Hyperspectral Image Classification Using SVM and Neural Network

A Dissertation Submitted in Partial Fulfillment of the Requirement

for the Award of the Degree of

MASTER OF ENGINEERING

in

Wireless Communication Submitted By

ASEEM MAHAJAN

Roll No: 801563002

Under Supervision of

Dr. Amit Mishra

Assistant Professor, ECED

ELECTRONICS AND COMMUNICATION ENGINEERING DEPARTMENT

THAPAR UNIVERSITY

PATIALA-147004 (PUNJAB)

JUNE, 2017

DECLARATION

I, Aseem Mahajan hereby declare that the work presented in this dissertation, entitled " **Hyperspectral Image Classification Using SVM and Neural Network,** " in partial fulfillment off the requirement for the award of degree of Master of Engineering in Wireless Communication Engineering, submitted at Electronics and Communication Department (ECED), Thapar University, Patiala, is an authentic record of work carried out under the supervision of **Dr. Amit Mishra** (Assistant Professor, ECED, Thapar University, Patiala) from July 2015 to July 2017. The matter presented in this has not been submitted either in part or full to any other university or institute for the award of any other degree.

Date: 17.07.17

Aseem Mahajan

801563002

It is certified that the above statement made by the student is correct to the best of my knowledge and belief.

Date: 17.07.17

Dr. Amit Mishra

Assistant Professor
ECED Thapar University,
Patiala

ACKNOWLEDGEMENT

We cannot achieve anything in the field of technical education until or unless the theoretical education acquired in the classroom is effectively joined to its practical approach that is taking place in the research. Although an engineer can only be successful through hard work, but the contribution of his teachers and all those who have been helpful cannot get unnoticed.

I wish to express my deep gratitude and sincere thanks to my supervisor, **Dr. Amit Mishra**, Assistant Professor, Electronics and Communication Department (ECED), Thapar University, Patiala, for his invaluable guidance, constant encouragement, and immense motivation, throughout the work. His valuable advice and suggestions for the corrections and modifications did enhance my work.

I also want to express my gratitude to **Dr. Alpana Aggarwal**, Head (ECED), **Dr. Hemdutt Joshi**, P.G Coordinator and **Dr. Ashutosh Kumar Singh** Program Coordinator, M.E. (WC) for providing adequate environment in carrying out the work.

Finally, I would like to thank all those persons who directly or indirectly helped me in carrying out this work in right direction.

Aseem Mahajan

801563002

ABSTRACT

A Satellite image classification is a significant method used in remote sensing for the automated analysis and pattern recognition of satellite data, which facilitate the automated understanding of a large amount of information. These days, there exist many types of classification algorithms, such as parallelepiped and minimum distance classifiers, but it is still essential to get better performance in terms of correctness rate. Alternatively, over the last few years, cellular automata have been utilized in remote sensing to implement procedure related to simulation. While there is little preceding research of cellular automata related to satellite image classification, they offer much reward that can improve the results of classical categorization algorithms.

In this research work, we firstly segment the pixels presented in the satellite image by using K - means to divide the image into clusters. Features are extracted with the help of PCA technique. After extracting the features, optimization is done by using PSO technique. Classification will be done by using Support vector machine (SVM) and Neural network (NN). The performance parameters like PSNR, MSE, and index error will be measured after simulating the proposed work.

Keywords: Image classification, Multispectral image classification, remote sensing, Swarm intelligence, Principle component analysis, Support vector machine, neural network and MATLAB.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

LIST OF ABBREVIATION

ABC	Artificial Bee Colony
ACO	Ant Colony Optimization
ANN	Artificial Neural Network
CBS	Constrained Band Selection
DBFE	Decision Based Feature Extraction
GA	Genetic Algorithm
ICA	Independent Component Analysis
LSMA	Linear Spectral Mixture Analyser
MNF	Minimum Noise Fraction
MSE	Mean Square Error
PCA	Principal Component Analysis
PSNR	Peak Signal to Noise Ratio
PSO	Particle Swarm Optimization
SPCT	Segmented Principal Component Analysis
SVM	Support Vector Machine

CHAPTER 1: INTRODUCTION

1.1 PREAMBLE

Image Classification is one of the imperative methodologies for perceiving distinctive landscape highlights. The investigators figure out which grouping technique meets his particular assignment. At present, various methods like Evolutionary calculations, ANN, ACO, PSO and Biogeography Based Optimization are being connected to picture classification. These sort of common registering method could even be named as nature roused system. Therefore each and every investigators picks which specific classifier is proper for the undertaking close by.

Remote detecting alludes to the innovation of securing data about the world's surface components (land and water) and climate utilizing space-borne stages. Satellite remote detecting has been perceived as an important device for deciding, watching, portraying, and examining, about our condition. Multi-spectral pictures catch diverse territory geologies like infertile land, water, vegetation, rough, and urban, that are particularly being arranged for promote necessity for picture examination. In this manner picture characterization and remote detecting are between twined to each other.

A Remote recognizing has been used as a piece of different organic applications with the course of action of comprehend and hinting at change an extensive variety of issues: soil quality examinations, water resource investigate, meteorology propagations, and characteristic protection, among others. To decide these issues, one must join together and prepare monstrous measures of satellite data, which make a standout amongst the most troublesome issues going up against remote recognizing [1]. It is the best way to deal with describe the things on the earth without compasses to it with the help of exceptionally actuated sensors and work on the multiplied signs which gets reflected by the ground objects like water, shake, vegetation's soil and various others in the shape go that are later on changed over in Digital Number.

Among the majority of the strategies utilized as a part of remote detecting to enable forecaster master to comprehend the information assembled, grouping calculations are the most helpful and capable. These arrangement calculations for satellite pictures group together picture pixels into a predetermined number of classes, which helps in decipher a lot of information controlled in the ghastly groups [2]. While applying a grouping calculation to a reliance picture, the data accomplish through the satellite telecom sensors as computerized levels are adjusted into an

all-out scale that is effortlessly decipher by investigator specialists. The resultant arranged picture is a topical guide of the one of a kind satellite picture, and pixels have a place with a similar class split alike ghastly character.

With the assistance of satellite picture different components of our tendency can be recognized. The general components on which researchers by and large work are water, infertile, urban and vegetation zones. We have taken a shot at Alwar area of Rajasthan as appeared in figure 1.1. Surely different components are being perceived yet these are ordinarily accessible in any sort of land regions. In this manner arrive covers all these normal classification of geologies. All through element extraction we endeavor toward recognizing every one of those elements that land is comprising of. As our dataset is Alwar locale of Rajasthan accordingly we have portrayed its picture to comprehend arrive cover include.

1.2 IMAGE CLASSIFICATION

Image Classification is a specific instance of Pattern Recognition. The general target of the characterization procedure is to consequently arrange all pixels in a picture into arrive cover classes in light of the predefined grouping model. The yield is a topical picture which relates with a constrained amount of highlight classes rather than a ceaseless picture with varying shades of dim or contrasting hues speaking to a persistent scope of ghastly reflectance's [3]. Picture arrangement is a noteworthy piece of the remote detecting, picture examination and example acknowledgment. In some occasion, the arrangement itself might be the question of the investigation. For example, association of land use from remotely detected information deliver a guide like picture as the last result of the examination. The picture order so shapes a basic instrument for examination of the advanced pictures. The term classifier alludes sadly to a PC program that gear a particular procedure for picture characterization.

Image Classification is basically used to extract the information class from the multispectral data. It make use of the reflectance statistics for individual pixels. These reflectance statistics are also known as the spectral signatures which is a unique quantity, thus comparing the spectral signature obtained and that with the known data we just able to classify the particular object or to which class the classified objects belong. There are many different approaches to perform the image classification and some of the important ones has been explained further in the upcoming pages

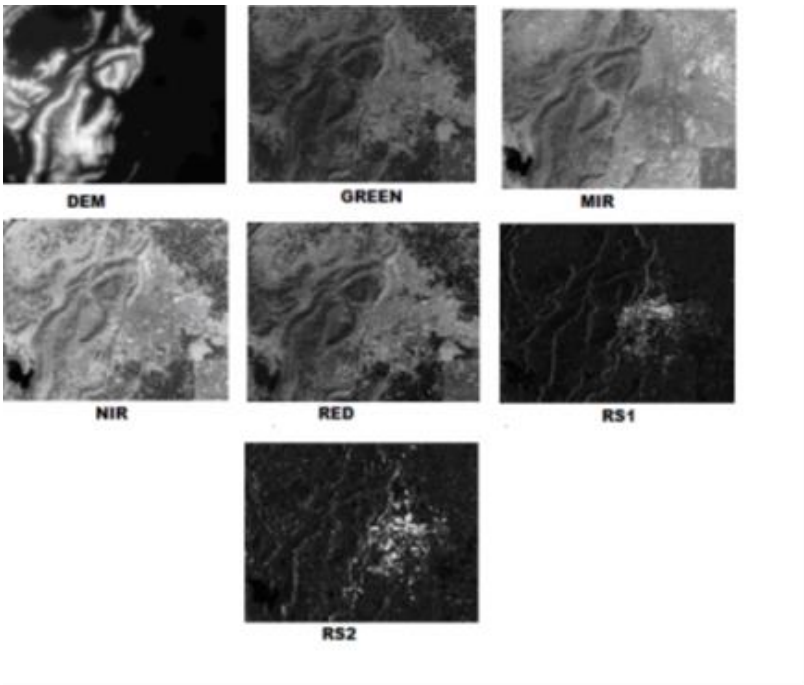

Figure 1.1: Image classification of Alwar taken from satellite

Classification process consists of following four main steps:

A. **Pre-processing:** It includes enhancement of image suffering from noise, enhanced the features required for further processing but the nature of priory information should be known to us.

B. **Detection and feature extraction of an object:** Detection includes detection of position and other characteristics of moving object image obtained from camera [4]. Extraction of the object includes estimating the trajectory of the object in the image plane.

C. **Training:** Selection of the particular attribute which best describes the pattern. Training represents the learning relation between data and the attributes from the fraction of input training data available.

D. **Classification of the object**: Object classification step categorizes the detected objects into predefined classes by using suitable method that compares the image patterns with the target patterns.

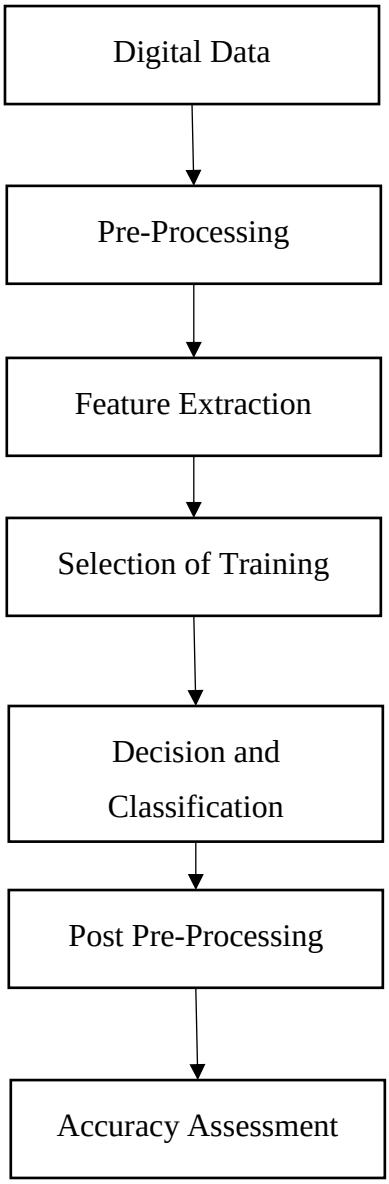

Figure 1. 2: Steps for Image classification

1.2.1 IMAGE CLASSIFICATION APPROACHES

Computerized picture characterization procedures aggregate pixels to speak to arrive cover geologies. Land cover could be rural, forested, urban, and different sorts of elements. There are seven referred to as image classification approaches [5] given as

1) Based on Characteristic Used:

a. Shape-based: This method classify the image on the basis of the 2- dimensional spatial information of the object. Generally the features that are being used for the classification of the object includes points, primitive geometric shapes (rectangle or oval) and skeleton.

b. Motion-based: Motion based approach extracts the static and dynamic characteristic features of moving targets and uses them to assign the targets to one of several predefined categories.

2) Based on Training Sample Used:

Pixels are the littlest unit exemplified in a photo. Picture grouping uses the reflectance insights for isolated pixels. The most two common strategies that has been used on the basis of training sample used are Supervised and Unsupervised Classification [6].

a. **Supervised Classification:** In this approach the prior information regarding the class is known to us. Labels are provided to the different classes and is basically used for the classification process. The new data entry is now going to be classify, that is to which class does it belongs will depend upon the label which has been provided to the different classes.

These label provided by us to the different classes is known as our training data set. Classification is being done on the basis of the spectral signatures available in our training data set. The digital picture classification application determines each class on what it resembles most in the training group.

Supervised Classification Steps: Mostly, For the most part, here are three boss strides required in the run of the mill regulated order methods as takes after:

i. Select training area: In this arrangement, in initial step the expert recognizes in the symbolism homogeneous delegate tests of the diverse surface cover sorts (data sets) specifically noteworthy. These illustrations are expressed towards by method for preparing zones. The variety of appropriate preparing zones depends on the examiner's nature with the topographical territory and their insight into the genuine surface cover sort’s display in the picture. Hence, the expert is "regulating" the order of an arrangement of distinct classes [7]. The arithmetical information introduce particularly in through and through otherworldly groups for the pixels involving these territories are utilized to "prepare" the PC to perceive frightfully comparable regions for each class.

ii. Generate signature record: The PC utilizes an extraordinary program or calculation (of which there are a few varieties), to decide the numerical "marks" for each instructional course.

iii. Classify: Once the PC has decided the marks for each and every class, each and every pixel given in the particular picture is coordinated to these particular marks and named as the class it most nearly "be like" carefully. Appropriately, in a directed grouping we are first distinguishing the data classes which are then used to decide the unearthly classes which speak to them, consequently what we surmise is that the earlier data with respect to the question be characterized is known to us.

b. Unsupervised Classification: In this unsupervised classification no prior information regarding the training samples are performed. It basically look for the patterns in the data but is not going to label it. In this pixels are going to be grouped on the basis of reflectance's, and this grouping of the pixels on the basis of reflectance is known as clustering.

The different strategies that are commonly used for the clustering are K-means and ISODATA [8].The unsupervised image classification strategy comes into play when no specific information regarding the data is available and there is need to look for the pattern in the data for classification.

Unsupervised Classification Stages:

- Create clusters
- Assign classes on the basis of generated clusters

Advantages of Unsupervised classification: It has two huge favorable circumstances over Supervised Classification.

1. The PC could easily allocate specific pixels to spectrally-separate classes that an analyst may perhaps not identify as present.

2. The computer can identify a much larger number of spectrally-distinct classes than an analyst might consider existing.

c. Object Based Classification: Question based picture examination give backings to the use of various groups expected for multi-determination division and in addition order. The way toward dividing the picture into the diverse otherworldly data classes as per the normal groupings found in the information. In this way making utilization of the examples of the known character (i.e. the pixels which has effectively allocated the data

classes) to order the pixels of the obscure personality (i.e. alternate pixels exhibit in the image).It is otherwise called Geographical Object based picture investigation. For example, infrared, rise generally winning shape documents could ready to simultaneously be used to arrange picture objects. Various layers could presumably have system by methods for each other. This setting begins in the plan of neighborhood affiliations, quickness and additionally remove amidst different layers.

Closest neighbor (NN) classifying is very like administered characterization. A while later multi-determination division, the administrator orders test destinations implied for each and every land cover class. The scientific figures to arrange picture objects are all around characterized. The question subordinate picture examination application around then orders objects relied upon their similitude toward the preparation areas and additionally the measurements very much characterized [9].

Object-dependent Adjacent Neighbor Classification Stages:

i. Execute multi-determination division
ii. Choose training zones
iii. Describe measurements
iv. Categorize

3.) Based on Supposition of Parameter on Information:

a. **Parametric classifier:** In this type of classifier the covariance matrix along with the constraints similar to the mean vector are generally used. These parameters are basically generated from the training data set. For example, direct discriminant analysis and Maximum probability.
b. **Non Parametric classifier:** There is no presumption about the information. Non-parametric classifiers don't make utilization of measurable parameters to ascertain class partition. For Instance: Expert system, ANN, choice tree classifier and SVM.

4) Based on the Pixel information Used:

a. **Per pixel classifier:** Traditional classifier creates the spectral signatures by summing up all the spectral information concerned with the pixels present in the training data for the particular feature. Thus the resulting signature contain some portion of spectral components from all the pixels present in the data set. It may be parametric or non-parametric but the accuracy does not meet up the end conditions because of the mixed pixel problem. For example ML, ANN, SVM and Minimum distance.

b. **Subpixel classifiers:** The ghostly estimation of every pixel is thought to be a direct or non-straight blend of characterized unadulterated materials called end individuals, giving relative enrollment of every pixel to each end part. Subpixel classifier has the capacity to deal with the blended pixel issue, reasonable for medium and coarse spatial determination pictures. Case: otherworldly blend investigation, subpixel classifier, Fuzzy-set classifiers.
c. **Per-field classifier:** The per-field classifier approach deals with natural diversity as it increases the classification accuracy. Basically used with the GIS based classification techniques.
d. **Object-arranged classifiers:** In this type of classifiers pixels are grouped together to form object and the classification is than performed on this object. It consists of two parts that is image segmentation followed by the classification of the image. Image segmentation is used to cluster the pixel into the object. For example: e Cognition. T

5) Based on the Number of Outputs for Each Spatial Element:

a. **Hard Classification:** Also known as crisp classification, in which each pixel is required or forced to be the member of the single class.eg ML, SVM, minimum distance, decision tree, and ANN.
b. **Soft Classification:** Correspondingly recognized as fuzzy arrangement and delivers more accurate results as compared to hard classification.

6) Based On The grounds of Spatial Information

a. **Contextual Classifiers:** This picture grouping utilizes the spatially neighboring pixel data. Example frequency based contextual classifier.
b. **Spectral Classifiers:** This picture cataloguing uses pure spectral data Case: Maximum probability, ANN along with minimum distance classifier.
c. **Spectral-relevant classifiers:** This grouping utilizes both otherworldly and spatial data starting characterization pictures are produced utilizing parametric or non-parametric classifiers and afterward relevant classifiers are actualized in the ordered pictures. Illustration: mix of parametric or non-parametric and logical calculations.

7) Multiple classifiers approach

Diverse classifiers have their own particular points of interest and drawbacks. In this approach distinctive classifiers are aggregate. A couple of the strategy for consolidating various classifier which are: Voting rules, Bayesian formalism, evidential thinking, and numerous neural system.

1.3 REMOTE SENSING

Remote detecting is the act of determining information seeing the world's property and additionally water surfaces using pictures acclimatized starting an overhead viewpoint, using electromagnetic radiation in at least one than one regions of the specific electromagnetic range, imitated or transmitted as of the world's surface [10]. This portrayal unquestionably doesn't cover all regions (e.g. meteorological or earthly remote detecting), it ensure serving great as a clarification of remote detecting. Remote sensing makes use of electromagnetic radiation and the finest well-known source of electromagnetic radiations is the sun that emit radioactive heat over the entire electromagnetic spectrum. This natural source of light is used in the remote sensing and its applications.

1.4 TYPES OF SENSING IMAGES

There are various sorts of remote detecting pictures that are as given underneath [11]:

• Multispectral Image: A multispectral picture contains of various groups of data. The groups in multispectral picture are less than 30. This be dependent upon the demand notwithstanding it is constantly under 30 groups.

• Super Spectral image: Late satellite sensors are equipped for catching pictures at numerous more wavelength groups [6]. For instance, a few satellite comprise of 36 phantom groups, packaging the wavelength zones extending initiating the wave infrared, noticeable, short, close infrared, toward the warm infrared. Transmission capacity of the groups is slimmer, allowing the better-quality ghastly components of the goals to be taken by the sensor. The expression "super ghostly" allude such sensors.

• Hyper-Spectral Image: A hyper spectral image comprises of at least hundred spectral bands framing a three-dimensional (two spatial measurements and one unearthly measurement) cube. In this every pixel contained the detailed spectrum. Larger the amount of information associated per pixels will lead to increasing capability to distinguish objects. Due to the increase dimensionality of the Hyper-Spectral Image, image analysis become more complex [12].

Thus efficient algorithms for automatic processing is required. Dimensionality decrease has been generally utilized as a part of hyper-Spectral image classification to lessen information volume and excess.

1.5 APPLICATION OF REMOTE SENSING

Remote detecting application happens in different fields like ranger service, sea and topography. Other imperative applications are:

1. Agriculture
2. Water asset administration
3. Biodiversity Characterization
4. Snow and Glacier Studies
5. Land Use/Cover mapping
6. Disaster checking and alleviation
7. Surveying and urban arranging
8. Coastal Studies
9. National spatial information framework
10. Infrastructure advancement arranging and checking
11. Mineral investigation
12. Telecommunication
13. Coral and Mangroves Studies
14. Wasteland Mapping

These are a portion of the exploration regions of researcher by help of multispectral pictures which are created by remote sensors.

1.6 FEATURE EXTRACTION IN HYPERSPECTRAL IMAGE

Data Information with an extensive number of elements will bring about higher computational cost, and the repetitive irrevealent elements may likewise crumble the order execution. Different methodologies have been proposed to manage the hyperspectral information. Lee and Land grebe proposed techniques for include extraction in light of choice limits that expand division of information in numerous two-class issues [7]. Jia and Richards built up the (SPCT) whereby the first groups are assembled into subsets of profoundly associated neighbouring groups to which the K-L change is connected [8]. Kumar et al. explored band joining strategies, motivated by best-basis functions as a means of feature extraction in a pairwise classifier framework [9]. From the previous discussion, it can be found that feature extraction methods are becoming important for classification of hyperspectral data with large dimensions. The primary reason of the feature extraction is to reduce the dimension without any critical loss of data loss of data. Highlight extraction is to discover the change from a higher measurement to

a lower dimensional element space with the greater part of the coveted data content safeguarded [9]. This change might be a straight or nonlinear mix of the first factors and might be administered or unsupervised [10]. The regularly utilized element extraction systems incorporate PCA [13], Independent Component Analysis (ICA), discriminant boundary feature extraction, DBFE [11]

Hyperspectral picture has exceptionally solid ghostly connection that implies there is huge measure of excess information and misdirecting data introduce in the picture and such data is should be evacuated before any additional data can be done. Consequently it is frequently helpful to decrease information dimensionality. Dimensionality decrease systems can be change based or determination based. The key refinement between these two is whether a dimensionality lessening procedures change or safeguard the importance of the first informational index during the time spent decrease. The decision of dimensionality lessening methods is regularly guided by the specific issue within reach.

1.6.1 TRANSFORMATION- BASED REDUCTIONS

Transformation based decrease approach is a semantics annihilating dimensionality diminishment strategy, which in the process irreversibly changes the first dataset highlights. These techniques are regularly utilized in the circumstances where the semantics of the first datasets are not required by additionally handle. Many methodologies, for example, foremost part examination, least clamor division change, discriminant investigation, non-parametric weighted component extraction [14] wavelet change and phantom blend examination have been utilized for highlight extraction in hyperspectral pictures so as to diminish information repetition or to extricate particular data.

This segment quickly examines a few such prominent procedures that have been utilized for hyperspectral pictures. Following are some of well-known component extraction techniques utilized for hyperspectral pictures.

1. Principal Component Analysis (PCA)
2. Minimum Noise Fraction (MNF)
3. Discriminant Analysis
4. Linear Spectral Mixture Analyser (LSMA)
5. Wavelet-based Methods

- **Principal Component Analysis**

  PCA is a technique used to compress a large amount of data set into smaller dataset. Mostly it minimizes the mean square error by compressing the data and by using orthogonal transform it will maximize the variances and de-correlation. These uncorrelated variables are known as principle components [15]. It will find applications in different fields like in image processing, data compression etc.

- **Minimum Noise Fraction (MNF)**

  MNF is an information change procedure. In hyperspectral picture preparing, MNF is regularly used to adjust the information along tomahawks of diminishing sign to commotion proportion (SNR). Since, hyperspectral pictures have a tendency to contain more clamor, MNF is a method to address the commotion exhibit in the picture. MNF expands the evaluated SNR. Hardly any commotion channels are utilized like Adaptive Filters (AF), and so forth. The fundamental preferred standpoint of Minimum Noise Fraction is that the MNF of a picture (i.e. higher request groups in hyperspectral picture) can be truncated consolidating low SNR that wipes out a great part of the commotion without losing much data. The principle burden is that if truncation comprises of excessively numerous groups, a lot of clamor is left in the picture. In the event that truncation comprises of excessively few groups, valuable flag might be barred from the subsequent picture.

- **Linear Discriminant Analysis**

  Linear Discriminant Analysis (LDA): Fisher's direct discriminant investigation is appropriate when information is ordinarily circulated. It is a parametric component extraction technique. LDA depends on following three diffuse grids: inside class, amongst class and blend (or aggregate) disseminate frameworks which can be utilized to plan criteria of class distinguishableness. Changes the information from a high dimensional space into low dimensional subspace with boosted class distinguishableness. In hyperspectral pictures, change is finished utilizing the proportion of number of preparing tests and phantom elements. The principle advantage is that it is sans dispersion, Fast and simple to apply and Gives better inside class scramble networks The fundamental hindrance is that the rank of between-class scramble network is the quantity of class’s short one. In this manner, the greatest number of components that can be separated continues as before which may not be adequate for accomplishing better precision in down to earth application. It can't perform well for non-ordinarily appropriated information. In hyperspectral picture preparing, the proportion of the quantity of preparing tests and elements is little because

of the way that marking is costly and that the component space has high dimensionality. Subsequently, it experiences the peculiarity issue since inside class covariance is particular.

- **Linear Spectral Mixture Analyser (LSMA)**

  Direct Spectral Mixture Analysis is broadly used to address the blended pixel issue. In hyperspectral picture investigation, LSMA is ordinarily utilized for subpixel location and blended pixel characterization. It depends on the presumption that the picture pixel is straight blended by materials (alluded to as targets). It decides the segment parts of blended pixels and the ghostly qualities of the end individuals where end individuals are the "immaculate" spectra comparing to each of the classes. The principle burden is that it is touchy to target information and clamor and wealth portion of targets introduce in a picture are thought to be obscure that should be assessed. To address these issues a few calculations are adjusted and created like Independent part examination (ICA) based direct ghostly irregular blend investigation (LSRMA) accepting plenitude portion as arbitrary.

- **Wavelet-based method**

  Wavelet-based diminishment for hyperspectral symbolism depends on the way that the phantom information from unique component space is changed to a decreased element space. The primary guideline of this strategy is to apply Discrete Wavelet Transform (DWT) to hyperspectral information in ghostly area and at every pixel area. Here, wavelet lessened information speak to a phantom dispersion like the first appropriation however in packed shape.

  The principle weakness is that it is a lossy pressure and wavelet based technique evacuated the high recurrence flag which may contain the helpful data for class division and ID. In this way, two noteworthy issues emerge from such change based dimensionality decrease methods. One is the quantity of information measurements required for dimensionality diminishment to maintain a strategic distance from huge loss of data. The other is that, since the information are changed, some urgent and basic data may have been traded off and twisted. Subsequently, the change based dimensionality lessening procedure more often than not changes the physical importance of the first information.

1.6.2 SELECTION-BASED REDUCTION

Determination based decrease approach is a semantics protecting strategy (alluded to as highlight choice or band choice in hyperspectral pictures) which endeavor to hold the

significance of the first dataset. The primary point of determination based dimensionality decrease strategy is to locate an insignificant subset of the first information without losing their physical significance. In hyperspectral pictures, to decrease the dimensionality of hyperspectral information, band choice method is ordinarily utilized. Band choice in hyperspectral pictures allude to choose a subset of unique groups containing most helpful data about the question, expelling excess, unimportant or deluding ghostly components from the hyperspectral information. As indicated by the accessibility of class data, band determination is ordered into two sorts: unsupervised band selection and supervised band selection.

- **Unsupervised Band Selection**

  At the point when class data is not known that is no priori information is available to us than unsupervised band selection comes into play and find the most informative and distinct band. These methods can be categorized based on the following classification as shown in Table 1.1

- **Supervised Band Selection**

  When class information is already known to us that is priori information regarding the class is known to us. It generally comprised of two major steps: subset generation and subset evaluation. Subset era is only a procedure of heuristic inquiry, where each point in the hunt space, indicates a competitor subset for assessment. Band determination in hyperspectral pictures requires a proficient seeking technique. Ideal pursuit systems, for example, Exhaustive Search or Branch and Bound are not relevant for hyperspectral pictures because of the gigantic measure of multifaceted nature, since in these methods all conceivable band mixes are tested and afterward ideal one will be chosen. Thus, to stay away from this, different suboptimal techniques, for example, SFS, SBS, SFFS and SBFS are utilized.

  Different stochastic enhancement based element (band) choice procedures for hyperspectral pictures have been utilized, for example, Genetic Algorithms (GA), Particle Swarm Optimization (PSO) and Ant Colony Optimization (ACO).

Table 1.1: Unsupervised band selection methods for hyperspectral images

Categories	Characteristics
Methods based on Information Theory	It incorporates the utilization of data hypothesis like Mutual information. Gives a system to discover correlation between different bands among picture groups. Disparity measures can be utilized to compare different probability distributions [17].
Methods based on Constrained Band Selection (CBS)	It is not quite the same as variance based techniques or data theoretic criteria-based strategies. This approach straightly obliges a band to limit the relationship or reliance of this specific band from alternate groups in a hyperspectral image [16, 17].
Methods based on eigen analysis and decorrelation of the input data	Most extreme Variance Principal Component Analysis (MVPCA) is one of the procedure in view of eigen investigation and decorrelation of information. In this, a joint band-prioritization and band decorrelation approach is embraced for hyperspectral picture order. The band prioritization relies upon an eigen investigation, breaking down a lattice into an eigen shape network from which a stacking factors grid has been built and used to organize groups [16, 18].
Methods based on clustering strategy	Clustering process system aggregate groups that limits the intra-group change and maximize the inter-cluster difference based on band information content such as mutual information, etc. [19].

In light of various assessment criteria, supervised techniques can be classified into two approaches: filter approach and wrapper approach [20].

- **Filter Approach:**

  Band choice technique in view of filter approach is reliant on general attributes of the information to assess and select component (band) subsets without including any order calculation so undesirable elements are overlooked before the grouping procedure starts. It utilizes some statistical techniques (i.e. separation measures, for example, the Mahalanobis Distance, Bhattacharyya Distance, and Jeffreys-Matusita Distance to distinguish an ideal subset of groups in hyperspectral image.
- **Wrapper Approach:** Band determination strategy in view of wrapper approach is reliant on predetermined classification algorithm and uses its performance as an evaluation criterion. It decides highlights (groups) which better suit the arrangement calculation as far as execution. Highlight Shaving is a case of this sort which expels the entire band including additional parameters, for example, most extreme size of smoothing window, a couple of keeps running of entire technique, a specialist learning, and so on.

1.7 OPTIMIZATION TECHNIQUES USED IN FEATURE EXTRACTION

The feature would be optimized in the research using nature based algorithms and these are as follows: Genetic Algorithm, Swarm intelligence algorithms namely ABC algorithm, Particle Swarm optimization, Ant Colony Optimization and Firefly Algorithm. In this Particle Swarm has been used as an Optimization Technique.

1.7.1 PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization (PSO) algorithm is developed by Jim Kennedy and Russ Eberhard in 1995, inspired by the social behaviour birds and fishes .It is a nature based algorithm which combines the self-experience with the social experience. Swarm basically refers to the group of insects or birds moving together. Their movement is such that they do not interfere with each other and occupy a particular position in space.

Even if they had to change the direction they adjust their position in such a way that no two birds interfere with each other. They adjust their position according to their own flying experience as well as the experience of the other birds. This type of behaviour shown by the

birds has led to the development of the optimization technique known as Particle Swarm Optimization (PSO).

The PSO algorithm is population-based optimized technique and its main aim is to find the global optimal solution for the real value function that is fitness function defined for the particular given search space.

1.7.2 ARTIFICIAL BEE COLONY OPTIMIZATION

The Artificial Bee Colony (ABC) calculation is a swarm based meta-heuristic calculation that was presented by Karaboga in 2005 for improving numerical issues. It was motivated by the shrewd rummaging conduct of bumble bees. The calculation is particularly in light of the model proposed by Tereshkova and Loengarov (2005) for the scavenging conduct of bumble bee settlements. The model comprises of three basic segments: utilized and unemployed searching honey bees, and nourishment sources. The initial two segments, utilized and unemployed scavenging honey bees, look for rich nourishment sources, which is the third part, near their hive. The model likewise characterizes two driving methods of conduct which are vital for self-arranging and aggregate knowledge: enlistment of foragers to rich sustenance sources bringing about positive input and deserting of poor sources by foragers causing negative criticism.

In ABC, a settlement of simulated forager honey bees (operators) look for rich manufactured sustenance sources (great answers for a given issue). To apply ABC, the considered streamlining issue is first changed over to the issue of finding the best parameter vector which limits a goal work. At that point, the counterfeit honey bees arbitrarily find a populace of beginning arrangement vectors and afterward iteratively enhance them by utilizing the procedures: moving towards better arrangements by methods for a neighbour seek component while surrendering poor arrangements.

1.7.3 ANT COLONY OPTIMIZATION

Ant colony optimization (ACO) is populace based absolutely metaheuristic that can be utilized to find rough answers for intense streamlining issues. In ACO, an arrangement of programming program retailers called manufactured ants search for right answers for a given enhancement issue. To utilize ACO, the advancement inconvenience is changed over into the inconvenience of finding the fine course on a weighted chart. The engineered ants (from now on ants) incrementally manufacture arrangements by method for moving at the diagram. The arrangement creation strategy is stochastic and is one-sided by methods for a pheromone form,

this is, a settled of parameters identified with chart added substances (either hubs or edges) whose esteems are adjusted at runtime by utilizing the ants.

1.8 CLASSIFICATION TECHNIQUES

The Classification is basically the algorithm and the techniques that has been designed to automatically group the pixels in the image into different classes or themes. Multispectral data is being used for classification and the spectral signature has been used as the basis of the categorization. In today world, we cannot say that the given classifier is best for all the situation because the characteristics of the image and the circumstances under which the categorization has been made varies a lot. In this work two classification techniques known as Support vector machine (SVM) and Artificial neural network (ANN) are used.

1.8.1 SUPPORT VECTOR MACHINE (SVM)

SVM is the supervised learning algorithm that is being used as a regression and a classifier but it is mostly used as a classifier for solving the classification problems. Classification in general can be viewed as separating the different classes present in the feature space. SVM can be used to classify or able to discriminate between the different classes whether they are linearly separable or non-linearly separable. SVM was invented by Vladimir Vapnik. On the basis of data whether it is linear or non-linear it is of two types that is Linear SVM and Non-Linear SVM. Although the linear SVM [21] gives higher accuracy but both of them have the advantages of their own.

1.8.2 ARTIFICIAL NEURAL NETWORK (ANN)

ANN (Artificial Neural Network) is the most important and the simple network which has been designed according to the neural structure of brain. Just like our brain consists of large number of neurons which act as a conducting or transporting medium to stimulate the respective body organ. Neural network also works in such a fashion. It basically consist of three layers and these are: Input layer, hidden layer and output layer. On the basis of the output layer weights are assigned to get output accordingly to input layer.

1. **Input layer:** This layer is the one through which our input trained data is being fed
2. **Hidden Layer:** All the processing of the input trained data is being performed here
3. **Output Layer:** After the processing of trained data in the hidden layer, the classification results are being taken out from this layer.

1.9 ORGANISATION OF THE DISSERATION

The Dissertation is organized in the form of chapters as discuss on the following pages

Chapter 1: Introduction, it deals with the overview of image classification, the various feature extraction techniques used in hyperspectral images, optimization algorithms and different classification techniques.

Chapter 2: Literature Survey, deals with the study of various existing techniques on the feature extraction in hyperspectral image, optimization techniques and classification techniques has been done. It also deals with the problem statement i.e. what is the basis and need of this work

Chapter 3: In this chapter we will discussed detail explanation regarding the various techniques and algorithm that has been used by us for the classification of hyperspectral image, objectives and flowchart of methodology so as to achieve satisfactory results.

Chapter 4: Results and Discussion, contains the observation and the results comparing the classification accuracy obtained from the different techniques.

Chapter 5 Conclusion and Future Scope, The work done has been concluded in this chapter and based on the observations future scope is given.

CHAPTER 2: LITERATURE SURVEY

2.1 Introduction

This chapter overviews the methods and work presented by various researchers in the domain of Hyperspectral using various methods and techniques, along with their various design methods from time to time. Various methods and techniques have been discussed, which provide an insight into the field of Hyperspectral Image. This survey provide ample emphasis in field of Hyperspectral Image thus provide objective for the dissertation. The Literature Review establish the need for additional research and define the problem statement.

2.2 Literature Review

Long *et al.* [22] in 2017 proposed anew localization framework, which was further classified into three processes: region proposal, classification, and accurate object localization process. For detecting all desired objects within the image a region method was used. Then a convolution neural network was used to extract the features. For increasing the accuracy, unsupervised score based bounding box regression algorithm was used. The proposed algorithm was very simple and robust.

Tangthaikwan *et al.* [23] in 2017 presented a method for the classification of land set. Multi-sensor scanner (MSS) are used to recognize the area of land use. The desired image is firstly pre-processed and then classifications using support vector machine has been used along with radial basis function (RBF) of SVM. 3×3 square neighbourhood data has been used. The result from simulation shows that a RBF kernel with a sigma value of 1.7 gives a higher correct classification rate.

Ushanandhini *et al.* [24] in 2016 proposed a new Gaussian normal affinity to illustrate the closest neighbour by ensemble process in a better way that ensures the reliability and diversity. Author also has applied wavelet transformation to texture feature for increasing the accuracy of classification.

Dixit *et al.* [25] in 2016 deals with the supervised classifiers namely minimum distance, support vector machine, maximum likelihood, and parallelepiped. Performance parameters like accuracy and Kappa coefficient has been calculated for measuring their performances. The overall accuracy obtained for the research work was 88% and Kappa coefficient was .82.

Chebbi *et al.* [26] in 2016 presented an approach used for storing and processing a huge satellite images by using Hadoop MapReduce framework. The simulation results showed that the proposed system provide an optimized results even when deals with the huge data volume.

Richards JA *et al.* [7] in 1999 proposed principal components transformation (PCT) for efficient hyper spectral remote-sensing image classification and display. The scheme requires, initially, partitioning the complete set of bands into several highly correlated subgroups. After separate transformation of each subgroup, the single-band separability are used as a guide to carry out feature selection. The selected features can then be transformed again to achieve a satisfactory data reduction ratio and generate the three most significant components for colour display.

The scheme reduces the computational load significantly for feature extraction, compared with the conventional PCT. A reduced number of features will also accelerate the maximum likelihood classification process significantly, and the process will not suffer the limitations encountered by trying to use the full set of hyperspectral data when training samples are limited. Encouraging results have been obtained in terms of classification accuracy, speed, and quality of colour image display using two airborne visible/infrared imaging spectrometer (AVIRIS) data sets.

Jiao L *et al.* [13] in 2009 represents the another feature extraction technique in view of resistant clonal choice (ICSA) and PCA is proposed for characterization of Hyperspectral remote detecting picture. As the Hyperspectral remote detecting picture is procured in extremely limit phantom channels, the subsequent high-dimensional capabilities may contain excess data. Along these lines, include extraction is important to arrange an information with expansive measurement, for example, the Hyperspectral remote detecting picture. PCA is prominently utilized for include extraction. In any case, in conventional PCA, choosing the bigger eigenvectors as important parts infers data misfortune. There is no precise approach to figure out which Principal Components (PCs) ought to be utilized. Another element extraction model to choose the ideal central segments utilizing ICSA is created. The information obtained by the NASA airborne AVIRIS instrument over the Kennedy Space Center, Florida is utilized for assessment. Test comes about demonstrate that our technique can improve comes about

Goel [27] in 2012 introduced a dynamic model of the mixed biogeography based improvement (BBO) for arrive cover attributes extraction. In the mixed BBO, the living spaces embody the competitor issue arrangements and the species movement exemplifies dissemination of

qualities (SIVs) among candidate arrangements as expressed by the fittingness of the natural surroundings which is called their HSI esteem.

Notwithstanding, it is accepted that these particular SIVs which are the measure of arrangement highlights, will as a rule remain steady for each natural surroundings and the HSI for every living space depends just on the migration and the displacement rates of species. This paper broadens the mixed BBO by considering the measurement that the specific no. of SIVs or else the choice factors potentially won't stay steady for all competitor arrangements (natural surroundings) that are a piece of the overall territory. Meanwhile the physiognomies of every single natural surroundings veer fundamentally thus, coordinating every one of the living spaces using the indistinguishable arrangement of SIVs might be misdirecting and furthermore may not prompt an ideal clarification.

Hereafter, in their dynamic prototypical, they consider the way that HSI of an answer is influenced by factors other than movement of SIVs which are arrangement attributes, in like manner. These extra perspectives can be demonstrated as a few meanings of HSI of a living space, each single definition in view of a various gathering of SIVs which mimics the impact of these extra angles.

They likewise decide the sanctioning of the prescribed model by running it on this present reality issue of land cover highlight extraction in a multi-ghostly satellite picture. They exhibit its execution on the dataset of Alwar area in Rajasthan where it turns out to be an effective element extractor as an expansion to the first biogeography based land cover include extractor.

Parminder Singh et.al. [28] in 2011 arrangement another figuring which relies upon Bacterial Foraging Optimization which is used to organize the satellite picture. Immediately we use a swarm data gathering procedure in light of sprout treatment by made bumble bees (FPAB) to bundle the satellite picture pixels. Those gatherings will be furthermore orchestrated using BFO. This new procedure shows an upgraded exceedingly exact results for the gathering of satellite picture when the proposed count is used. The proposed count has been executed to approve the multi assurance, multi-spooky, and multi-sensor pictures of Alwar region in Rajasthan.

Panchal *et.al.* [29] in 2009 concentrated on grouping of the satellite picture of a specific land cover utilizing the hypothesis of Biogeography based Optimization. The extraordinary BBO framework does not eat up the inbuilt property of grouping which is required amid picture characterization. Consequently modifications have been expected to the first calculation and

the adjusted calculation is utilized to arrange the satellite picture of an accepted zone. The outcomes assign that to a great degree correct land cover elements can be extricated adequately when the proposed calculation is utilized. They have taken a multi-unearthly, multi determination and multi-sensor picture of Alwar zone in Rajasthan.

Martínez-Usó et al. [19] in 2007 presents a strategy for dimensionality reduction to manage hyperspectral images. The proposed strategy depends on a progressive clustering structure to limit the intra cluster change and maximize the inter cluster variance. This point is sought after utilizing data measures, such has separations in view of shared data or Kullback–Leibler dissimilarity, with a specific end goal to diminish information repetition and non-helpful data among picture groups. The method that is exhibited has a steady conduct for various picture informational indexes and an observable exactness, for the most part while choosing little arrangements of groups.

Gupta *et.al*. [30] in 2011**,** proposed methodical blueprint of film registering technique in arrive cover highlight extraction. Layer registering is a novel branch of common calculation that has a gigantic arrangement of conveyance and additionally holds maximal parallelism. The bio-roused technique is used for picture arrangement notwithstanding these particular pictures are known as remote detecting satellite picture. The territory scenes like vegetation, water, rough, fruitless, notwithstanding urban are required to be ordered as their information mastermind gigantic help all through climatic behavioural changes, catastrophic event, and also in a few further locales of ecological varieties.

They have very much characterized imperatives of layer registering in arrive cover highlight reflection terms thusly showing that the specific figure of speech of film processing is significant for picture grouping notwithstanding then they have anticipated a strategy of P structure. The prescribed technique is actualized to Alwar district of Rajasthan of 472X576 estimation that additionally contains 7 Band Indian Resources at Satellite Digital Numbers. The technique has secured about a wide range of the landscape structures of these specific zone. It shows basically 99% skill on vegetation and also water range. The KHAT figures shows that recommended system has a total capability of around 0.68812.

Chang *et al*. [16] in 2006 utilize Constrained vitality minimization (CEM) has indicated successful in hyperspectral target recognition. It straightly obliges a coveted target signature while limiting meddling impacts caused by other obscure marks. This paper investigates this thought for band choice and builds up another way to deal with band determination, alluded to as constrained band selection (CBS) for hyperspectral imagery. It deciphers a band picture as

a coveted target signature vector while considering other band pictures as obscure mark vectors. Accordingly, the proposed CBS utilizing the idea of the CEM to straightly oblige a band picture, while additionally limiting band relationship or reliance given by other band pictures, is alluded to as CEM-CBS.

Johal *et.al*. [31] in 2010 manages picture grouping by methods for using swarm processing system. In this work, they used a novel swarm information grouping technique subordinate upon bloom fertilization by methods for simulated honey bees toward bunching the satellite picture pixels. The goal of bunching was to part a gathering of information focuses into self-comparable accumulations. Those particular bunches will maybe be more grouped using Biogeography Based Optimization. The results point towards that to a great degree exact grouping of the satellite picture is achieved by means of using the proposed strategy. The calculation has been connected to multi-phantom, multi determination and multi-sensor 7-band picture of Alwar Town (in Rajasthan, India).

Lorenzo Bruzzone *et al*. [20] in 2008 presents a novel way to deal with feature extraction determination for the arrangement of hyperspectral images. The proposed approach goes for choosing a subset of the first arrangement of components that displays two principle properties: i) high capacity to separate among the considered classes, ii) high invariance (stationarity) in the spatial space of the explored scene. The element determination is proficient by characterizing a multi target measure that considers two terms: i) a term that surveys the class distinctness, ii) a term that assesses the spatial invariance of the chose highlights.

Kumar *et.al*. [32] in 2013 arranged the photos swarm knowledge ends up being the generally available event by means of the exploration researchers as a result of its essential quality notwithstanding adaptability concerning the work. In a line of this exploration, they are putting forth a crisp approach toward grouping a photo by methods for having the likelihood view of bumble bees. This paper mostly centres upon the likelihood event of nectar hamburger toward picking their particular nectar quality. On that premise, they will effortlessly arrange their photo objects. In this paper, they offered a likelihood based bumble bee technique as a capable classifier utilized for high determination multi-ghastly satellite picture of Alwar area. The particular kappa coefficient of 0.941 also bolsters their calculation's capability at this particular time.

Kaur *et.al*. [33] in 2012 proposed PSO system. In their work they have utilized dataset of Alwar area. At the point when acclimatized with old-style bunching methodology, difficulties for example nearby optima and in addition affectability towards instatement are consolidated,

subsequently finding a more prominent district using overall pursuit. This fragmented picture is later arranged using coefficient of Kappa

2.3 PROBLEM STATEMENT

Classification is the process in which the data is classified on the basis of the reflectance property or you can say that the spectral signatures. The spectral signature being the unique identity. By comparing the spectral signatures obtained with that of the known material we came to know about the data .This has been widely used in remote sensing studies to classify about the land covers, forest area, barren land, vegetation area that is being available over the particular region and keep a record of this for future use, to know that how the particular area has been affected with time and due to the environmental changes.

At present, classification is being performed to acquire the information regarding the land use/cover. Various method has been developed in the past few years for the remote sensing, but none of them has been achieved good accuracy. Therefore we have proposed the technique which will enhance the accuracy using various algorithms like for feature extraction process principle component algorithm (PCA), optimization process particle swarm optimization (PSO) and classification techniques like artificial neural network (ANN) [34] and support vector machine (SVM) [35]. Following metrics will be used for the simulation of the work and for the comparison of the proposed work with the existing work:

i. **Peak signal to noise ratio (PSNR):** PSNR is defined as the ratio of the signal power to that of the noise power of the signal that has been corrupted by noise. Greater the PSNR more easily it become to recover the signal. It is measured in dB It depends upon MSE that is lesser is the mean square error greater will be the PSNR.

ii. **Mean square error (MSE):** Mathematically MSE is defined as

$$\text{MSE}=\frac{1}{PQ}\sum_{i=1}^{m}\sum_{j=1}^{n}\left(B_{ij}-C_{ij}\right)2 \quad (3.1)$$

Here, B_{ij} is perfect image, C_{ij} is the multilayer image.

iii. **Classification Accuracy**

Classification accuracy is parameter which tell us how accurately we have classified the data and up to what extent it has been accurately classified. It depends upon the Mean Square Error

CHAPTER 3 : OBJECTIVE AND METHODOLOGY

3.1 OBJECTIVES

The main objective and aim of the thesis work is being summarize as follow

a. To study the various existing works on the satellite image classification techniques.

b. To propose a better classification technique we present an appropriate feature extraction using PCA.

c. To optimize the PCA feature use PSO algorithm with a novel objective function.

d. To classify the bands and reduce the error rate use artificial neural network.

e. To compare the existing work with proposed work, we calculate the performance matrices like PSNR, SSIM, MSE, and error index.

3.2 METHODS OF CLASSIFICATION OF HYPERSPECTRAL IMAGE

3.2.1 PRINCIPAL COMPONENT ANALYSIS

PCA is a technique used to compress a large amount of data set into smaller dataset. Mostly it minimizes the mean square error by compressing the data and by using orthogonal transform it will maximize the variances and de-correlation. These uncorrelated variables are known as principle components. It will find applications in different fields like in image processing, data compression etc.

It is a ` data reduction technique developed by Hoteling in 1933. PCA transforms the original data matrix or data set into other dimensions of reduced components such that the orthogonality of the components are preserved. With the help of PCA correlated structure is converted into the uncorrelated data structure but with the reduced dimension. PCA works in the following fashion. For the multidimensional random vector x we need to find the linear transform W such that the components obtained are uncorrelated.

$$y = Wx \tag{3.1}$$

Such that

$$\Sigma_y = E\{(y - E\{y\})(y - E\{y\})^T\} \tag{3.2}$$

Where Σ_y represents the covariance matrix (comes out to be diagonal). The vector of the expected value of Y that is $E\{y\}$ and Σ_y can be expressed in terms of the vector of expected value and covariance of random vector$x$.

$$E\{y\} = E\{Wx\} = WE\{x\} \tag{3.3}$$

$$\Sigma_y = W\,\Sigma_x\,W^T \tag{3.4}$$

A_x Represents the matrix corresponding to normalized Eigen vectors of the covariance matrix Σ_x such that

$$C_x = A_x^T\ \Sigma_x\,A_x \tag{3.5}$$

The C_x represents the diagonal Eigen value matrix. Since in eq. (3.2).The Σ_y should be diagonal so we observed that $W = A_x^T$ leads to the PCA solution.

$$y = A_x^T\,x \tag{3.6}$$

The components of y represents the principal components and the Eigen vector matrix A_x is called principal component transform (PCT)

PCA allows calculating a linear alteration that maps information as of a high dimensional space to a lower dimensional space.

$$B_1 = t_{11}a_{11} + \cdots \ldots \ldots t_{1n}a_{1n} \tag{3.7}$$

$$B_2 = t_{21}a_{11} + \cdots \ldots \ldots t_{2n}a_{2n} \tag{3.8}$$

Geometrical interpretation

- PCA projects the data along the directions where the data varies the most.
- These directions are determined by the eigenvectors of the covariance matrix corresponding to the largest Eigen values.
- The magnitude of the Eigen values corresponds to the variance of the data along the eigenvector directions.
- To choose the principal components
- As shown in the figure 3.1. Vector x can be reconstructed using it's the principle component.

PCA is not always an optimal dimensionality-reduction procedure for classify- action purposes. Principal Component Analysis is the tool for extracting the features from the image. It forms the feature vector according to the variance and arrange them in the descending order. The variance value are also known as Eigen value. Lower Eigen value can be neglected as most of the information are lie on the starting values, these values are also known as Principal Components. In this way the Eigen values with the highest variance are selected and that with the lower values are neglected and the dimensionality is reduced.

It is a classic method used to compress higher dimensional data to lower dimension data with reduced components such that the orthogonality of the components are preserved. Various steps which are being followed in the Principal Component Analysis has been explained in detail as below.

Steps involved in the Principal Component Analysis

Step 1: Get normalizes data of the image vector.

Step 2: The next step is to take the mean of the image vector. The mean can be calculated using the expression

$$\mu = (\sum_{i=1}^{n} Xi)/n \tag{3.9}$$

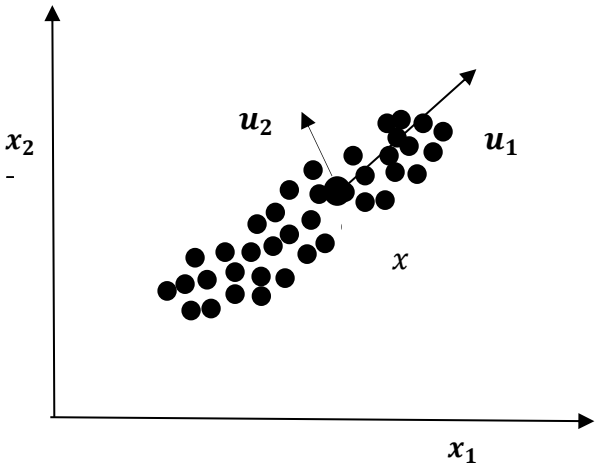

Figure 3.1 PCA Geometrical Representation

where μ represent the mean image which corresponds to the common feature obtained and to get the unique feature we has normalize the mean and the normalized mean is achieved by subtracting the μ from the entire image

$$\tilde{X}_i = X_i - \mu \tag{3.10}$$

Where $\tilde{X}_i$ represents the normalized mean or the unique feature

Step 3: Compute the covariance matrix.

$$\mathrm{C} = cov(\tilde{X}_i) \tag{3.11}$$

Step 4: Analyse the Eigenvectors and Eigen values of the covariance matrix.

Step 5: Eigen values obtained are arranged in the descending order and the value with the highest value represent the component containing large amount of information.

$$F_v = \left(E_1, E_2, E_{3,\ldots\ldots\ldots\ldots\ldots\ldots} E_n\right) \tag{3.12}$$

Where F_v is Feature Vector and $\left(E_1, E_2, E_{3,\ldots\ldots\ldots\ldots\ldots\ldots} E_n\right)$ represents the Eigen values arrange in the descending order.

Step 6: Derive the new data set once we have chosen the components, we simply take the transpose of the vector and increase it on the left of the original data set, transposed.

$$F_d = R_{fv} * R_{da} \tag{3.13}$$

where F_d is the final data and R_{fv} is the Row Feature Vector matrix with the eigenvectors in the columns transposed so that the eigenvectors are now in the rows, with the most major eigenvector at the top, and R_{da} is the Row Data Adjust is the mean-adjusted data transposed.

When no prior information is available, then there is need to carry out unsupervised characteristic extraction. In this example, the data and distance measures among lessons cannot be computed or predicted and the primary aim of function extraction shifts to reduction of information redundancy. The slim bandwidths associated with the spectral bands of hyperspectral statistics result in correlation between the adjacent bands ensuing in an especially excessive stage of redundancy inside the statistics. Primarily based in this commentary, you will really continue to perform characteristic choice by using analysing the correlation matrix and selecting only a few bands from each institution of notably correlated bands. A better approach is to convert the records such that the ensuing functions are decorrelated.

The variance of the individual components is taken into consideration to be an indicator of information content; massive values advocate high tiers of facts and low values indicate the presence of generally noise. Primarily based in this, only the functions with excessive variance are selected for similarly processing. In case of PCA [15], the separability is completed through decorrelation. While considering hyperspectral statistics, each band corresponds to a factor of random pixel vectors and constitutes a feature of the data. On this context, the picture cube ensuing after PCA processing has the bands decorrelated and looked after in step with their variance. This indicates that most of the statistics may be retrieved from the primary few features and that maximum of the ultimate features have near zero variance (indicating a loss of statistics)

Based totally on this statement, PCA is regularly used for function extraction by way of losing the bottom variance components. Decorrelation based characteristic extraction has been

discovered to be much less green while managing small classes (i.e. lessons having small spatial extents) in the photo. In other phrases, small lessons are sparsely represented in the facts (i.e. small lessons include only a few pixels). Due to their length, these instructions tend to have little impact on the band variance leading to the opportunity of being discarded within the decrease variance bands. In the context of goal detection, loss of data regarding small objectives (that correspond to small lessons in the picture) impacts the accuracy of the function extraction algorithms. Consequently, there is no assure that band reduction the use of PCA will successfully hold the complete data content of the photograph cube.

3.2.2 PARTICLE SWARM OPTIMIZATION

The Particle Swarm Optimization (PSO) algorithm is developed by Jim Kennedy and Russ Eberhard in 1995, inspired by the social behaviour birds and fishes .It is a nature based algorithm which combines the self-experience with the social experience. Swarm basically refers to the group of insects or birds moving together. Their movement is such that they do not interfere with each other and occupy a particular position in space.

Even if they had to change the direction they adjust their position in such a way that no two birds interfere with each other. They adjust their position according to their own flying experience as well as the experience of the other birds. This type of behaviour shown by the birds has led to the development of the optimization technique known as Particle Swarm Optimization (PSO).

The PSO algorithm is population-based optimized technique and its main aim is to find the global optimal solution for the real value function that is fitness function defined for the particular given search space. It uses a number of agents also known as particles which constitute the swarm, moving around the search space looking for the best solution. A single particle will not provide us with the accurate solution thus when the particles are working in the group they are able to search or find an optimized solution. This process is repeated until, we will get a better solution for the problem.

Each particle in the search space adjusts its flying according to its own flying experience as well as the flying experience of the other particles. Each particle keep the track of its best solution that is personal best (pbest) and the best value of the any particle that is the global best (gbest). The particles modified their position according to its current position, current velocity, the distance between the current position and pbest and the distance between its current position and the gbest. This is shown in the eq. (3.11).

$$v_k^{i+1} = v_k^i + c_1 r_1 (p_k^i - x_k^i) + c_2 r_2 (p_k^g - x_k^i) \quad (3.14)$$

Where c_1 refers to the acceleration factor related to pbest, c_2 refers to the acceleration factor related to gbest, $r_1 and\ r_2$ represent the random number whose value lies between 0 and 1 , v_k^{i+1} is the modified velocity, v_k^i is the current velocity , p_k^i represent personal best, p_k^g represent global best and x_k^i represents the current position. The PSO can be more easily understood from the flowchart as shown below in the figure 3.2.

3.2.3 SUPPORT VECTOR MACHINE (SVM)

SVM is the supervised learning algorithm that is being used as a regression and a classifier but it is mostly used as a classifier for solving the classification problems. Classification in general can be viewed as separating the different classes present in the feature space. SVM can be used to classify or able to discriminate between the different classes whether they are linearly separable or non-linearly separable. SVM was invented by Vladimir Vapnik. On the basis of data whether it is linear or non-linear it is of two types that is Linear SVM and Non-Linear SVM. Although the linear SVM [21] gives higher accuracy but both of them have the advantages of their own.

Consider the binary classification problem of linear SVM in which the two classes has to be classified, than what will be the optimal solution or the line that can separate the classes in more efficient and accurate way. The best solution to this problem is to obtain the optimal separating hyperplane. Now the question arises what is hyperplane. Hyperplane is a decision surface that separates or split the space in two parts. Consider the equation (1.15) as shown.

$$g(x) = w^t x + b \quad (3.15)$$

Where w^t represent the weight vector which is perpendicular to the hyperplane and represent the orientation of the hyperplane in the d-dimension space feature vector, b is the bias term which represent the position of that hyperplane in the d-dimension space and x is the feature vector.

The equation (3.15) represents the straight line in 2-dimension, represents the equation of plane in 3-dimension and if the dimension is greater than three than it represents the equation of the hyperplane.

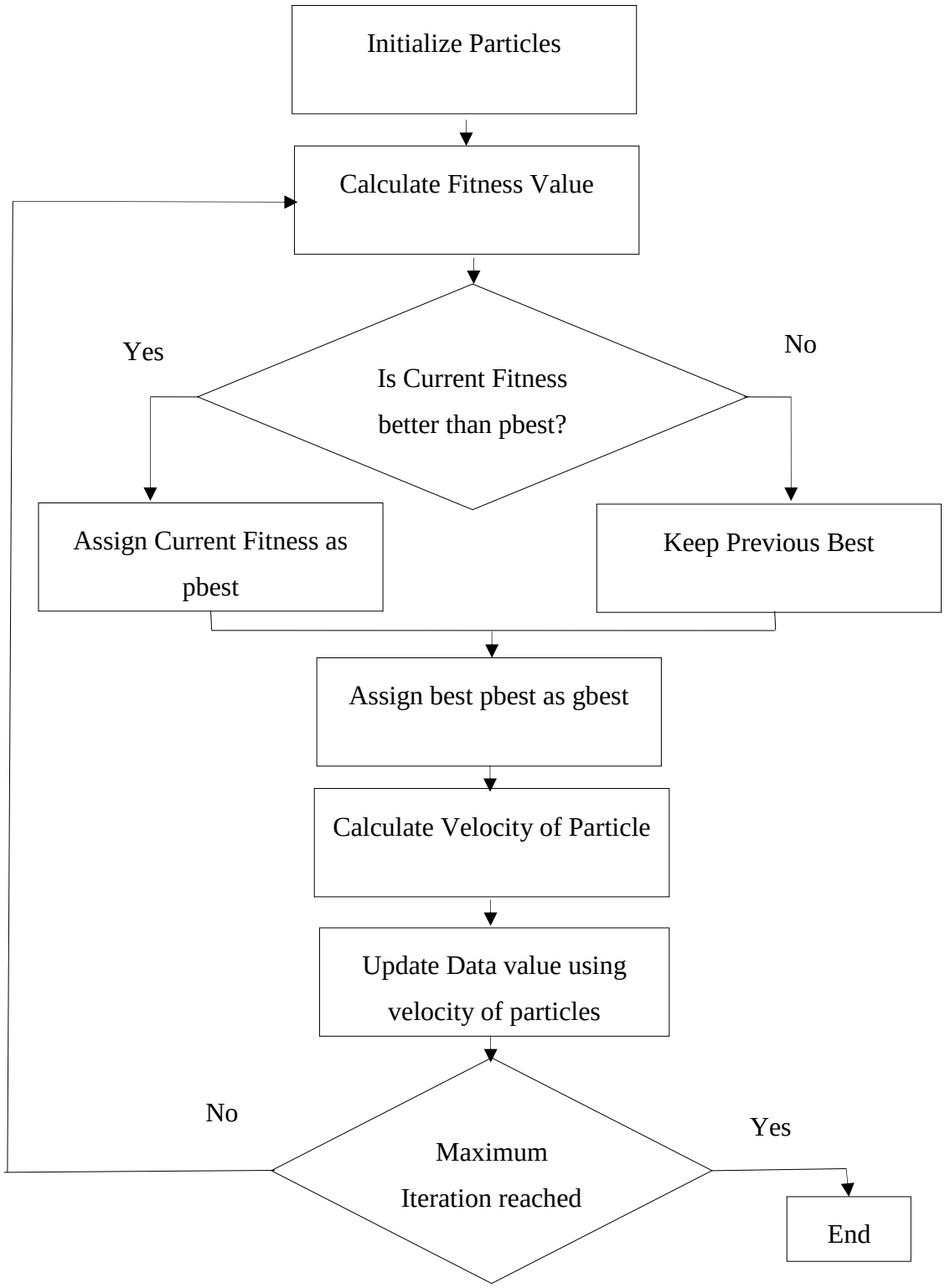

Figure 3.2 Flowchart of Particle Swarm Optimization

Optimal Separating hyperplane means that the nearest data point to the hyperplane should be as far as possible. This can be more easily understood by the figure 1.4 as shown.

Actually to obtain the optimal separating hyperplane we need to maximize the distance of the closest data point to the separating hyperplane. It means that the closest distance from the either class is at the same distance from the hyperplane. The distance of the closest data point to that of the hyperplane is known as margin. The goal of finding the optimal separating hyperplane is essentially to find a classifier such that the margin M is as large as possible.

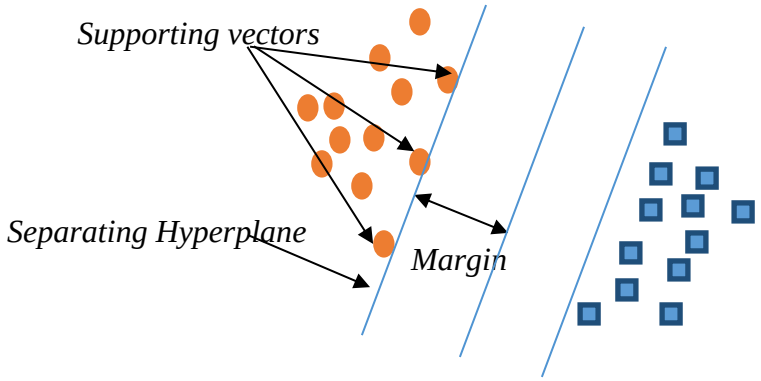

Figure 3.3 Linear SVM

Thus for every feature vector x we has to compute the linear function as defined by eq. (3.15). Now consider the specific vector x_1, we have $g(x_1)$ as shown

$$g(x_1) = w^t x_1 + b \tag{3.16}$$

If the vector x_1 is on the positive side of the hyperplane than we have eq. (3.16) greater than zero, if this x_1 is on the negative side of the hyperplane the eq.(3.16) will be less than zero and if the x_1 lies on the hyperplane than we have eq.(3.14) equals to zero. So it is the hyperplane which divides the d-dimensional space into two half spaces.

In one of the half spaces if we take the feature vector x than $g(x)$ for that particular vector of x will be positive and if we take the vector x on another half space than $g(x)$ for that vector is negative. We have the classification rule that if $w^t x_1 + b > 0$ than it belongs to class c_1 and if the $w^t x_1 + b < 0$ than it belongs to classc_2. Our aim is to find the value of w and b so that the value of $g(x)$ with that particular value of w and b can be used as a classifier. For classification we are using supervised learning, a large number of samples has been given. Some of them belonging to the class c_1 and some are coming from the classc_2.

Using these training samples to train our network and training is done in an iterative fashion. For every training sample belonging to c_1 we started with the initial value of $w\ and\ b$ and for every training sample belonging to class c_1 , $w^t x + b$ should be greater. If it is not than we need to modified the value of $w\ and\ b$ in such a way that the position and the orientation of the hyperplane is so modified that the vector x which is taken from the class c_1 is moved to the positive side of the hyperplane.

The similar procedure is followed for classc_2. SVM basically try to maximize the distance of the separating boundary between two classes by maximising the distance of separating hyperplane of each of the feature vector. For every x_i there exist y_i such that$y_i = \pm 1$. To maximize the margin we need to minimize the equation subject to some constraints as discuss below

$$\frac{1}{2}\|w\|^2 \qquad (3.17)$$

$$y_i\ (\vec{w}.\vec{x}_i + b) \geq 1 \qquad (3.18)$$

The Lagrange multiplier equation which try to optimize is given by equation

$$\frac{1}{2}\|w\|^2 - \ \Sigma_i \alpha_i (y_i\ (\vec{w}.\vec{x}_i + b) - 1) \qquad (3.19)$$

Thus solving the above lagrangian optimization equation will give us the value of w, b and α_i parameters to give us the unique margin. The SVM uses the kernel to divide the data into two groups. This division is due to the kernel trick [].

3.3.4 ARTIFICIAL NEURAL NETWORK (ANN)

ANN (Artificial Neural Network) is the most important and the simple network which has been designed according to the neural structure of brain. Just like our brain consists of large number of neurons which act as a conducting or transporting medium to stimulate the respective body organ. Neural network also works in such a fashion. It basically consist of three layers and these are: Input layer, hidden layer and output layer. On the basis of the output layer weights are assigned to get output accordingly to input layer. Describe below the working of neural network.

1. **Input layer:** This layer is the one through which our input trained data is being fed
2. **Hidden Layer:** All the processing of the input trained data is being performed here
3. **Output Layer:** After the processing of trained data in the hidden layer, the classification results are being taken out from this layer.

Basically what happen is that the input training data set is being fed to the input layer of the neural network. The information from here is being transmitted to the output layer with the help of hidden layers. Hidden layers are basically the group of neurons which are responsible for the information to be transmitted from one layer to another. The transmission of data take place in such a way that the output of one neuron act as the input to the other neuron. The gap between the two neurons that are present in the different layer is known as synaps and it is the place where the weight has been assigned to each link. Since each layer contain the neurons which acts as transmission agent, the following process is being performed at the each neuron. First the input trained data set "x" is multiplied by the scalar weight "w" and the result obtained is being added to bias "b" to form the net input. The net input is being fed to the transfer function "f" which produces the necessary output. Various learning rule has been used in the neural networks but the most common that is being used these days is the Delta rule [21].

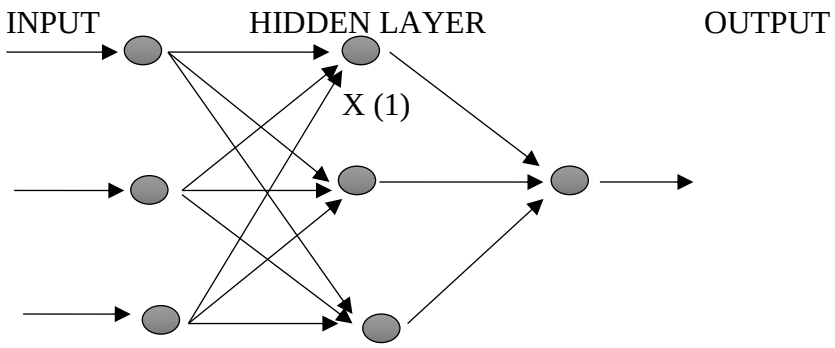

Figure 3.4 Neural network architecture

Figure 1.6 depicts the general architecture of the neural network. Three different layers has been shown, here in the hidden layer weights are being multiplied by the input data that has been transmitted from the input layer and then the data is being transmitted to the output layer with the help of hidden layer.

3.3 FLOWCHART OF WORK

The flowchart of the work is as shown in the figure 3.5 and various steps that follows are:

Step 1: Upload satellite image.

Step2: Apply pre-processing in order to remove noise from the test image and to extract the bands from the satellite image.

Step 3: Develop a PCA code for the feature extraction from the each band.

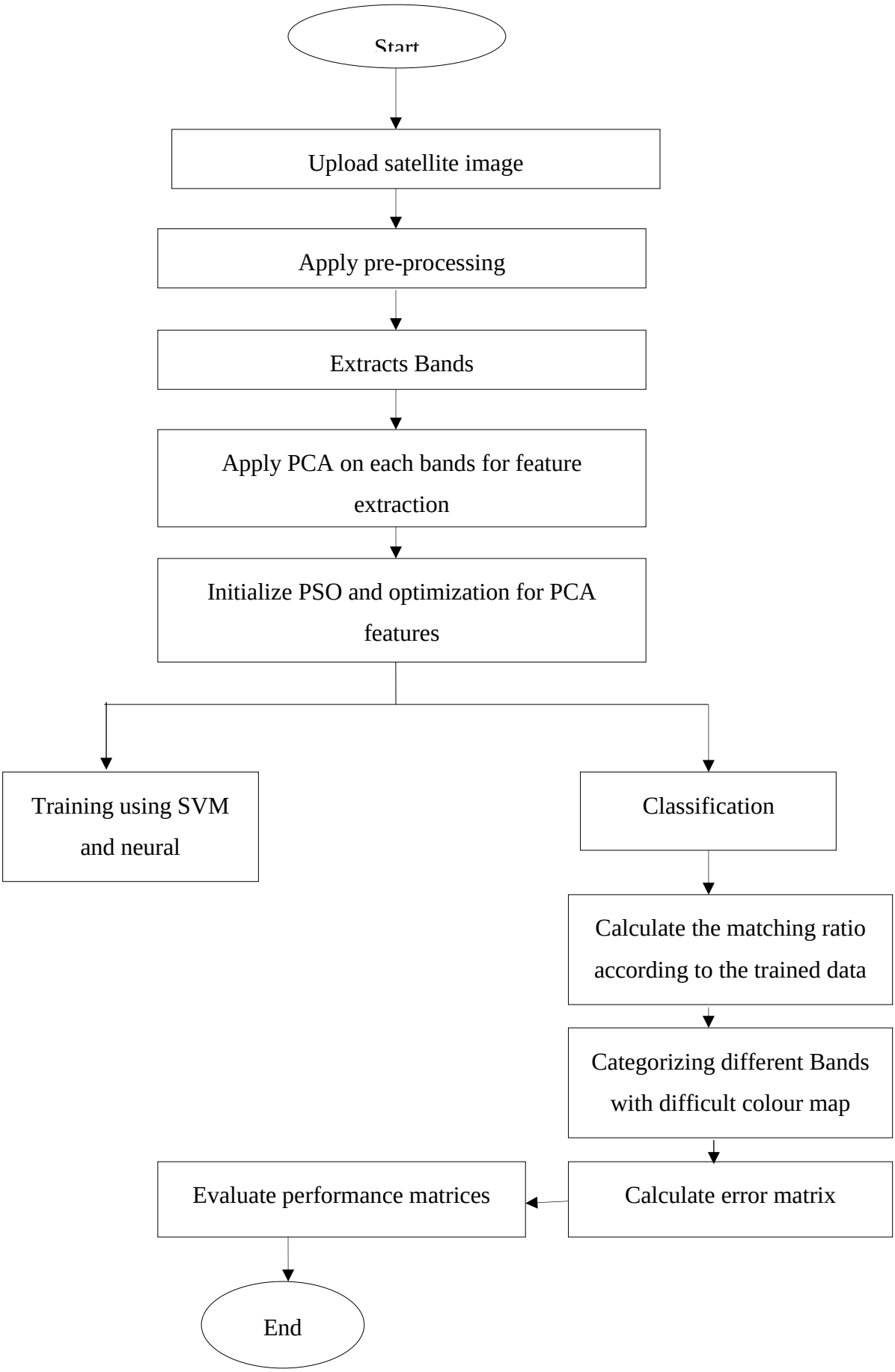

Figure 3.5 Flowchart of work

Step4: Initialize the PSO algorithm to optimize the PCA feature so we generate a novel objective function according to the requirement.

Step 5: Develop a code for the SVM classifier as well as ANN classifier to train and classification of satellite image.

Step6: Store the trained structure of SVM and ANN to classify the bands.

Step 7: Calculate the error matrix of classification according to the trained data.

Step 8: After the classification represent the band using the color map with different color.

Step 9: Calculate the performance metrics for the two different classification techniques.

CHAPTER 4: RESULTS AND ANALYSIS

4.1 INTRODUCTION

This chapter deals with the simulation and results of the two different classification techniques that is the Support Vector Machine and the Artificial Neural Network. First considering hyperspectral data, since this data dimension is very high, one is not able to retrieve the image of it on the screen. In this every pixel contained the detailed spectrum. Larger the amount of information associated per pixels will lead to increasing capability to distinguish objects. Due to the increase dimensionality of the Hyper-Spectral Image, image analysis become more complex. Thus efficient algorithms for automatic processing is required. Only considering one of the band of the hyperspectral image one can plot that band. The next step is to perform the segmentation of that image which is done by using k-means algorithm. This is unsupervised type of algorithm in which no prior information regarding the training samples are performed. K-means basically look for the patterns in the data but is not going to label it. In this pixels are going to be grouped on the basis of reflectance's, and this grouping of the pixels on the basis of reflectance is known as clustering.

After the clustering has been performed, now our aim is to extract the feature from the clustered plot and this feature extraction is performed by the well-known algorithm known as Principal Component Analysis. PCA extract feature on the basis of the Eigen values or the Principal Components which are also known as variance. The Principal components obtained are sorted out in the decreasing order and the one with the highest value results in the most of the information contained in it. After the feature has been extracted the next step is to optimize the feature, which is performed by the well-known algorithm known as Particle Swarm Optimization. This optimization algorithm derived its concept from the nature that has it origin from the flocking of the birds and swarm of fishes. The optimized feature obtained is than passed through the two different classification techniques that is SVM and ANN. The results obtained from these two different classification techniques are compared and the different performance parameters are calculated.

4.2 SIMULATION RESULT

This section includes the simulation results of the two different classification techniques whose input has been optimized using PSO and feature has been extracted using PCA

.

4.2.1 Results obtained by using SVM as a classification technique

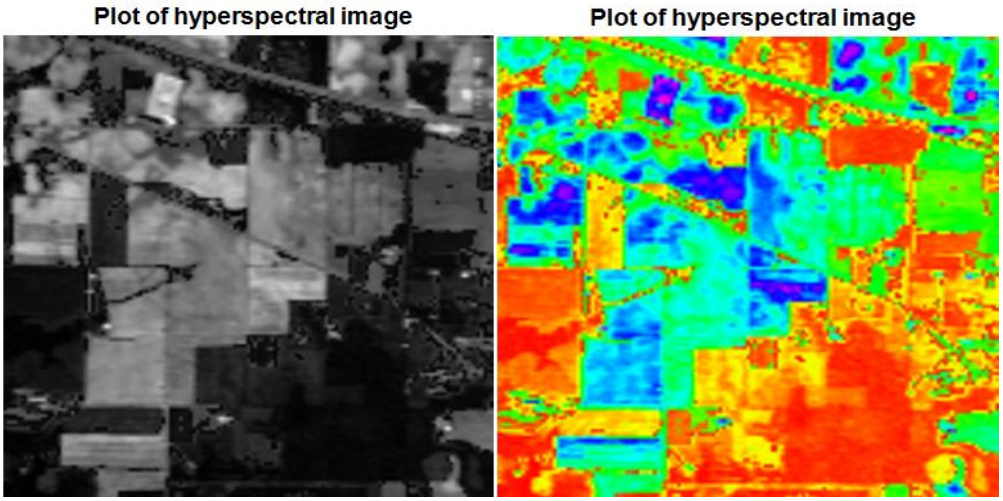

Figure 4.1: Gray Scale and RGB plot of one of the bands of Indian Pines

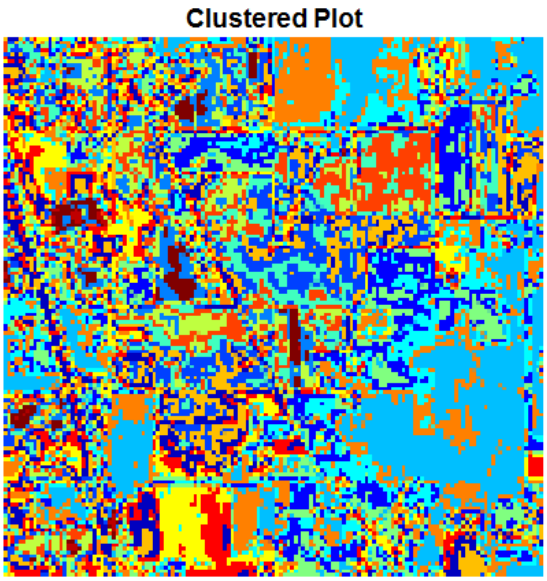

Figure 4.2: Clustered Plot obtained using K-means Algorithm

Clustered image obtained using k-means (which is unsupervised learning algorithm) contains information from a number of spectral channels into a single image. This image will provide useful information as compared to single image. Here, different colours represent various bands. In the clustered image we have applied PCA which is the feature extraction technique. By setting ground truth value, we have extracted the features from the multilayer image.

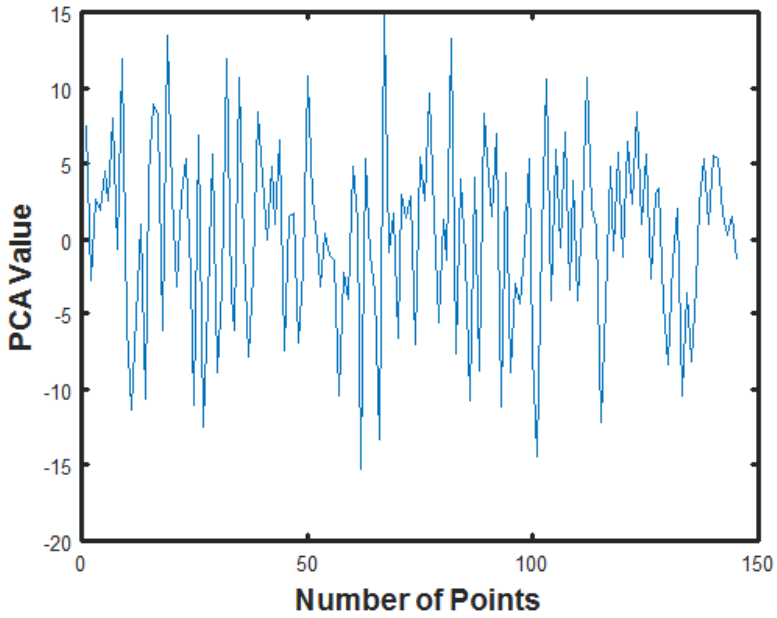

Figure 4.3: graph between PCA value and number of points

The above graph shows that there are 150 number of points or pixels in the PCA image that we have selected and the value of feature extraction. The above features obtained from the PCA are being processed by the very popular optimization technique known as Particle Swarm Optimization. We apply PSO technique on PCA feature set to reduce the unwanted feature from feature set.

The figure 4.4 as shown below depicts that, the red colour circle represent cluster which is created by using K-mean algorithm. The blue colour line represents the optimized feature value which is obtained by PSO technique. K-mean algorithm is used for clustering of different classes about which we do not have any prior information.

So it is unsupervised type of classification algorithm. It basically look for the patterns in the data but is not going to label it. In this pixels are going to be grouped on the basis of reflectance's, and this grouping of the pixels on the basis of reflectance is known as clustering.

The fig 4.5 as shown below represents the classification of spectral image according to their image and we represent different classes in different colour. The plot is obtained using support vector machine as a classifier to classify the spectral image according to PCA feature with PSO as an optimization technique.

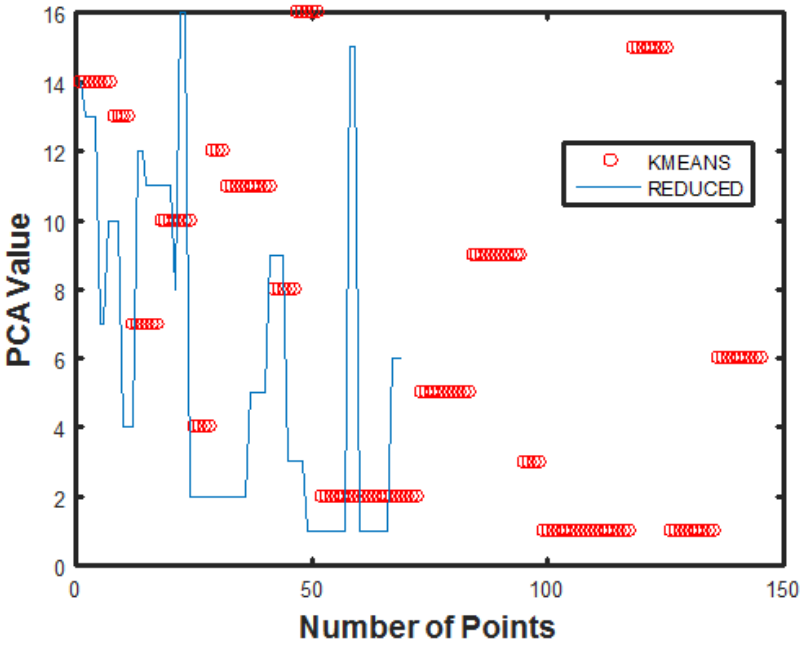

Figure 4.4: k-means verses Reduced

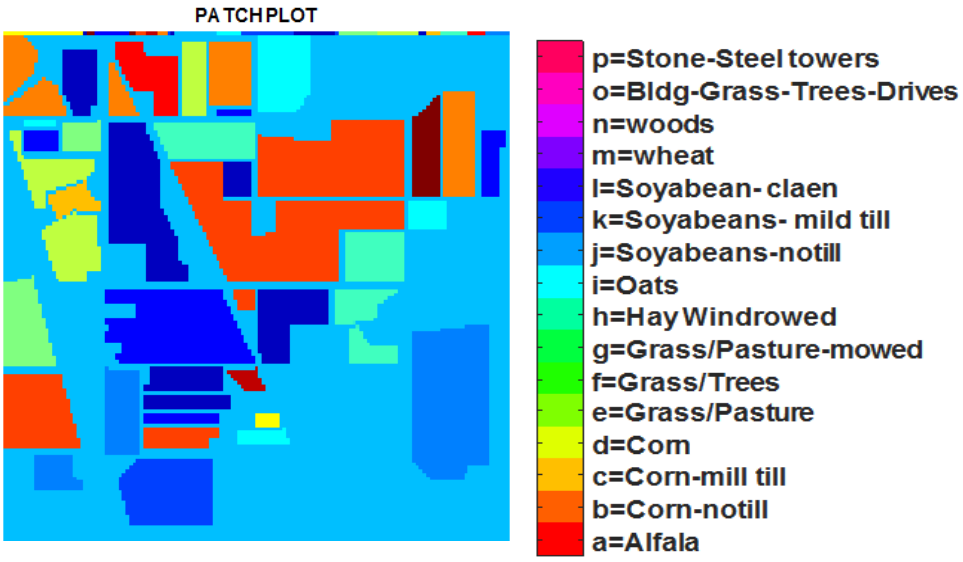

Figure 4.5: Patch Plot obtained after PSO-SVM Classification

4.2.2 Results obtained using ANN as a classification technique

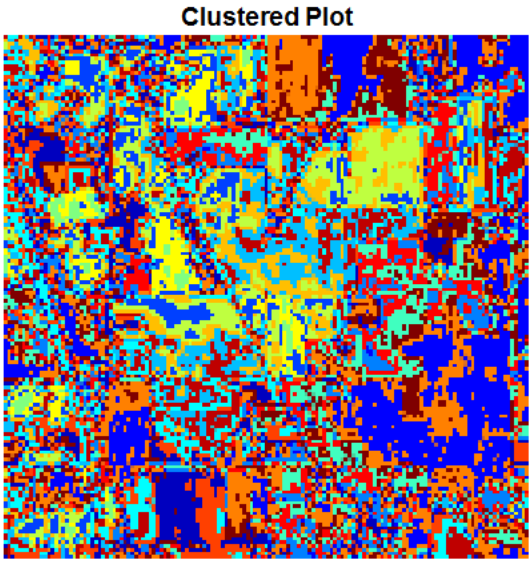

Figure 4.6 Clustered image obtained using k-means algorithm for neural network

In the above figure, an image with different colours has been displayed. Different colours represent different number of bands. This image is used for neural network classification. The image obtained is passed through the feature extraction algorithm known as Principal component Analysis. After the features has been obtained from PCA it is then passed through the optimization algorithm PSO. The distorted part in PCA image has been removed by using PSO techniques which helps us to give better results.

The figure 4.7 as shown below depicted that, the red colour circle represent cluster which is created by using K-mean algorithm. The blue colour line represents the optimized feature value which is obtained by PSO techniques.

The above image represent the patch plot which has been obtained from the ANN classification. The clusters has been formed using the K- Mean algorithm. After the clusters has been performed our aim is to extract the feature vector from the clusters and the features are being selected by the well-known algorithm known as the Principal Component Analysis based on the Eigen values.

After the feature has been selected, they are being apply to the well- known Optimization technique known as the Particle Swarm Optimization. This Optimization technique is nature based technique that has it origin from the flocking of the birds. From this we have obtained the optimized feature and is then passed through the neural network training tool to obtain the desired result.

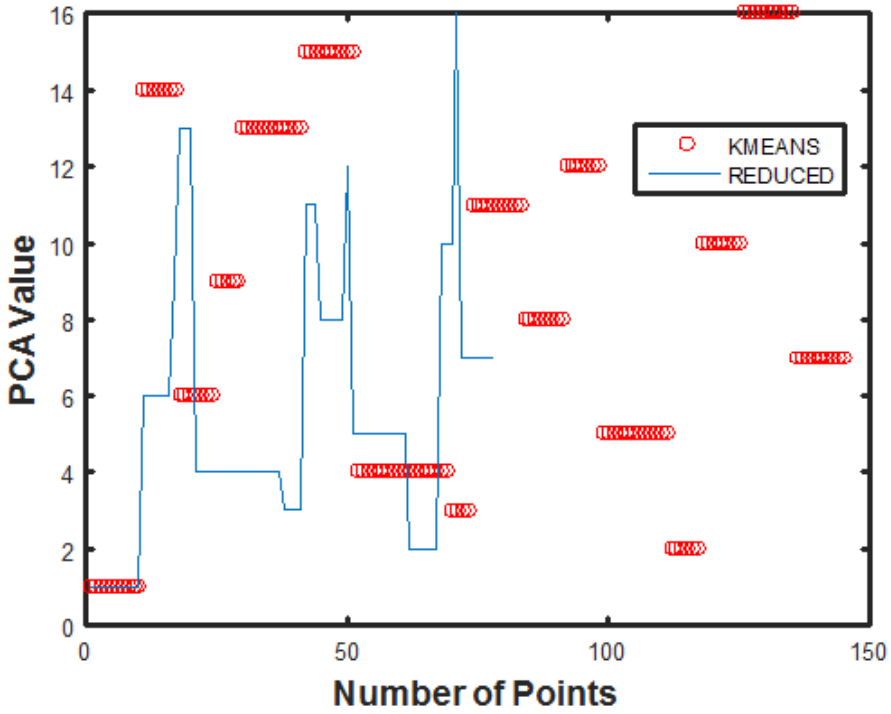

Figure 4.7: K-mean verses reduced graph for ANN

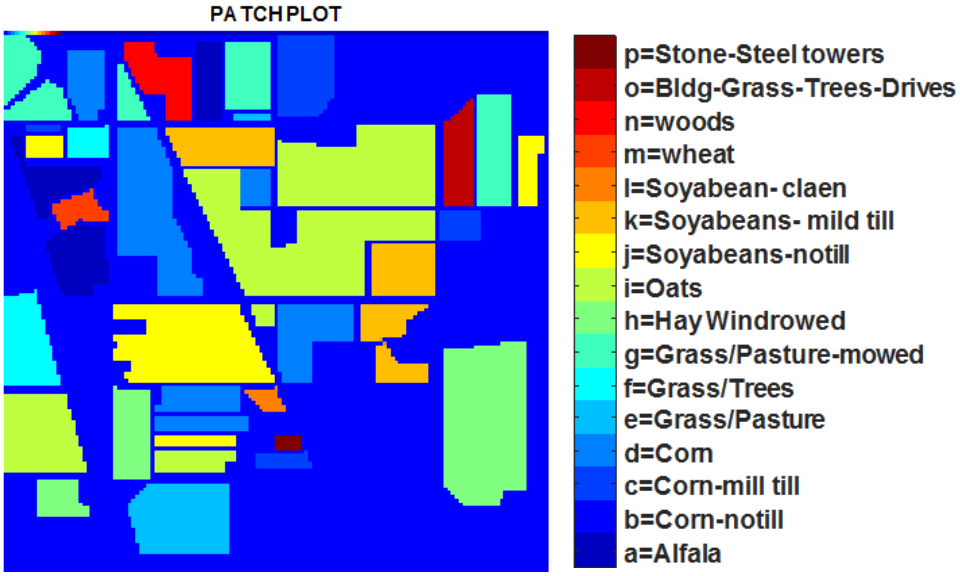

Figure 4.8: Patch plot of the image obtained from neural network for INDIAN PINES

The above image represent the patch plot which has been obtained from the ANN classification. The clusters has been formed using the K- Mean algorithm which is unsupervised learning algorithm and the pixels are being clustered on the basis of reflectance. After the clusters has been performed our aim is to extract the feature vector from the clusters and the features are being selected by the well-known algorithm known as the Principal Component Analysis which extract the feature according to the Eigen values.

After the feature has been selected, they are being apply to the well- known Optimization technique known as the Particle Swarm Optimization. This Optimization technique is nature based technique that has it origin from the flocking of the birds.

From this we have obtained the optimized feature and is then passed through the neural network training tool to obtain the desired result. Different colours in the patch plot represent the various classes which has been classified accordingly the neural network training tool.

Table 4.1: Comparison of various performance parameters obtained

Classes	Classification Accuracy using PCA-SVM	Classification Accuracy PSO-SVM	Classification Accuracy PSO-ANN
Alfalfa	99.2449	100	85.9147
Corn-no till	96.0550	98.6799	68.3082
Corn - mill till	91.1237	92.2905	61.6247
Corn	80.5793	80.6025	93.7399
Grass/Pasture	78.3294	62.4518	43.6590
Grass/Trees	71.5065	69.4587	96.4787
Grass/Pasture-Mowed	52.5327	68.0895	93.7399
Hay windrowed	66.6792	91.1905	99.6087
Oats	75.9562	94.97198	65.1559
Soya beans - no till	83.2183	98.2033	98.4350
Soya beans- mild till	87.0400	99.0833	96.4787
Soya beans-claen	93.5239	99.4133	72.3522
Wheat	96.5327	99.4133	83.9281
Woods	99.5338	99.6700	98.4350
Bldg.-Grass-Tree-Drives	99.5338	99.9175	55.9162
Stone-Steel Towers	99.9037	88.3371	99.1517
Average Accuracy	85.7058	88.3371	82.0582
Average MSE	14.2942	11.6629	16.199
PSNR	36.6132	37.4967	36.0699

CHAPTER 5: CONCLUSION AND FUTURE SCOPE

5.1 CONCLUSION

A hyper spectral image comprises of at least hundred adjacent ghostly groups framing a three-dimensional (two spatial measurements and one unearthly measurement) picture shape. In this every pixel contained the detailed spectrum. Larger the amount of information associated per pixels will lead to increasing capability to distinguish objects. Due to the increase dimensionality of the Hyper-Spectral Image, image analysis become more complex. Thus efficient algorithms for automatic processing is required. Till now many algorithms has been designed regarding the classification problem but none of them has been able to assemble the data set belonging to the real world with great competency and without error. In order to review the quality of satellite images, there is need to design or model the algorithm based on the statistical ground. In this research, we have presented a hybrid technique of PSO-SVM and PSO-ANN network for Optimization to classify multispectral satellite image.

Firstly perform the segments [36] of the satellite image which can be done by using K-mean algorithm. After the clustering has been performed, now our aim is to extract the feature from the clustered plot and this feature extraction is performed by the well-known algorithm known as Principal Component Analysis. PCA [37] extract feature on the basis of the Eigen values or the Principal Components which are also known as variance. The Principal components obtained are sorted out in the decreasing order and the one with the highest value results in the most of the information contained in it. After the feature has been extracted the next step is to optimize the feature, which is performed by the well-known algorithm known as Particle Swarm Optimization [38]. This optimization algorithm derived its concept from the nature that has it origin from the flocking of the birds and swarm of fishes. The optimized feature obtained is than passed through the two different classification techniques that is SVM [39] and ANN [40]. The results obtained from these two different classification techniques are compared and the different performance parameters are calculated.

5.2 FUTURE SCOPE

In today's world detail spectrum analysis along with the spatial information is required for classification. Hyperspectral images contained detailed spectral information but lacks the spatial information. So in future, classification can be performed by combining different techniques which increases the spatial resolution content in hyperspectral image resulting in the enhancement of the classification accuracy.

REFERENCES

[1] Jensen JR. *Remote sensing of the environment: An earth resource perspective.* India: Pearson Education India, 2009.

[2] Rees WG. *Physical principles of remote sensing.* United Kingdom: Cambridge University Press, 2013.

[3] Han, J. and Kamber M. *Data Mining Concepts and Techniques*. San Francisco: Morgan Kaufmann Publishers, 2006.

[4] Lee Rok Sang (2010). A Coarse-to-Fine Approach for Remote-Sensing Image Registration Based on a Local Method, *International Journal on Smart Sensing and Intelligent Systems*, 3(4), 690-702.

[5] Lin M *et al.* (2007). Blended tools for remote sensing education, *Geoscience and Remote Sensing Symposium, (IGARSS),* [Barcelona, Spain: 2007], pp. 2217-2219.

[6] Engelbrecht, P Andrei's. *Fundamentals of Computational Swarm Intelligence*. United Kingdom: Wiley and Sons, 2006.

[7] Jia X, Richards JA (1999). Segmented principal components transformation for efficient hyperspectral remote-sensing image display and classification, *IEEE Trans. Geoscience. Rem. Sens.*, 37(1), 538-542.

[8] Kumar S, Ghosh J, Crawford MM (2001). Best basis feature extraction algorithms for classification of hyperspectral data, *IEEE Trans. Geosci. Rem. Sens*, 29(7), 1368-79.

[9] Ifarraguerri A, Chang C (2000). Unsupervised hyperspectral image analysis with projection pursuit, *IEEE Trans. Geosci. Remote Sens*., 38(6), 2529–2538.

[10] Jin NH *et al.* (2005). Pca-based feature extraction using class information, *IEEE International Conference on Systems, Man and Cybernetics*, [Waikoloa, HI, USA: 2005], pp.341 – 345.

[11] Liangpei Z, Yanfei Z (2007), Dimensionality Reduction Based on Clonal Selection for Hyperspectral Imagery, *IEEE Trans. Geosci. Remote Sens*, 45 (12), 4172-4186.

[12] Lee CD, Grebe LA (1993). Analysing high-dimensional multispectral data", *IEEE Trans. Geosci. Remote Sens*., 31(4), 792–800.

[13] Jiao L *et al.* (2009). Feature Extraction Combining PCA and Immune Clonal Selection for Hyperspectral Remote Sensing Image Classification, *International Conference on Artificial Intelligence and Computational Intelligence*, [Shanghai, China: 2009], pp. 472-476.

[14] Shalem RH *et al.* (2014). Compression of Hyperspectral Images Containing a Subpixel Target, *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*, 7(6), 2246-2255.

[15] Li J (2016), A novel remote sensing image classification algorithm based on PCA and hidden Markov random field theory, *Online International Conference on Green Engineering and Technologies (IC-GET)*, [Coimbatore, India: 2016], pp. 1-5.

[16] Yang H et al. (2007). Unsupervised band selection for hyperspectral image analysis, *IEEE International Geoscience and Remote Sensing Symposium*, [*Barcelona*, Spain: 2007], pp. 282-285.

[17] Wang S *et al.* (2006). Constrained Band Selection for Hyperspectral Imagery, *IEEE transactions on geoscience and remote sensing*, 44(6)000, 752-763.

[18] Du Q *et al.* (1999). A Joint Band-Prioritization And Band Decorrelation Approach To Band Selection for Hyperspectral Image Classification, *IEEE transactions on geoscience and remote sensing*, 37(6), 52-63.

[19] Pla F *et al.* (2007). Clustering-based Hyperspectral Band Using Information Measures, *IEEE transactions on geoscience and remote sensing*, 45(12), 6752-6766

[20] Persello C et al. (2008). A Novel Approach to the Selection of Robust and Invariant Features for Classification of Hyperspectral Images, *IEEE International Geoscience and Remote Sensing Symposium*, [Boston, MA: 2008], pp. 66-69.

[21] Dey S et al. (2015). A morphological segmentation and curve-let features extraction on text region classification using SVM, *IEEE International Conference on Computational Intelligence and Computing Research (ICCIC)*, [Madurai, India:2015], pp. 1-6.

[22] Liu Q. (2017). Accurate Object Localization in Remote Sensing Images Based on Convolutional Neural Networks, *IEEE Transactions on Geoscience and Remote Sensing*, 55(5), 2486-2498.

[23] Agsornintara A. *et al.* (2017). Multiclass support vector machine for classification spatial data from satellite image, *International Conference on Knowledge and Smart Technology (KST)*, [9th: Chonburi, Thailand: 2017], pp. 111-115.

[24] Rajakani M. *et al.* (2016). Wavelet based semi supervised learning for satellite image classification using ensemble projection, *International Conference on Intelligent Systems and Control (ISCO)*,[10th : Coimbatore, India: 2016], pp. 1-6.

[25] Dixit M. *et al.* (2016). Supervised classification of satellite images, *Conference on Advances in Signal Processing (CASP)*, [Pune, India: 2016], pp. 93-98.

[26] Farah IR e*t al.* (2016). Improvement of satellite image classification: Approach based on Hadoop/MapReduce, *International Conference on Advanced Technologies for Signal and Image Processing (ATSIP)*, [2nd: Monastir, Tunisia: 2016], pp. 31-34.

[27] Goel L. *et al.* (2012) "Dynamic Model of Blended Biogeography Based Optimization for Land Cover Feature Extraction, *International Conference on contemporary computing*, [5th: Noida, India: 2012], pp. 8-19.

[28] Singh P, Kaur N, Kaur L (2011). Satellite Image Classification by Hybridization of FPAB Algorithm and Bacterial Chemotaxis, *International Journal of Computer Technology and Electronics Engineering (IJCTEE)*, 1(3), 2249-6343.

[29] Kundra H *et al.* (2009). Biogeography based Satellite Image Classification, *International Journal of Computer Science and Information Security (IJCSIS)*, 6(2), 269-274.

[30] Gupta, Daya, Bidisha Das, and V. K. Panchal, "A methodical study for the extraction of landscape traits using membrane computing technique," *WORLDCOMP, Las Vegas, USA* (2011).

[31] Kundra H *et al.* (2010). A hybrid FPAB/BBO algorithm for satellite image classification, *International Journal of Computer Applications,* 6(5), 31-36

[32] Singh V *et al.* (2013).Land Features Extraction from Image by the Probability Concept of Honey Bees: A Remote Sensing Perspective, *International Journal of system, Algorithms and Applications,* 2(13), 1-4.

[33] Kaur J, Kaur K. (2012). Remote Image Classification Using Particle Swarm Optimization, *International Journal of Emerging Technology and Advanced Engineering*, 2(7), 470-473.

[34] Singh S.P. *et al.* (2015). Hybrid clustering algorithm and Neural Network classifier for satellite image classification, *International Conference on Industrial Instrumentation and Control (ICIC)*, [1st : Pune, India : 2015], pp. 1378-1383.

[35] Jiang Xiaoqian *et al.* (2008). Privacy-preserving SVM classification, *Knowledge and Information Systems, springer*, 14(2), 161-178.

[36] Tian S *et al.* (2015). Building surface texture segmentation in urban remote sensing image using improved ORTSEG algorithm, *IEEE International Conferencing Geoscience and Remote Sensing Symposium (IGARSS)*, [Milan, Italy: 2015], pp. 4348-4351.

[37] Vlasenko B et al. (2015). Exploring dataset similarities using PCA-based feature selection, *International Conference on Affective Computing and Intelligent Interaction (ACII)*, [Xi'an, China: 2015], pp. 387-393.

[38] Mostaghim M *et al.* (2017).PSO-based Search mechanism in dynamic environments: Swarms in Vector Fields, *IEEE Congress on Evolutionary Computation (CEC)*, [Don Ostia, San Sebastian, Spain: 2017], pp. 1263-1270.

[39] Raghuwanshi MM et al. (2016). Classification of remote sensing image using SVM kernels, *World Conference on Futuristic Trends in Research and Innovation for Social Welfare (Start-up Conclave)*,[Coimbatore, India: 2016], pp. 1-5.

[40] Said A *et al.* (2016). Neural Network Classification for Iris Recognition Using Both Particle Swarm Optimization and Gravitational Search Algorithm, *World Symposium on Computer Applications & Research (WSCAR)*, [Cairo, Egypt: 2016], pp. 12-17.

LIST OF PUBLICATIONS

[1] Mahajan A and Mishra A. Hyper Spectral satellite image classification using SVM and Artificial neural network. Turkish Journal of Electrical Engineering and Computer Sciences (Communicated)