectral images. In future ABC can be seen as a promising technique for reducing the dimension of hyperspectral images.

9. PUBLICATIONS FROM THESIS

Conference Name: International Conference on Computer Science & Engineering (ICCSE-2012)

Paper Title: "Remote Sensing Image Classification Using Artificial Bee Colony Algorithm"

Authors: Srideepa Banerjee, Akanksha Bharadwaj , Daya Gupta and V.K. Panchal

ISBN NUMBER: 978-93-81693-96-4

STATUS: Paper was presented and the work was highly appreciated, was awarded the 'Best Paper' and sent with good recommendation for review in the journal.

Location: Nainital, India

Conference date: 19th May 2012

Publisher/Proceedings: If the paper is accepted for the conference, it will be published in any one of the following journals, published by Interscience Open Access Journals.

International Journal of Computer Science and Informatics (IJCSI)

International Journal of Internet Computing (IJIC)

International Journal of Smart Sensors and Ad Hoc Networks (IJSSAN)

10. CONCLUSION AND FUTURE SCOPE

Artificial Bee Colony algorithm is being globally explored for various optimization problems. It has not yet been used for supervised classification of remote sensing image. We have proposed a framework of ABC in land cover feature extraction. It has proved itself a good classifier and even a refined one. Artificial Bee Colony is a recent approach which tries to model natural behaviour of real honey bees in food foraging. The proposed work called ABC for Remote Sensing Image Classification can be adapted according to the database of expert knowledge for a more focussed satellite image classification.

With the help of satellite image various features of our nature are identified. The general features on which scientists generally work are water, barren, urban and vegetation areas. Indeed other features like sand, snow, marshy etc can be recognized but the dataset we worked upon did not have such terrain. During feature extraction we attempt to identify all those features that land is consisting of.

The proposed work focuses on extracting the natural terrain features from satellite image using artificial bee colony.

ABC works on single unit of object that is; it processes the image pixel wise thus it focuses on heterogeneous issues more prominently. Our till date soft computing technique classified image with high accuracy but they have little focus on sparse regions. Classification accuracy can be highly increased if heterogeneous regions are correctly mapped. Images generally has sparse region where lots of features can be identified. Most optimization techniques consider a global model and overlook these conflicting regions which get wrongly classified. This heterogeneous region has been easily captured by ABC. Here lies a significant difference than other computational intelligence techniques. In section 7.3 we have presented a comparison of how different soft computing techniques like MDC, MLC, Fuzzy classifier, BBO, MC, Hybrid ACO/BBO, CAntMiner and some more to show how they deal with the feature extraction from heterogeneous regions and proved that using ABC as the classification we are able to distinctly separate the different features from these regions.

We have shown that the classification efficiency of heterogeneous regions is independent of the Kappa coefficient calculated for the overall image. This is because for calculating Kappa we consider some random pixels from whole image which have very less probability of falling in these small regions.

Our proposed classifier is flexible enough to classify any land cover very efficiently as none of the parameters are dependent on feature type and hence can be adapted according to the application.

The proposed algorithm is applied to two different images and in both of them it has shown efficient and good results. It was first applied on the 7 band multispectral satellite image of Alwar region in Rajasthan (Fig 6.1) and then to the 6 band multispectral image of Saharanpur region in Uttar-Pradesh (Fig 6.2). Both these regions have a good number of terrain features to be identified. We obtained good classification results for both datasets as shown in figure 6.6 for Alwar output and figure 6.8 for Saharanpur output. The validation of our classification accuracy also turned out good with high Kappa coefficient for Alwar as 0.9174 and for Saharanpur 0.9586. The Error matrix for each has been shown in table 7.1 and 7.2. We also present a comparative study of the results of applying our proposed classifier and the other recent soft computing classifiers such as BBO, Fuzzy sets, membrane computing and FPAB/BFO when compared with the traditional probabilistic classifiers such as Minimum Distance to Mean Classifier and the Maximum Likelihood Classifier on the 7-band Alwar Image based on kappa coefficient as the parameter. This comparison is represented in fig 7.7.

The results presented are preliminary and there is scope for improvement to develop this algorithm as efficient classifier. To prove our classifier as a universal we need to test it on some more datasets with varied terrain features. The ABC implemented in our work is in the beginning stage where we can modify the algorithm to improve the performance. We can refine the neighbourhood concept and introduce some randomization. We also recommend for certain modification to the algorithm so that the Kappa coefficient can be improved further.

In our work we have worked on multispectral images, have used artificial bee colony and have been successful in extracting all features from them. The scope of ABC has not been explored much yet but we can utilize the bee optimization in reducing the dimension of hyperspectral images. Dimension reduction of Hyperspectral images is very critical for extracting information from these images. There are various techniques for this process, we have analysed them strategically and presented a theory which uses artificial bee colony for optimal dimension reduction.

ABC has better performance in terms of global optimization and hence it can be used to optimally select the bands from the hyperspectral images. In future ABC can be seen as a promising technique for reducing the dimension of hyperspectral images.

The problem of mixed pixels (like boundary pixels) can also be solved by the classification algorithm proposed in this work.

REFERENCES

[1] Omkar S.N., Kumar Manoj, Muley Dipti, Urban Satellite Image Classification using Biologically Inspired Techniques in proceedings of IEEE, 2007.

[2] Karaboga D, “ An idea based on honey bee swarm for numerical optimization” TR-06, Erciyes University, En-ginering Faculty, Computer Engineering Department, 2005

[3] Dorigo M, Gambardella L M. “Ant colony system: a cooperative learning approach to the travelling salesman problem”. IEEE Transactions on Evolutionary Computation, 1997.1(1).pp. 53-66.

[4] Ümit Haluk Atasever, Coşkun Özkan, Filiz Sunar,‘The Use of Artificial Intelligence Optimization Algorithms in Unsupervised Classification’, EARSeL, 2011

[5] T.M. Lillesand, R.W. Kiefer & J.W. Chipman, ―Remote Sensing & Image Interpretations‖ 5th edition, Wiley & Sons Pvt. Ltd., pp 586-592, 2005, England.

[6] L.A. Zadeh, ―Fuzzy sets‖, Information & Control, vol.8, no.3, pp 338-353, June,1965.

[7] Z. Pawlak, ―Rough Set Theory & its Application to Data Analysis‖, Cybernetics & Systems, vol. 29, no. 7, pp 661-688, 1998

[8] E. Bonabeau, M. Dorigo, G. Theraulaz, ―Swarm Intelligence from Nature to Artificial System‖, First Edition, Oxford University Press, pp 1-24, USA, 1999.

[9] L.A. Zadeh, ―Fuzzy sets‖, Information & Control, vol.8, no.3, pp 338-353, June,1965.

[10]S.N.Omkar, K.M. Manoj, D. Mudigree & D. Muley, ―Urban Satellite Image Classification using Biologically Inspired techniques", In Proceedings of IEEE International Symposium on Industrial Electronics, Vigo, Spain, pp. 1767-1772, 2007.

[11]Daya Gupta, Lavika Goel, V.K.Panchal, ―Embedding Expert Knowledge to Hybrid Bio-Inspired techniques- An Adaptive Strategy towards Focussed Land Cover feature Extraction‖, International Journal of Computer Science & Information Security, (IJCSI), vol.8, no.2, 2010.

[12] D. Chandrakala1 and S. Sumathi2, ‘Application of Artificial Bee Colony Optimization Algorithm for Image Classification Using Color and Texture Feature Similarity Fusion’, ISRN Artificial Intelligence Volume 2012, Article ID 426957, 10 pages doi:10.5402/2012/426957, 2012

[13] Shefali Aggarwal, ‘Principles Of Remote Sensing’, *Photogrammetry and Remote Sensing Division,IIRS, Satellite Remote Sensing and GIS Applications in Agricultural Meteorology.*

[14] Introduction to Remote Sensing and Image Processing, IDRISI Guide to GIS and Image Processing Volume 1.

[15] B.Bhatta, ‘ Remote sensing and GIS’, Oxford University Press. 2008.

[16]Url:http://www.nrcan.gc.ca/earth-sciences/geography-boundary/remote sensing/fundamentals/1920

[17] A J2SE Application Enhances and Classifies Map Data‖, http://software.intel.com/en-us/articles/

[18] Tong Hau Lee, Mohammad Faizal, Ahmad Fauzi and Ryoichi Komiya, " Segmentation of CT Brain Images Using K-means and EM Clustering", 5th IEEE International Conference on Computer Graphics, Imaging and Visualization, 2008.

[19]Tatsuya Yamazaki, "Introduction of EM Algorithm into Color Image Segmentation", ATR Adaptive Communications Research Laboratories, 1998.

[20] Thales Sehn Korting and Luciano Vieira Dutra, " Improvements to Expectation-Maximization approach for unsupervised classification of remote sensing data", National Institute for Space Research, 2007.

[21] Adil Baykasolu, Lale Özbakır and Pınar Tapkan "Artificial Bee Colony Algorithm and Its Application to Generalized Assignment Problem", Swarm Intelligence: Focus on Ant and Particle Swarm Optimization, tech Education and Publishing, December 2007.

[22] J. Kennedy and R. C. Eberhart. "Swarm Intelligence". Morgan Kaufmann. 2001

[23] K. M. Passino, "Biomimicry of bacterial foraging for distributed optimization and control," *IEEE Control Systems Magazine*, vol. 22, no. 3, pp. 52–67, 2002.

[24] Simon, D. (2008), "Biogeography-based Optimization", in IEEE Transactions on Evolutionary Computation, vol. 12, No.6, IEEE Computer Society Press. pp. 702-713.

[25] V.K.Panchal Parminder Singh Navdeep Kaur Harish Kundra, "Biogeography based Satellite Image Classification", *(IJCSIS) International Journal of Computer Science and Information Security, Vol. 6 , No.2, 2009*

[26] Gheorghe Păun, Grzegorg Rozenberg, ―An Introduction to & an Overview of Membrane Computing

[27] Păun, Gheorghe. "Computing with membranes". TUCS Report 208, Turku Center for Computer Science (1998).

[28] Daya Gupta, Bidisha Das, V. K. Panchal, 'A methodical study for the extraction of landscape traits using membrane computing technique', The 2011 World Congress in Computer Science, Computer Engineering, and Applied Computing, 2011.

[29] Daya Gupta, Lavika Goel, V.K.Panchal, "Embedding Expert Knowledge to Hybrid Bio-Inspired techniques- An Adaptive Strategy towards Focussed Land Cover feature Extraction", International Journal of Computer Science & Information Security, (IJCSI), vol.8, no.2, 2010.

[30] Dervis Karaboga , Bahriye Akay, "A comparative study of Artificial Bee Colony algorithm", Elsevier, 2009

[31] Wenping Zou, Yunlong Zhu,Hanning Chen, and Xin Sui, 'A Clustering Approach Using Cooperative Artificial Bee Colony Algorithm', Discrete Dynamics in Nature and Society Volume 2010, Article ID 459796, 2010.

[32]D. Karaboga and B. Basturk, "Artificial bee colony (ABC) optimization algorithm for solving constrained optimization problems," in Proceedings of the 12th International Fuzzy Systems Association World Congress on Foundations of Fuzzy Logic and Soft Computing (IFSA '07), vol. 4529 of Lecture Notes in Artificial Intelligence, pp. 789–798, Springer, 2007.

[33] D. Karaboga and B. Akay, "Artificial bee colony (ABC) Algorithm on training artificial neural networks," in Proceedings of the 15th IEEE Signal Processing and Communications Applications (SIU '07), pp. 1–4, June 2007.

[34] C. Ozturk and D. Karaboga, “Classification by neural networks and clustering with artificial bee colony (ABC) algorithm,” in Proceedings of the 6th International Symposium on Intelligent and Manufacturing Systems, Features, Strategies and Innovation, Sakarya, Turkey, October 2008.

[35] Camazine, S. and Sneyd, J, ‘A model of collective nectar source by honey bees: selforganization through simple rules. Journal of Theoretical Biology 149, 547–571,1991.

[36] D. Teodorovi´c, T. davidovi´c, M.Selmi´c, ‘Bee Colony Optimization: The Applications Survey’, ACM Transactions on Computational Logic, 2011.

[37] S. L. Sabat, *et al.*, "Artificial bee colony algorithm for small signal model parameter extraction of MESFET," Engineering Applications of Artificial Intelligence, vol. 23, pp. 689-694, 2010.

[38] Mohd Afizi Mohd Shukran, Yuk Ying Chung, Wei-Chang Yeh, Noorhaniza Wahid, Ahmad Mujahid Ahmad Zaidi, ‘Artificial Bee Colony based Data Mining Algorithms for Classification Tasks’, Modern Applied Science, Vol. 5, No. 4; August 2011.

[39] Karaboga D & B Basturk, “ On the performance of artificial bee colony (ABC) algorithm”, Applied Soft Computing, 687-697, 2008.

[40] D. LU and Q. WENG, “A survey of image classification methods and techniques for improving classification performance”, International Journal of Remote Sensing Vol. 28, No. 5, , 823–870, 10 March 2007

[41] Parminder Singh, Navdeep Kaur, Loveleen Kaur, “Satellite Image Classification by Hybridization of FPAB Algorithm and Bacterial Chemotaxis”, International Journal of Computer Technology and Electronics Engineering (IJCTEE) Volume 1, Issue 3, 2011.

[42] Fernando E.B. Otero, Alex A. Freitas, and Colin G. Johnson, *c*Ant-Miner, “ An Ant Colony Classification Algorithm to Cope with Continuous Attributes”, Springer-Verlag Berlin, Heidelberg 2008

[43] D. Bratton and J. Kennedy, “ Defining a Standard for Particle Swarm Optimization”, Proceedings of the 2007 IEEE Swarm Intelligence Symposium, Honolulu, Hawaii, USA, April 2007

[44] N. Holden and A.A. Freitas, “ A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data”, In: Proc. 2005 IEEE Swarm Intelligence Symposium (SIS-05), pp. 100-107, IEEE, 2005.

[45] N. Holden and A.A. Freitas, “ A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data. In: Proc. 2005 IEEE Swarm Intelligence Symposium (SIS-05), pp. 100-107, IEEE, 2005.

[46] Lavika Goel, V.K Panchal, Daya Gupta, “Embedding Expert Knowledge to Hybrid Bio-Inspired Techniques- An Adaptive Strategy towards Focussed Land Cover Feature Extraction”, International Journal of computer Science and Information Security (IJCSIS), vol. 8, no. 2, 2010.

[47] Akanksha Bharadwaj, Srideepa Banerjee, Daya Gupta, V.K. Panchal, “Comparative Analysis of Nature Inspired Techniques for Heterogeneous region in Remote Sensing”, International Conference on Computer Science & Engineering, 2012,978-93-81693-96-4

[48] Marcus Borengasser, William S. Hungate, Russell Watkins, ―Hyperspectral Remote Sensing Principles and Applications‖, Taylor & Francis Series in Remote Sensing Applications, Indiana State University.

[49] Yang, X.,(2009), ―Firefly Algorithm for Multimodal Optimization‖, SAGA 2009, LNCS 5792,pp.169-178,2009.

[50] S. T. Monteiro, K. Uto, Y. Kosugi, N. Kobayashi, E. Watanabe, K. Kameyama, ―Feature Extraction of Hyperspectral Data for under Spilled Blood Visualization Using Particle Swarm Optimization,‖ IJBEM, vol. 7, pp. 232-235,2005.

[51] F.Samadzadegan, F.Tabib Mahmoudi, ―Optimum Band Selection in Hyperspectral Imagery Using Swarm Intelligence Optimization Algorithms‖, ICIIP ,2011

[52] J. Kennedy and R. C. Eberhart, Swarm Intelligence. San Francisco: Morgan Kaufmann Publishers, 2001.

[53] Sildomar Takahashi Monteiro, Yukio Kosugi, ―A Particle Swarm Optimization-based Approach for Hyperspectral Band Selection‖, *CEC 2007*

[54] He Yang, Qian Du, ―Particle Swarm Optimization-based dimensionality reduction for hyperspectral image classification‖, IGARSS 2011

[55] Dorigo M, Gambardella L M. ―Ant colony system: a cooperative learning approach to the traveling salesman problem‖. IEEE Transactions on Evolutionary Computation, Vol.1, No.1, pp. 53-66, 1997.

[56] Shuang Zhou,Jjunping Zhang, Baoku Su*,"* Dimensionality reduction and classification based on ant colony algorithm for Hyperspectral remote sensing image*", IGARSS 2008.*

[57] Yang-Lang Chang, Jyh-Perng Fang and Jin-Nan Liu, Hsuan Ren, Wen-Yew Liang, ―A Simulated Annealing Feature Extraction Approach for Hyperspectral Images‖, IEEE,2007.

[58] J. P. Fang, Y. L. Chang, H. Ren, C. C. Lin, W. Y. Liang, and J. F. Fang, ―*A simulated annealing band selection approach for hyperspectral imagery"*, Proc. SPIE 6378, pp. 63781G–1–10, 2006. doi:10.1117/12.685683

[59] Yang-Lang Chang*,* Kun-Shan Chen, Bormin Huang, Wen-Yen Chang, Jon Atli Benediktsson, and Lena Chang,― A Parallel Simulated Annealing Approach to Band Selection for High-Dimensional Remote Sensing Images‖, ieee journal of selected topics in applied earth observations and remote sensing,Vol. 4, no. 3, september 2011.

[60] Minshan Cui, Saurabh Prasad, Majid Mahrooghy, Lori M. Bruce, James Aanstoos, " Genetic algorithms and linear discriminant analysis based dimensionality reduction for remotely sensed image analysis"‖, IGARSS 2011

[61] Li Zhuo , Jing Zhengb, Fang Wang, Xia Li , Bin Ai , Junping Qian, ―A Genetic algorithm based wrapper feature selection method for classification of hyperspectral images using support vector machine‖, *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences. Vol. XXXVII. Part B7. Beijing 2008.*

[62] Abhishek Agarwal, Tarek El-Ghazawi, Hesham El-Askary, and Jacquline Le Moigne, "Efficient Hierarchical-PCA Dimension Reduction for Hyperspectral Imagery"‖, International Symposium on Signal Processing and Information Technolog*,* IEEE,2007

[63] Nandakishore Kambhatla Todd K. Leen, ―Dimension Reduction by Local Principal Component Analysis‖, *Neural Computation* 9, 1493 1516 ,MIT,1997

[64] Leonidas Fountanas Thesis on ―Principal Components based techniques for Hyperspectral Image data‖, Naval Postgraduate School Monterey, California, December 2004.

[65] Jing Wang, Chein-I Chang, ―Independent Component Analysis-Based Dimensionality Reduction With applications in Hyperspectral image analysis‖, IEEE Transactions on Geoscience and remote sensing, vol. 44, no. 6, june 2006

[66] J. Wang and C.-I Chang, ―Mixed PCA/ICA spectral/spatial compression for hyperspectral imagery,‖ presented at the *OpticsEast, Chem. Biol. Standoff Detection III (SA103)*, Boston, MA, Oct. 2005.

[67] J.B. Tenenbaum, V. de Silva, J. C. Langford, ―A global geometric framework for nonlinear dimensionality reduction. Science, Vol. 290, No. 5500, pp. 2319–2323. 2000.

[68] Dong Guangjun Zhang Yongsheng Ji Song, ―Dimensionality Reduction of Hyperspectral

[69] S.T. Roweis and L.K. Saul. ―Nonlinear dimensionality reduction by Locally Linear Embedding. Science‖, 290(5500):2323–2326, 2000.

[70] M. Belkin and P. Niyogi. ―Laplacian Eigenmaps and spectral techniques for embedding and clustering,‖ In Advances in Neural Information Processing Systems, volume 14, pages 585–591, Cambridge, MA, USA, 2002. The MIT Press.

[71] Lori Mann Bruce, Cliff H. Koger, and Jiang Li,―Dimensionality Reduction of Hyperspectral Data Using Discrete Wavelet Transform Feature Extraction‖, IEEE Transactions on Geoscience and remote sensing, vol. 40, no. 10, october 2002.

[72] Sinthop Kaewpijit, Jacqueline Le Moigne, Tarek El-Ghazawi, ―*Automatic Reduction of Hyperspectral Imagery Using Wavelet Spectral Analysis", IEEE Transactions on Geoscience and remote sensing, vol. 41, no. 4, APRIL 2003.*

APPENDIX A

ABBREVIATIONS

EM – Electromagnetic

ABC – Artificial Bee Colony

MC – Membrane Computing

DN – Digital Number

ACO – Ant Colony Optimization

BBO–Biogeography Based Optimization

PSO – Particle Swarm Optimization

BFO – Bacterial Foraging optimization

MDC – Minimum Distance Classifier

MLC – Maximum Likelihood Classifier

SI – Swarm Intelligence

SC – Soft Computing

FL –Fuzzy Logic

RS – Rough set theory

NIR – Near Infra-Red

SVM – Support Vector Machine

CBR – Case Based Reasoning

MIR – Middle Infra-Red

RS1 – Radarsat-1

DEM – Digital Elevation Model

RS2 – Radarsat2

ANN – Artificial Neural Network

AIS – Artificial Immune System

APPENDIX B

An Introduction to Image Processing in MATLAB

MATLAB is a high performance language for technical computing. It integrates computation, visualization and programming in an easy-to-use environment where problems and solutions are expresses in familiar mathematical notation. It is an interactive system whose basic data element is an array that does not require dimensioning. This allows you to solve many technical computing problems, especially those with matrix and vector formulations in a fraction of the time it would take to write a program in a scalar non interactive language such as C , C++ or JAVA.

Image formats supported by Matlab

The following image formats are supported by Matlab:

- BMP
- HDF
- JPEG
- PCX
- TIFF
- XWB

Working formats in Matlab

If an image is stored as a JPEG-image on your disc we first read it into Matlab. However, in order to start working with an image, for example perform a wavelet transform on the image, we must convert it into a different format.

Intensity image (gray scale image)

It represents an image as a matrix where every element has a value corresponding to how

bright/dark the pixel at the corresponding position should be colored. There are two ways to represent the number that represents the brightness of the pixel: The double class (or data type). This assigns a floating number ("a number with decimals") between 0 and 1 to each pixel. The value 0 corresponds to black and the value 1 corresponds to white. The other class is called uint8 which assigns an integer between 0 and 255 to represent the brightness of a pixel.

Binary image

This image format also stores an image as a matrix but can only color a pixel black or white (and nothing in between). It assigns a 0 for black and a 1 for white.

Indexed image

This is a practical way of representing color images. An indexed image stores an image as two matrices. The first matrix has the same size as the image and one number for each pixel. The second matrix is called the *color map* and its size may be different from the image. The numbers in the first matrix is an instruction of what number to use in the color map matrix.

RGB image

This is another format for color images. It represents an image with three matrices of sizes matching the image format. Each matrix corresponds to one of the colors red, green or blue and gives an instruction of how much of each of these colors a certain pixel should use.

Reading Image Files

The command to read an image from file *filename* and store it in matrix variable p is:

p = imread('*filename*');

The filename may be an absolute pathname or a relative pathname from the current working directory. Omitting the ';' at the end of the command causes the value to be printed to the command window.

Writing Image Files

The command to write an image from variable p and store it in file *filename* is:

imwrite (p, '*filename*');

The filename may be an absolute pathname or a relative pathname from the current working directory.

Displaying Image Files

The command to display the image from variable p to a figure window is:

imshow(p);

An additional parameter may be used to set the number of display levels or set the range of display levels. Various controls, such as dynamic display of index and value for the cursor position, are available in the image display tool.

The figure command can be used to create a new current figure for the display:

figure, imshow (p);

Standard Arrays

MATLAB has standard arrays for creating arrays of a defined size with zeros, ones, true, false, or random values. For example: p = zeros(M,N);

Built-in Functions (used in this work)

MATLAB has many useful builtin functions. For example: max(p(:))

Gives the largest value in the matrix.

Note, given matrix p, max (p) treats the matrix as an array of column vectors and returns a vector of the largest value in each column.

Standard deviation – std : If X is a matrix, std(X) returns a row vector containing the standard deviation of the elements of each column of X. If X is a multidimensional array, std (X) is the standard deviation of the elements along the first non-singleton dimension of X.

APPENDIX C

An Introduction to ERDAS

URL: http://www.ERDAS.com/

ERDAS is pleased to provide ERDAS IMAGINE® version 8.4. Many private and commercial users who need to extract and interpret information from imagery recognize ERDAS IMAGINE as a must have. With ERDAS IMAGINE 8.4, ERDAS' latest, most advanced release of ERDAS IMAGINE, production workflows are enhanced and simplified like never before. As an example, the Batch Wizard streamlines repetitive procedures such as importing; reprojecting, and exporting large numbers of files at once, using a wizard approach to record and —re-play‖ commonly used procedures. Also featured is the IMAGINE Expert Classifier™ – a tool for graphically building and executing geographically aware, rules-based expert systems. This tool can be used to build decision support systems, classifiers for high-resolution imagery, GIS analysis techniques, etc. These can then be distributed to other users for use with their own data.

Key Features Summary

- IMAGINE Expert Classifier
- Direct read and edit of ESRI's Shapefiles
- Direct read of ESRI's SDE data
- Enhanced and expanded native raster file handling
- Re-projection of raster data on-the-fly
- Batch processing wizard
- Enhanced Viewer functionality
- Improved print versatility on Windows NT
- International 2-byte font support in Annotation layers
- Support for ERDAS IMAGINE .img files larger than 2 GB