Jyothi Hills, Panjal Road,
Vettikattiri PO, Cheruthuruthy , Thrissur,
Kerala 679531

Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SEMINAR REPORT

SATELLITE IMAGE CLASSIFICATION USING SUPPORT VECTOR MACHINE

Submitted by

SRUTHI ELSA SHAJI
JEC17CS100

Supervised by

Ms.NINU FRANCIS
Asst. Prof., Dept. of CSE

---

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY (B.Tech)

in

COMPUTER SCIENCE & ENGINEERING
of

A P J ABDUL KALAM TECHNOLOGICAL UNIVERSITY

CREATING TECHNOLOGY LEADERS OF TOMORROW

DECEMBER 2020

Jyothi Hills, Panjal Road,
Vettikattiri PO, Cheruthuruthy , Thrissur,
Kerala 679531

Jyothi Engineering College
NAAC Accredited College with NBA Accredited Programmes*
Approved by AICTE & affiliated to APJ Abdul Kalam Technological University
A CENTRE OF EXCELLENCE IN SCIENCE & TECHNOLOGY BY THE CATHOLIC ARCHDIOCESE OF TRICHUR
JYOTHI HILLS, VETTIKATTIRI P.O, CHERUTHURUTHY, THRISSUR. PIN-679531 PH : +91- 4884-259000, 274423 FAX : 04884-274777
NBA accredited B.Tech Programmes in Computer Science & Engineering, Electronics & Communication Engineering, Electrical & Electronics Engineering and Mechanical Engineering valid for the academic years 2016-2022. NBA accredited B.Tech Programme in Civil Engineering valid for the academic years 2019-2022.

DEPARTMENT OF COMPUTER SCIENCE AND ENGINEERING

SEMINAR REPORT

SATELLITE IMAGE CLASSIFICATION USING SUPPORT VECTOR MACHINE

Submitted by

SRUTHI ELSA SHAJI
JEC17CS100

Supervised by

Ms.NINU FRANCIS
Asst.Prof., Dept. of CSE

---

in partial fulfillment for the award of the degree

of

BACHELOR OF TECHNOLOGY (B.Tech)

in

COMPUTER SCIENCE & ENGINEERING
of

A P J ABDUL KALAM TECHNOLOGICAL UNIVERSITY

DECEMBER 2020

Department of Computer Science and Engineering

JYOTHI ENGINEERING COLLEGE, CHERUTHURUTHY

THRISSUR 679 531

DECEMBER 2020

BONAFIDE CERTIFICATE

This is to certify that the seminar report entitled **SATELLITE IMAGE CLASSIFICATION USING SUPPORT VECTOR MACHINE** submitted by **SRUTHI ELSA SHAJI (JEC17CS100)** in partial fulfillment of the requirements for the award of **Bachelor of Technology** degree in **Computer Science and Engineering** of **A P J Abdul Kalam Technological University** is the bonafide work carried out by her under our supervision and guidance.

Ms. Ninu Francis	**Dr. Swapna B Sasi**	**Dr. Vinith R**
Seminar Guide	Seminar Coordinator	Head of The Dept
Assistant Professor	Assistant Professor	Professor
Dept. of CSE	Dept. of CSE	Dept. of CSE

DEPARTMENT OF

COMPUTER SCIENCE & ENGINEERING

COLLEGE VISION

Creating eminent and ethical leaders through quality professional education with emphasis on holistic excellence.

COLLEGE MISSION

- To emerge as an institution par excellence of global standards by imparting quality engineering and other professional programmes with state-of- the-art facilities.
- To equip the students with appropriate skills for a meaningful career in the global scenario.
- To inculcate ethical values among students and ignite their passion for holistic excellence through social initiatives.
- To participate in the development of society through technology incubation, entrepreneurship and industry interaction.

DEPARTMENT OF

COMPUTER SCIENCE & ENGINEERING

DEPARTMENT VISION

Creating eminent and ethical leaders in the domain of computational sciences through quality professional education with a focus on holistic learning and excellence.

DEPARTMENT MISSION

- To create technically competent and ethically conscious graduates in the field of Computer Science & Engineering by encouraging holistic learning and excellence.
- To prepare students for careers in Industry, Academia and the Government.
- To instill Entrepreneurial Orientation and research motivation among the students of the department.
- To emerge as a leader in education in the region by encouraging teaching, learning, industry and societal connect

PROGRAMME OUTCOMES (POs)

1. **Engineering Knowledge**: Apply the knowledge of mathematics, science, engineering fundamentals, and an engineering specialization to the solution of complex engineering problems.
2. **Problem Analysis**: Identify, formulate, review research literature, and analyze complex engineering problems reaching substantiated conclusions using first principles of mathematics, natural sciences, and engineering sciences.
3. **Design/Development of Solutions**: Design solutions for complex engineering problems and design system components or processes that meet the specified needs with appropriate consideration for the public health and safety, and the cultural, societal, and environmental considerations.
4. **Conduct Investigations of Complex Problems**: Use research-based knowledge and research methods including design of experiments, analysis and interpretation of data, and synthesis of the information to provide valid conclusions.
5. **Modern Tool Usage**: Create, select, and apply appropriate techniques, resources, and modern engineering and IT tools including prediction and modeling to complex engineering activities with an understanding of the limitations.
6. **The Engineer and Society**: Apply reasoning informed by the contextual knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to the professional engineering practice.
7. **Environment and Sustainability**: Understand the impact of the professional engineering solutions in societal and environmental contexts, and demonstrate the knowledge of, and need for sustainable development.
8. **Ethics**: Apply ethical principles and commit to professional ethics and responsibilities and norms of the engineering practice.
9. **Individual and Team Work**: Function effectively as an individual, and as a member or leader in diverse teams, and in multidisciplinary settings.
10. **Communication**: Communicate effectively on complex engineering activities with the engineering community and with society at large, such as, being able to comprehend and write effective reports and design documentation, make effective presentations, and give and receive clear instructions.
11. **Project Management and Finance**: Demonstrate knowledge and understanding of the engineering and management principles and apply these to one's own work, as a member and leader in a team, to manage projects and in multidisciplinary environments.
12. **Life-Long Learning**: Recognize the need for, and have the preparation and ability to engage in independent and life-long learning in the broadest context of technological change.

PROGRAMME EDUCATIONAL OBJECTIVES (PEOs)

1. The graduates shall have sound knowledge of Mathematics, Science, Engineering and Management to be able to offer practical software and hardware solutions for the problems of industry and society at large.
2. The grad uates shall be able to establish themselves as practising professionals, researchers or Entrepreneurs in computer science or allied areas and shall also be able to pursue higher education in reputed institutes.
3. The graduates shall be able to communicate effectively and work in multidisciplinary teams with team spirit demonstrating value driven and ethical leadership.

Programme Specific Outcomes (PSOs)

1. An ability to apply knowledge of data structures and algorithms appropriate to computational problems.
2. An ability to apply knowledge of operating systems, programming languages, data management, or networking principles to computational assignments.
3. An ability to apply design, development, maintenance or evaluation of software engineering principles in the construction of computer and software systems of varying complexity and quality.
4. An ability to understand concepts involved in modeling and design of computer science applications in a way that demonstrates comprehension of the fundamentals and tradeoffs involved in design choices.

Course Outcomes (COs)

C418.1 **Presentation Skills in terms of Content** : Students will be able to show competence in identifying relevant information, defining and explaining topics under discussion. They will demonstrate depth of understanding, use primary and secondary sources; they will demonstrate the working, complexity, insight, cogency, independent thought, relevance, and persuasiveness. They will be able to evaluate information and use and apply relevant theories.

C418.2 **Presentation Skills in terms of Organization** : Students will be able to show competence in working with a methodology, structuring their oral work, and synthesizing information. They will make a detailed study on the previous works related to their topic and will present the observations.

C418.3 **Presentation Skills in terms of Delivery** : Students will use appropriate registers and vocabulary, and will demonstrate command of voice modulation, voice projection, and pacing. They will be able to make use of visual, audio and audio-visual material to support their presentation, and will be able to speak cogently with or without notes.

C418.4 **Discussion Skills** : Students will be able to judge when to speak and how much to say, speak clearly and audibly in a manner appropriate to the subject, ask appropriate questions, use evidence to support claims, respond to a range of questions, take part in meaningful discussion to reach a shared understanding, speak with or without notes, show depth of understanding.

C418.5 **Listening Skills** : Students will demonstrate that they have paid close attention to what others say and can respond constructively. Through listening attentively, they will be able to build on discussion fruitfully, supporting and connecting with other discussants.

C418.6 **Argumentative Skills and Critical Thinking** : Students will develop persuasive speech, present information in a compelling, well-structured, and logical sequence, respond respectfully to opposing ideas, show depth of knowledge of complex subjects, and develop their ability to synthesize, evaluate and reflect on information.

Course Outcome							
Programme Outcomes		C418.1	C418.2	C418.3	C418.4	C418.5	C418.6
	1	3	3	3	3	3	3
	2	3	3	3	3	3	3
	3	3	3	3	3	3	3
	4	3	3	3	3	3	3
	5	3	3	3	3	3	3
	6	3	3	3	3	3	3
	7	3	3	3	3	3	3
	8	3	3	3	3	3	3
	9	3	3	3	3	3	3
	10	3	3	3	3	3	3
	11	3	3	3	3	3	3
	12	3	3	3	3	3	3

PO - CO Mapping

PEO - CO Mapping

		Course Outcome					
		C418.1	C418.2	C418.3	C418.4	C418.5	C418.6
Programme Educational Objective	1	3	3	1	1	-	2
	2	3	3	3	3	1	3
	3	1	2	3	3	1	3

PSO - CO Mapping

		Course Outcome					
		C418.1	C418.2	C418.3	C418.4	C418.5	C418.6
Programme Specific Outcomes	1	3	3	3	3	3	3
	2	3	3	3	3	3	3
	3	3	3	3	3	3	3
	4	3	3	3	3	3	3

Seminar Outcome

1. Studied about the concept of Machine learning algorithms.
2. Analyzed and studied the general architecture of support vector machine.
3. Studied about different kernel functions.
4. Studied about different applications of image classification.
5. Analyzed and compared with other classification methods.
6. Analyzed the methodology of satellite image classification using svm.

Seminar Outcome - CO Mapping

	Course Outcome						
		C418.1	C418.2	C418.3	C418.4	C418.5	C418.6
Seminar Outcome	1	3	3	3	1	3	3
	2	3	3	1	1	3	3
	3	3	3	3	1	3	1
	4	3	3	3	3	1	1
	5	3	1	3	3	1	1

ACKNOWLEDGEMENT

I take this opportunity to express my heartfelt gratitude to all respected personalities who had guided, inspired and helped me in the successful completion of this seminar. First and foremost, I express my thanks to **The Lord Almighty** for guiding me in this endeavour and making it a success.

I take immense pleasure in thanking the **Management** of Jyothi Engineering College and **Dr. Sunny Joseph Kalayathankal**, principal, Jyothi Engineering College for having permitted me to carry out this seminar. Our sincere thanks to **Dr.Vinith R**, Head of the Department of Computer Science and Engineering for permitting me to make use of the facilities available in the department to carry out the seminar successfully.

I express my sincere gratitude to **Mr. Shaiju paul& Dr. Swapna B Sasi**, seminar coordinators for their invaluable supervision and timely suggestions. I am very happy to express my deepest gratitude to my mentor **Ms.Ninu Francis**, Assistant Professor,Department of Computer Science and Engineering, Jyothi Engineering College for her able guidance and continuous encouragement.

Last but not least, I extend my gratefulness to all teaching and non-teaching staff who directly or indirectly involved in the successful completion of this seminar work and to all friends who have patiently extended all sorts of help for accomplishing this undertaking.

ABSTRACT

SVM plays a vital role in image classification.Support vector machines are supervised learning models with associated learning algorithms that analyze data used for classification and regression analysis.Here is an approach for the classification of Landsat (MSS) satellite images to identify the areas of land use. The image is pre-processed and classified using SVM with the RBF Kernel as it is an efficient supervised classification technique. In this research, pixel – base classification method is performed according to the value of spectral pixels with Multi-Spectral Scanner satellite image and used data corresponds to a 3x3 square neighborhoods. The work consists of two main stages. At the first stage, the optimal parameter, sigma value of RBF kernel, for SVM is studied. At the second stage, the obtained classification result is compared with other classification methods. In this study, sigma value, a parameter of RBF kernel, is varied between 1.0 and 2.0.The sigma value at 1.7 lead to the best classification result which has over 90% accuracy. The result from this SVM method has a higher accuracy compared to other methods.

CONTENTS

List of Figures

List of Abbreviations

SVM	: *SUPPORT VECTOR MACHINE*
MSS	: *MULTISPECTRAL SCANNER SYSTEM*
RBF	: *RADIAL BASIS FUNCTION*
MLP	: *MULTILAYER PERCEPTRON*
PCA	: *PRINCIPAL COMPONENT ANALYSIS*

CHAPTER 1
INTRODUCTION

1.1 Overview

Image classification is one of the interesting and challenging task concerned in remote sensing. Now a days, satellite imagery is widely used in remote sensing and other research areas for various applications. The wide range of applications of satellite imagery includes, Oil and Gas exploration,Mining, Geographic Information Systems, Energy and Infrastructure, Engineering and Construction,Disaster Response, Media and Entertainment,Defense and Intelligence, and exploration of various natural resources. The classification of land use information from remotely sensed data produces a map like image as the result of classification.Remote sensing surveys provide a rapid means of data collection of satellite images and an appropriate classification scheme is required to perform the classification of land use information from these images for various observations and interpretations.Various supervised learning algorithms are preferred for such classification purposes. In this work, we propose a new approach for the classification of satellite images to identify the areas of land use.This is mainly classified using support vector machine with the radial basis function kernel as it is an supervised classification technique.

Satellite image classification is a process of grouping pixels of the image into number of different defined classes [1].The classification is the process for extracting the information contained in different bands of the satellite sensor and the information is extracted in terms of digital numbers which is then converted to a spatial class. The method of classification can be supervised or unsupervised. Unsupervised classification is referred as a clustering for unclear grouping when no sample sets are available. Supervised classification requires input from analyst and identifiers different classes based on the sample training sets.Pixel – based classification method is widely used for the purpose of detecting land use and land cover with satellite image. The research work classification of satellite images into Landsat Multi-Spectral Scanner (MSS) to identify the areas of land use. The image is pre-processed and classified using Support Vector Machine with the Gaussian Radial Basis Function (RBF) Kernel. As it is an efficient supervised classification technique.

In this research, pixel – base classification method is performed according to the value of spectral pixels with Multi-Spectral Scanner satellite image.The main purpose of this work is to automatically classify the agriculture land use.

1.2 Objective

The main objective of this seminar is to introduce an satellite image classification using SVM.The purpose of this work is to automatically classify the agriculture land use.The research has classification spatial data from a satellite image.The objective of the study is to use SVM technique for classifying multispectral satellite image dataset and compare the overall accuracy with the conventional image classification method.

1.3 Organization Of The Report

The report is organised as follow:

- **Chapter 1:Introduction** Gives an introduction to supervised technique for classification spatial data from satellite image.
- **Chapter 2:Literature Survey** Summarizes the research on different satellite image classification.
- **Chapter 3: Satellite Image Classification** Discusses in depth about the role & methodology of SVM in satellite image classification.
- **Chapter 4: Results** Determines that SVM classification method gives a high correct classification rate.
- **Chapter 5:Advantages & Disadvantages** List out the advantages and disadvantages of support vector machine for image classification.
- **Chapter 6:Real World Applications**List out the real world applications.
- **Chapter 7:Conclusion** The overall development and its inferred results are concluded with probable best practice.
- **References** Includes the references for satellite image classification for future purpose.

CHAPTER 2
LITERATURE SURVEY

2.1 Satellite Image Classification Methods & Techniques

ABSTRACT

Satellite image classification process involves grouping the image pixel values into meaningful categories. Several satellite image classification methods and techniques are available. Satellite image classification methods can broadly classified into three categories 1)automatic 2) manual and 3) hybrid. All three methods have their own advantages and disadvantages. Majority of the satellite image classification methods fall under first category. Satellite image classification needs selection of appropriate classification method based on the requirements. The current research work is a study on satellite image classification methods and techniques.

1. **Need Of Satellite Image classification**

   Satellite image classification is a process of grouping pixels into meaningful classes. It is a multi-step workflow.Satellite image classification can also be referred as extracting information from satellite images. Satellite image classification is not complex, but the analyst has to take many decisions and choices in satellite image classification process.Satellite image classification involves in interpretation of remote sensing images, spatial data mining, studying various vegetation types such as agriculture and foresters etc. and studying urban and to determine various land uses in an area.

   Satellite image classification plays a major role in extract and interpretation of valuable information from massive satellite images. Satellite image classification is required for:

   - Spatial data mining
   - Extract information for an application
   - Thematic map creation
   - Visual and digital satellite image interpretation
   - Field surveys

• Effective decision making

• Disaster management

2. **Satellite Image Techniques**:

   Satellite image classification methods can be broadly classified into three categories:

   • Automated

   • Manual

   • Hybrid

2.1.1 AUTOMATED

Automated satellite image classification methods uses algorithms that applied systematically the entire satellite image to group pixels into meaningful categories. Majority of the classification methods fall under this category. Automated satellite image classification methods further classified into two categories 1) supervised 2) unsupervised classification methods.

3. **Supervised classification:**

   Supervised classification methods require input from an analyst.The input from analyst is known as training set.Training sample is the most important factor in the supervised satellite image classification methods. Accuracy of the methods highly depends on the samples taken for training[5].Training samples are two types, one used for classification and another for supervising classification accuracy.• Major supervised classification methods uses the following statistical techniques:

   • Artificial Neural Network (ANN)

   • Binary Decision Tree (BDT)

   • Image Segmentation

   Supervised satellite image classification steps include:

   • Create training sites based on the ground truth sample positions(assuming the samples have been categorized to habitat type,then an area around the sample positions is used to create a training site)

   • prepare images

• Create signatures and apply to the images(Training sites extract data from images and creates signature for each habitat category that is then applied to the whole image)

•Transform output to create suitable map.

Various classification techniques deals with different kinds of similarity matching methods. Supervised classification includes additional functionality such as analyzing input data,creating training samples and signature files, and determining the quality of the training samples and signature files.

4. **ANN**

   Artificial Neural Network Algorithms fall under Artificial Neural Network (ANN) simulate human learning process to associate the correct meaningful labels to image pixels.Advantage of ANN based satellite image classification algorithms is easy to incorporate supplementary data in the classification process and improves classification accuracy.

5. **BDT**

   Binary Decision Tree (BDT) satellite image classification algorithms are machine learning techniques. Decision tree technique includes a set of binary rules that define meaningful classes to be associated to individual pixels. Different decision tree software are available to generate binary rules. The software takes training set and supplementary data to define effective rules.

6. **IMAGE SEGMENTATION**

   Segmentation plays a vital role in satellite image processing,analysis and pattern recognition. Satellite image segmentation techniques/algorithms are not directly related to image classification. Image segmentation groups pixels which are relatively homogeneous into segments. Image segmentation algorithms provide variables that support analyst to specify relative size and shape of the segments. Segmented image can be classified at segmentation level, instead of pixel level. Segmentation level satellite image classification algorithms are much faster than pixel level classification methods.

7. **Unsupervised classification**

   Unsupervised classification technique uses clustering mechanisms to group satellite image pixels into unlabeled classes/clusters. Later analyst assigns meaningful labels to the clusters and produces well classified satellite image. Most common unsupervised satellite image classification is ISODATA, Support Vector Machine (SVM) and K-Means.

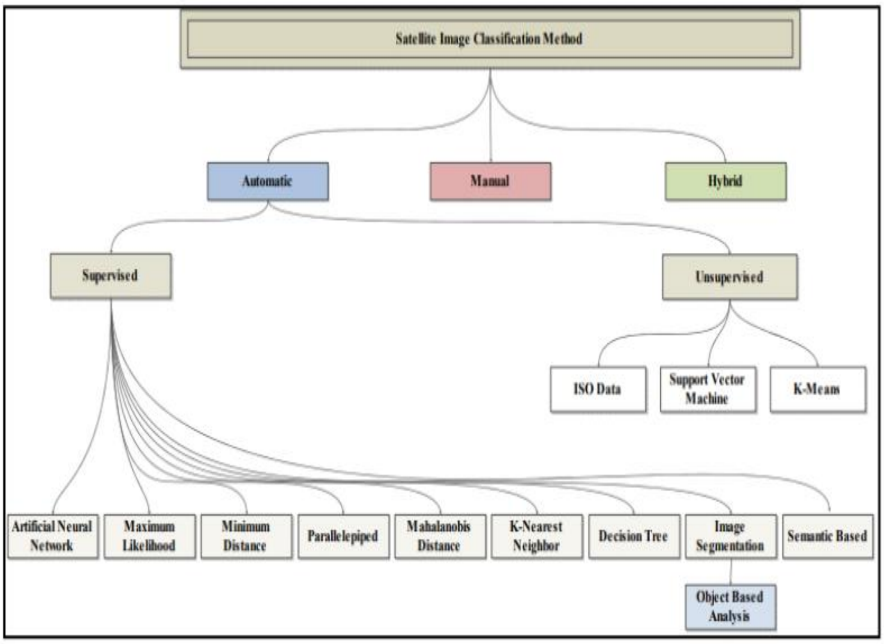

Figure 2.1: Satellite Image Classification Hierarchy

2.1.2 MANUAL

Manual satellite image classification methods are robust,effective and efficient methods. But manual methods consume more time. In manual methods the analyst must be familiar with the area covered by the satellite image. Efficiency and accuracy of the classification, depends on analyst knowledge and familiarity towards the field of study.

2.1.3 HYBRID

Hybrid satellite image classification methods combines the advantages of automated and manual methods. Hybrid approach uses automated satellite image classification methods to do initial classification, further manual methods are used to refine classification and correct errors.

Satellite Image Classification Methods

• High resolution satellite image classification using **fuzzy rule** is an classification method to classify images into specific classes using fuzzy logic. This method classifies satellite images into five major classes: shadow,vegetation, road, building and bare land. This method uses image segmentation and fuzzy techniques for satellite image classification. It applies two levels of segmentation, first level segmentation identifies and classifies shadow, vegetation and road. Second level segmentation identifies buildings. Further it uses contextual check to classify unclassified segments and regions. Fuzzy techniques are used to improve the classification accuracy at the borders of objects[8]

• **Support Vector Machine** (SVM) is a non- parametric supervised/unsupervised statistical classification method. This method can be used to extract land-use map. SVM works on the assumption that there is no information on how to distribute the overall data. SVM reduces satellite classification cost,increases speed and improves accuracy[13]

• **Decision tree** is supervised image classification technique. This method extracts features from satellite image based on pixel color and intensity. Extracted features assist to determine objects reside in the satellite images. The methods classifies satellite images using decision tree with the support of identified objects[11]

• **K-Means** is a popular statistics and data mining technique. It partitions n observations into k clusters based on Euclidean mean value. Advantages with the K-Means technique are simple to process and fast execution. Limitation with this method is analyst should know priori number of classes[2]

• **ISODATA** technique[3] is most common unsupervised satellite classification method. It creates predefined number of unlabeled clusters/classes in a satellite image. Later meaningful labels are assigned to the clusters. ISODATA parameters needs several parameters that control number of clusters and iterations to be run. In few cases clusters may contain pixels of different classes. In such situations ISODATA[7] uses cluster-busting technique to label the complex classes.

• **Minimum distance** approach calculates mean spectra of each predefined class and assigns the pixel to a group that has the least distance to mean. It easy to execute and simple to process. But minimum distance method considers only mean value. Mahalanobis distance method is very similar to minimum distance method. It uses statistics technique covariance matrix for satellite image classification.

• **Parallelepiped** executes based on parallelepiped-shaped boxes for each class. Parallelepiped boundaries for each class are pre-determined. Pre-determined boundaries identifies checks pixels of test images and determine class of the pixel. Parallelepiped method is fast and easy to run, but overlap may produce false results.

Comparison of satellite image classification methods

Various researchers have been performed comparison on unsupervised, supervised satellite image classification methods and on the combination of both with respect to classifi-

Researcher	Classification Methods Taken for Comparison	Test Data	Better Method from the Researcher Study
K. Kanika et al., [23]	K-Nearest Neighbour Minimum Distance Maximum Likelihood	IRIS Plants Dataset	K-Nearest Neighbour
R. Offer et al., [24]	ISODATA Maximum Likelihood Hybrid Method	Desert Outlay Datasets	Hybrid Method
A. Aykut et al., [25]	Maximum Likelihood Minimum Distance Parellelpiped	Landsat 7 ETM+ Images	Maximum Likelihood
T. Jamshid et al., [26]	Parallelepiped Minimum Distance Chain Method	Landsat 5TM images	Chain Method
H. N. Shila et al., [27]	Unsupervised Supervised Hybrid Method	Landsat7 ETM+ data	Hybrid Method
N. Maryam et al., [28]	Support Vector Machine Maximum Likelihood Mahalanobis Distance Minimum Distance, Spectral Information Divergence Binary Codes Parallelepiped	Landsat7 ETM+ data	Support Vector Machine
Manoj Pandya et al., [29]	K-Means ISODATA Minimum Distance Maximum Likelihood Parallelpiped Seeded region Growing Enhanced Seeded region Growing	Landsat, SPOT and IRS Datasets	Enhanced Seeded Region Growing
T. Subhash et al., [30]	Maximum Likelihood Minimum Distance Mahalanobis Distance	Landsat7 ETM+ data	Maximum Likelihood
W. Malgorzata et al., [31]	Pixel-based Classification Object-Oriented Classification	Multi-Spectral Satellite Images	Object-Oriented Classification

Figure 2.2: comparison of satellite image classification method

cation accuracy and kappa coefficient. This section compares comparison summary of various researchers.Above figure provides comparison summary of different researcher's conclusion. From the comparison summary, researcher's opinion on better satellite image classification method is not consistent. Further there is a need to study weather satellite image classification methods performance depends on test dataset.

Conclusion

This paper gives a summary on automated satellite image classification methods and compares several reviews done by various researchers. Automated satellite image classification methods can be classified into 1) supervised 2) unsupervised. Supervised and unsupervised satellite image classification methods differ in the way of grouping pixels into meaningful categories. In the literature, researchers have presented survey on satellite image classification methods and evaluated the performance against different datasets. This paper summarizes the various reviews on satellite image classification methods

2.2 Satellite Image Classification Based on Fuzzy with Cellular Automata

Abstract:

Satellite image classification is a significant technique used in remote sensing for the computerized study and pattern recognition of satellite information, which make possible the routine explanation of a huge quantity of information. Nowadays cellular automata are implemented for simulation of satellite images and also cellular automata relates to categorization in satellite image is used simultaneously. Based on information of stored image value to the cell and dimension of neighbourhood cells. Inoder fine tune classification rate of cellular automata algorithm fuzzy rules with cellular automata are used . In this paper cellular with fuzzy rules have been implemented for classifying the satellite image and quality of classified image is analyzed.

Introduction

Remote sensing is the acquisition of physical information of an object without touch or contact[13]. The representation is acquired by the sensor which is fitted to the scanning object and it is sent back to the earth station for processing and analysis.

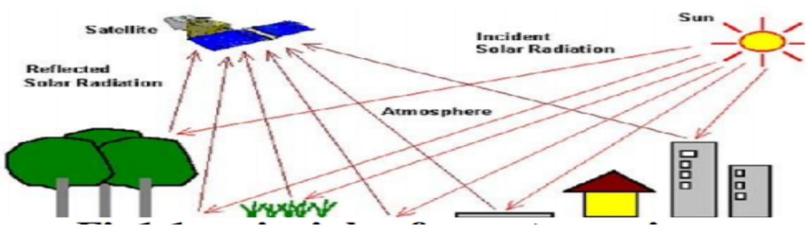

Figure 2.3: Remote sensing

There are numerous image investigation techniques obtainable and the methods used depend on the requirements of the precise trouble concerned. In image segmentation and classification algorithms are used to outline diverse areas in an image into thematic classes. The resulting creation is a thematic map of the revise area. This thematic map can be combined with erstwhile databases of the test area for supplementary analysis and utilization. The classifications algorithms are one of the most significant methods used in remote sensing that assist developers to understand the information contained in the satellite images. The aspire of satellite images categorization is to split image pixels into discrete classes. The consequential classified image is really a thematic map of the unique image. These algorithms have reached a immense progress in the last years. The analysts utilize the classification algorithms to deduce the information contained in the satellite images.

2.2.1 Cellular Automata

A cellular automata is a discrete dynamical system that consists of a regular network of finite state automata (cells) that change their states depending on the states of their neighbors, according to a local update rule[6]. All cells change their state simultaneously, using the same update rule. The process is repeated at discrete time steps. It turns out that amazingly simple update rules may produce extremely complex dynamics when applied in this fashion. A well known example is the Game-of-life by John Conway. Cellular automata are:

- discrete in both space and time.
- homogeneous in space and time
- local in their interactions

The fundamental proposal is very straightforward: a cellular automaton evolves in discrete time-steps by updating its states according to a evolution rule that is applied commonly and synchronously to every cell at every time-step. The importance of all cell is determined based on a geometric composition of neighbour cell.

Working:

States: Each cell can acquire an integer value that corresponds to its existing state. There is a limited set of states.

Neighbourhood: A set of cells that correlate with the present one.

Transition function f : Takes as input arguments the cell and neighbourhood states, and returns the recent state of the existing cell.

Rules: The transition function f utilizes a lay down of rules that recognize how the states of the cells vary.

Iterations: The transition assignment f is applied to every cell of the lattice across several iterations.

2.2.2 Implementation

fuzzy Implementation:

The FIS Editor GUI implements authorizing the maximum level features of the fuzzy inference system, such as the quantity of input and output variables. The FIS Editor is the high-level display for whichever fuzzy logic inference system. It agree to to call the various other editors to function on the FIS.

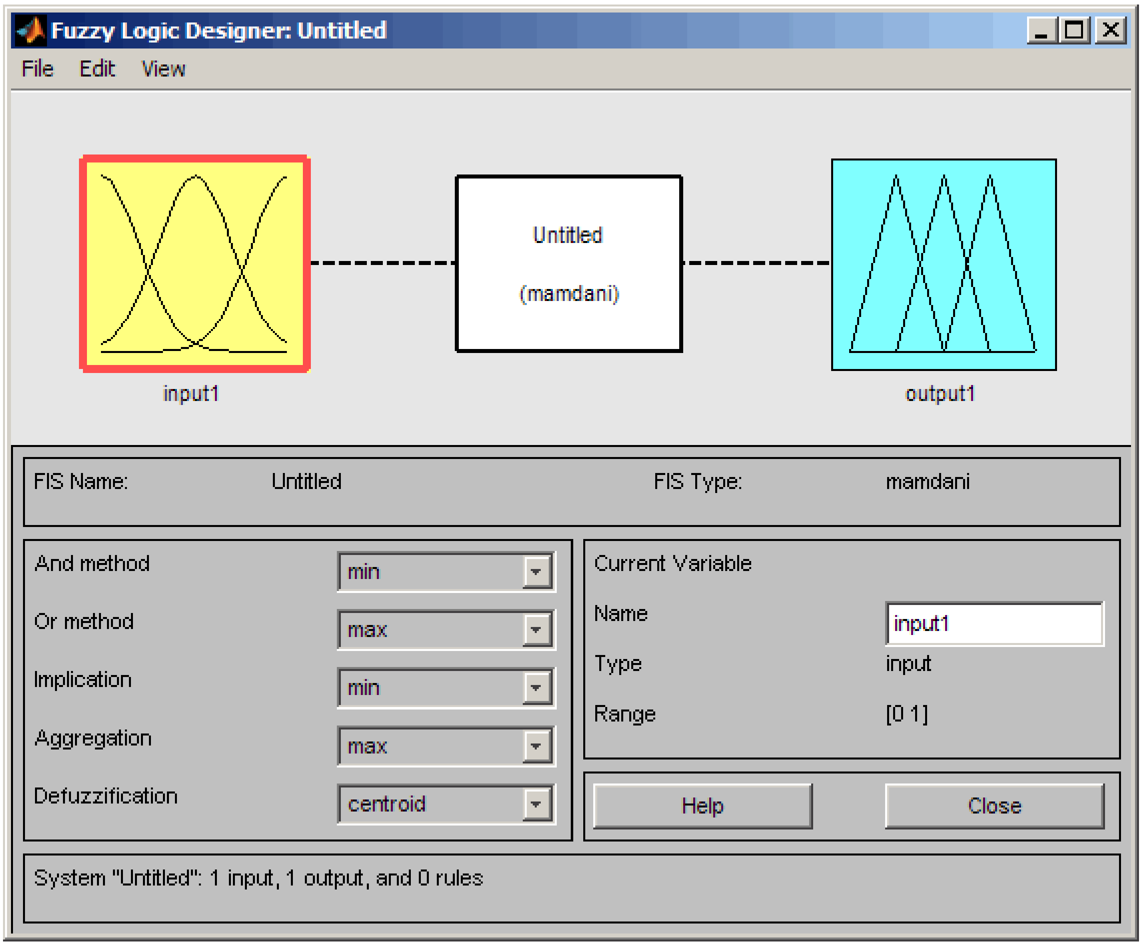

Figure 2.4: FIS Editor Window

Membership function:

The Membership Function Editor is the tool that lets you display and edit all of the membership functions associated with all of the input and output variables for the entire fuzzy inference system.Membership functions are used in the fuzzification and defuzzification steps of a FLS (fuzzy logic system), to map the non-fuzzy input values to fuzzy linguistic terms and vice versa.

A membership function for a fuzzy set A on the universe of conversation X is defined as μa :X → [0,1], where each element of X is mapped to a rate between 0 and 1. This rate, called membership value or degree of membership, quantifies the standing of membership of the constituent in X to the fuzzy set A. Membership function agree to to graphically represent a fuzzy set. The x axis represents the creation of conversation, whereas the y axis represents the degrees of membership in the [0,1] interval. Easy functions are used to build membership functions. since are defining fuzzy concepts, using supplementary complex functions does not include extra accuracy.

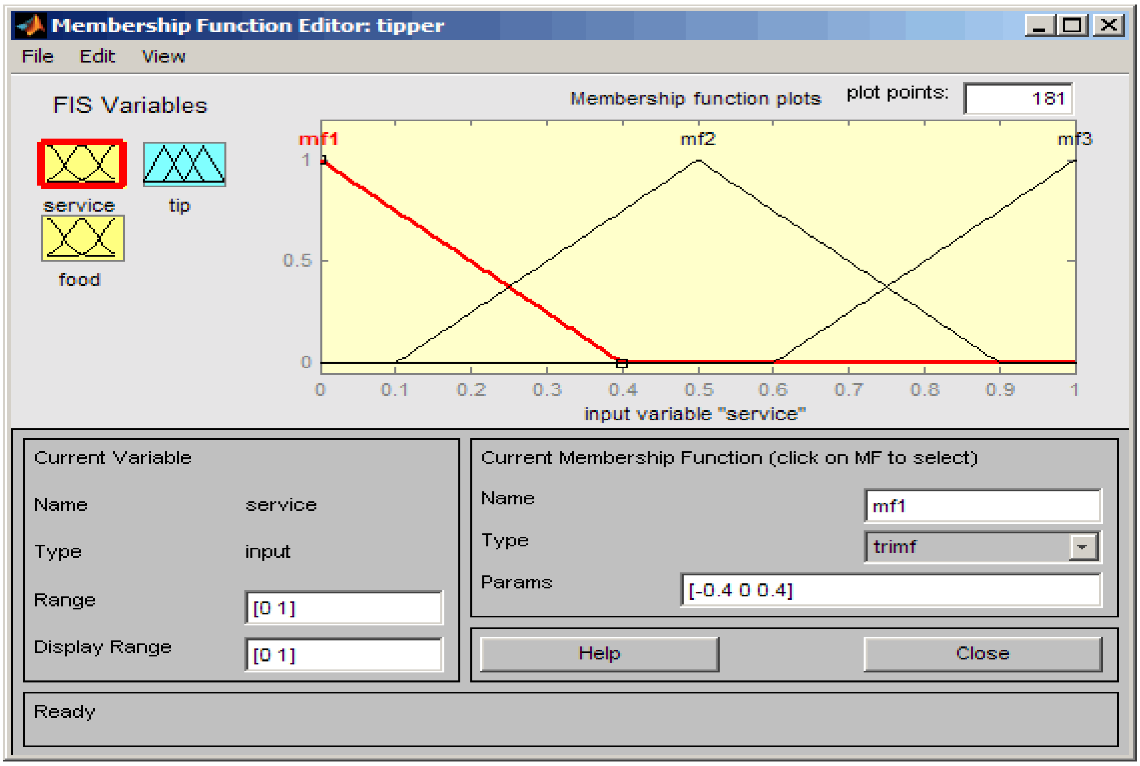

Figure 2.5: Membership Function Editor

Rule Viewer

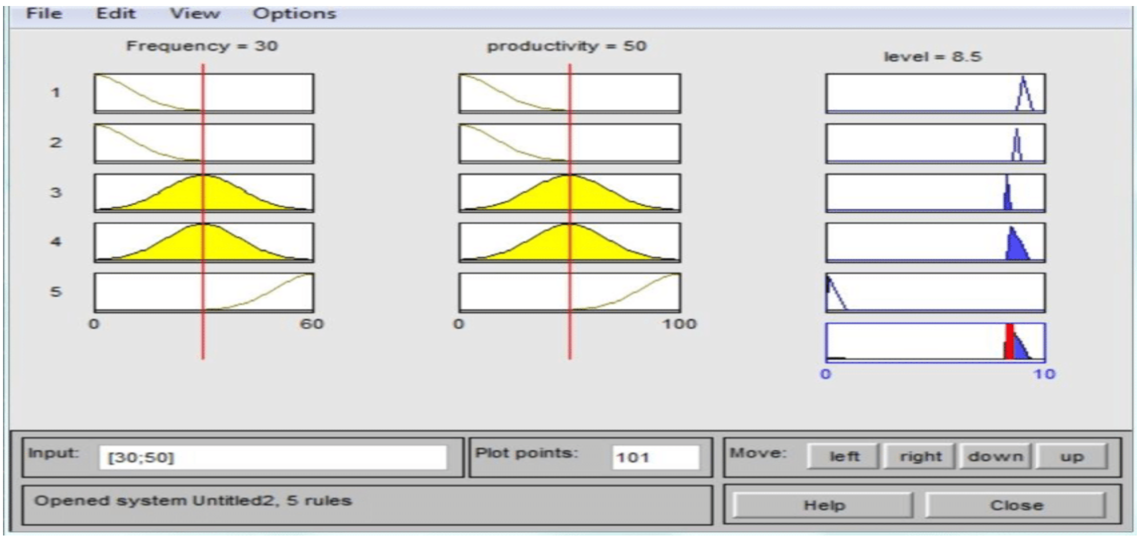

Figure 2.6: Rule Viewer Window

The Rule Viewer tolerates to assume the absolute fuzzy inference progression at once. The Rule Viewer also demonstrates the plan of exact membership functions influences on the whole outcome. It presents a sort of micro view of the fuzzy inference system. To distinguish the whole output surface of system then the complete length of the output set based on the complete duration of the input to unlock up the Surface Viewer. You can adjust the input values and view the corresponding output of each fuzzy rule, the aggregated output fuzzy set, and the defuzzified output value. To view the inference process, specify the input and output variables of FIS, their corresponding membership functions, and the fuzzy rules for the system.

Above figure explains:

The first column in this preparation corresponds to the input variables.

The second column communicates to the output changeable.

The third column displays the weight applied to every rule.

Surface Viewer Window

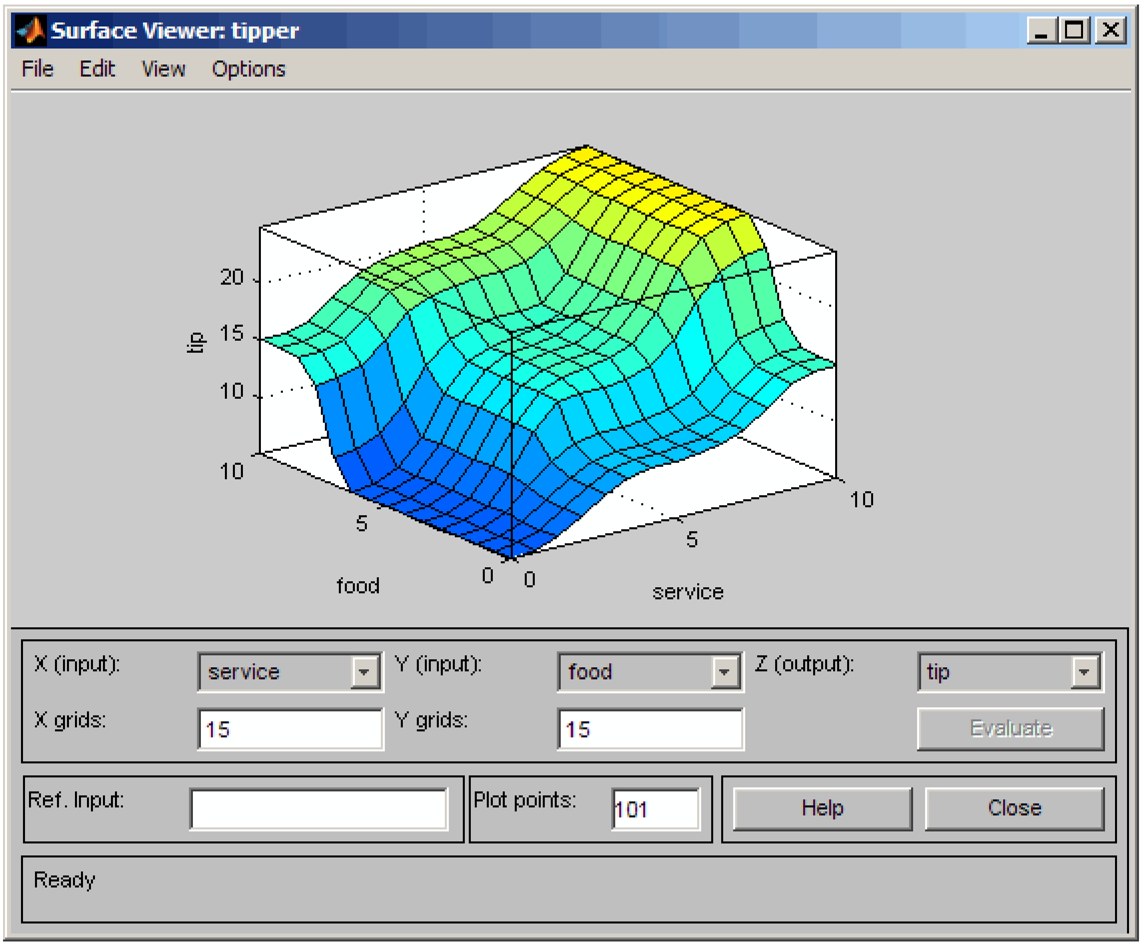

Figure 2.7: Surface Viewer

METHODOLOGY:

DATASET: The input image is taken for production of land covers is central coast of Tamil Nadu comprising Cuddalore and Villupuram districts and various land features are needed to be identified, the land features consists of rivers, agricultural land, settlement area and waste land needed to be classified by means of supervised and unsupervised algorithm and it accuracy assessment every class and algorithm used is analyzed.

METHOD OF IMPLEMENTATION: The Indian remote sensing satellite data product is implemented and tested in MATLAB environment and image value is about 256 x256 is used, cellular automata is applied as image classification process based on the neighbourhood value pixel assigned to each class in different iterations ,in this method upto 10 iterations carried out and percentage of edge, focus, noisy and uncertain pixel is show in command window if percentage of uncertain pixel is more then incorporate fuzzy rule to cellular to eliminate uncertain pixel and it may enhance the classification accuracy.

RESULTS:

Pixel type	percentage
Edge pixel	0.38
Focus pixel	0.17
Noisy pixel	0.06
Uncertain pixel	0.43

Figure 2.8: output of cellular automata

Pixel type	Percentage
Edge pixel	0.54
Focus pixel	0.17
Noisy pixel	0.06
Uncertain pixel	0.23

Figure 2.9: output of fuzzy with cellular automata

Conclusion

The accuracy rate of cellular automata is very much increased when fuzzy rules are implemented to this systems, uncertain pixels which are still present in the classification process is eliminated and uncertain pixel is classified to each class its give paths to well classified image with high accuracy.

2.3 Satellite Image Classification using convolutional learning

Introduction

The classification task is one of the most important topics in computer vision, especially in satellite image classification. Many classification methodologies have been proposed for this task[1]. Traditionally, classification schemes include a feature extractor and a classifier. For decades, various types of classifiers as well as various feature extraction methods have been applied to various tasks. Since most of conventional classifier schemes such as Support Vector Machine(SVM) cannot learn directly from the original data, their accuracy rates heavily depend on the feature extractors they adopt including Discrete Cosine Transform (DCT), Hough Transform, and Hue-Saturation-Value (HSV). On the contrary, deep learning addresses this problem of dependency on feature extractors. As a class of deep learning models, Convolutional Neural Network (CNN) has been known as the first truly successful deep networks architecture designed for 2 dimensional data. In this paper,explains about the advantages of CNN for a satellite image classification problem and a CNN architecture is also proposed for the satellite image classification problem.[9]

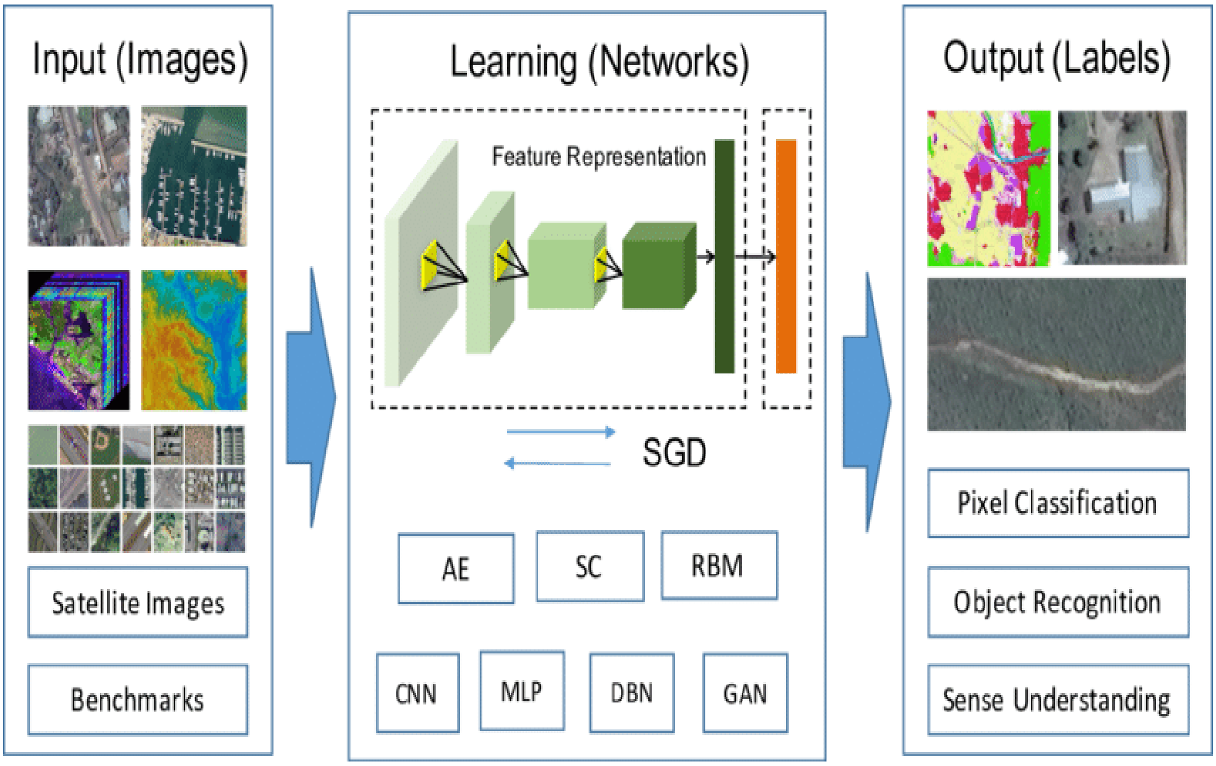

Figure 2.10: Framework of CNN

2.3.1 CNN

CNN is a class of deep neural networks,most commonly applied to analyzing visual imagery.it is an deep learning algorithm which can take in an input image,assign importance to various aspects/objects in the image and be able to differentiate one from other.CNN work because it's a good extension from the standard deep learning algorithm.However,CNN is more efficient because it reduces the number of parameters.CNNs are used for image classification and recognition because of its high accuracy.One of the most popular algorithm used in computer vision today is Convolutional Neural Network or CNN. Convolutional Neural Networks have a different architecture than regular Neural Networks. Regular Neural Networks transform an input by putting it through a series of hidden layers. Every layer is made up of a set of neurons, where each layer is fully connected to all neurons in the layer before. Finally, there is a last fully-connected layer — the output layer — that represent the predictions[10]. Convolutional Neural Networks are a bit different. First of all, the layers are organised in 3 dimensions: width, height and depth. Further, the neurons in one layer do not connect to all the neurons in the next layer but only to a small region of it. Lastly, the final output will be reduced to a single vector of probability scores, organized along the depth dimension.

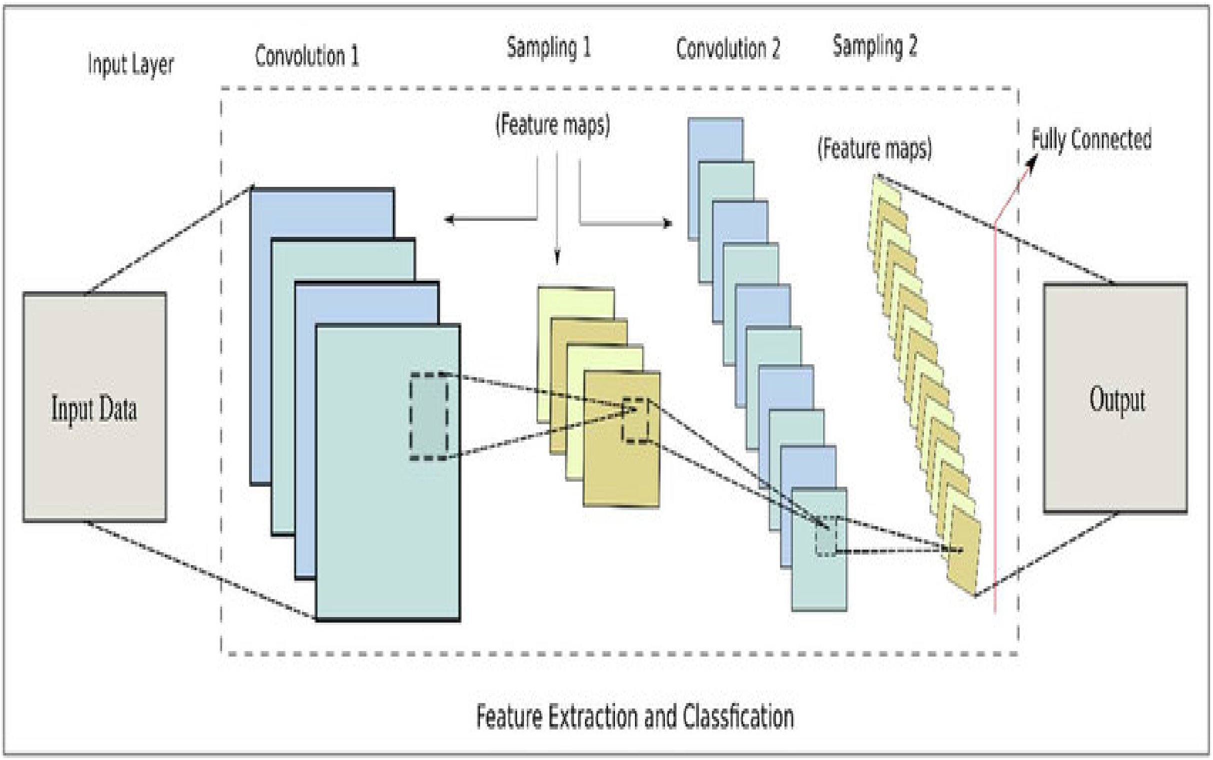

Figure 2.11: General Architecture

CNN is a class of multi-layer neural networks. It is known as the first successful deep architecture for images. As the generic deep learning algorithm, the abstract level increases from layers to layers. Designing of CNN consists of two types of hidden layers: convolution and sub-sampling layer. The key idea of CNN is that it minimizes the parameters required to learn by pooling filter responses over the convolution and sub-sampling layers to obtain the high level abstract feature.The convolution layers perform convolutions over feature maps in previous layers. The last layer of a CNN is usually fully connected with the previous layer. It can be considered as a linear classifier of features obtained by repeating convolution and sub-sampling processes.

Satellite Image Classification

Convolutional Neural Networks (CNNs) is the most popular neural network model being used for image classification problem. The big idea behind CNNs is that a local understanding of an image is good enough.CNNs are used for image classification and recognition because of its high accuracy.

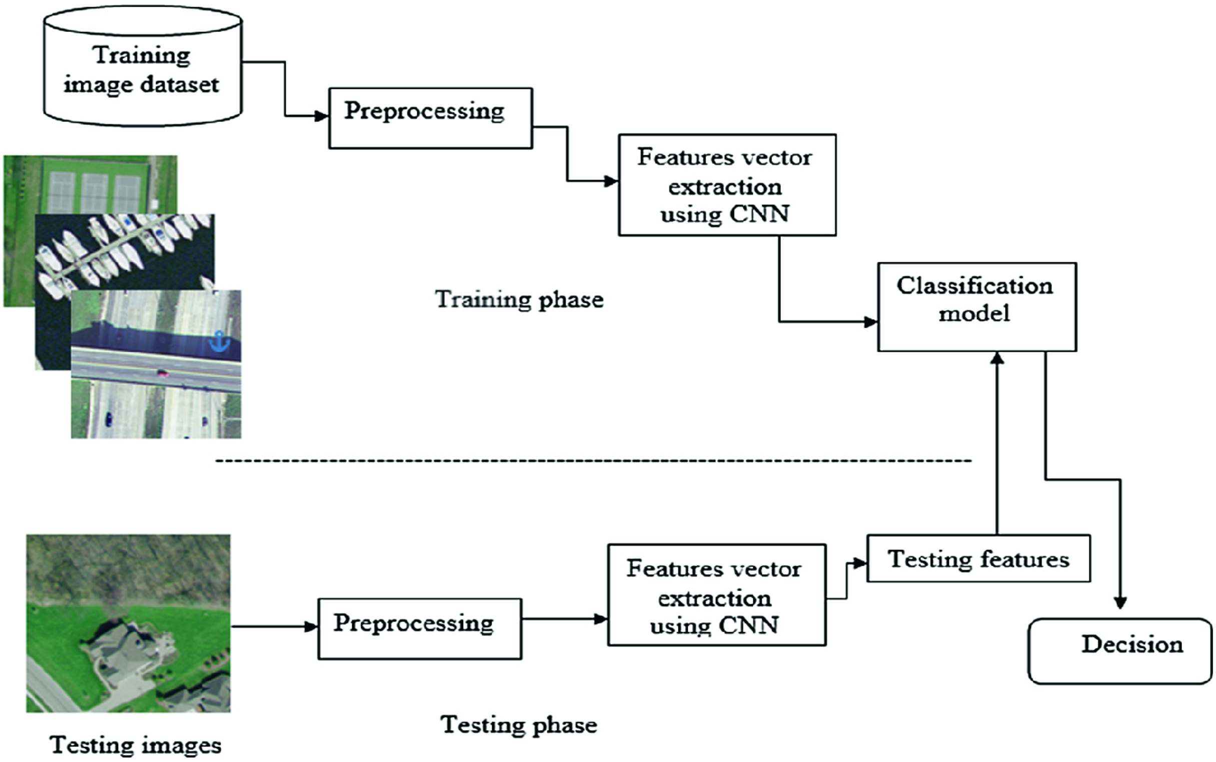

Figure 2.12: Satellite Image Classification

Working:

After observing the characteristics of satellite image dataset, here proposed a CNN architecture for training these features.

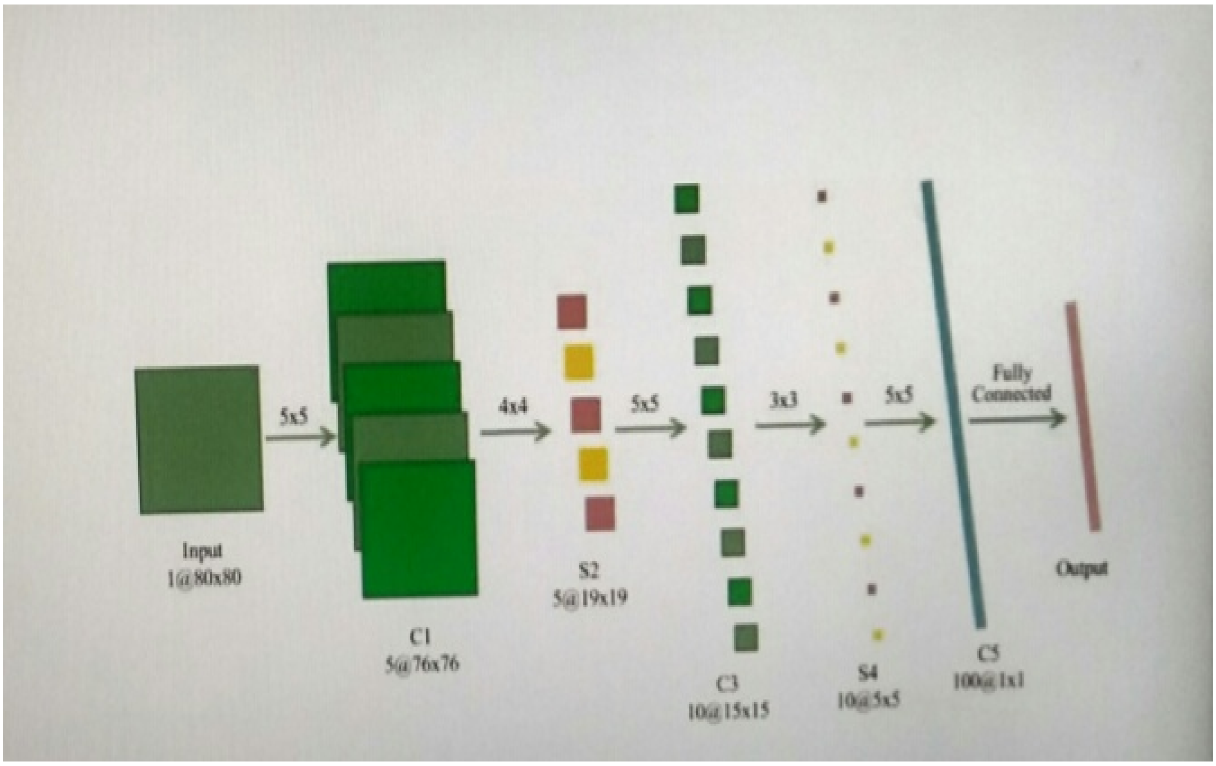

Figure 2.13: Architecture of the proposed CNN

Unlike other object image dataset, satellite images usually have a high similarity among samples in each class since the images are obtained from long distance. Therefore, not need a CNN with many feature maps in each layer. Furthermore, in the convolution layer, instead of convolving all the maps in the previous layer, only two maps are selected. The architecture of this network with six layers is shown in above figure. In the input layer, the original image is converted to gray image and resized to our standard 80×80 size. The first convolution layer (C1) performs the convolution operation from the input using five kernels of the size 5 × 5 and 5 biases to produce five 76 × 76 maps. As a total, there are 130 parameters to be trained in this later. The next subsampling layer (S2) simply computes spatial sub-sampling of each map in C1. Hence, S2 has the same number of maps as C1. It uses 5 weights and 5 biases to produce five 19×19 maps. There are 10 trainable parameters in this layer.In C3, any two possible difference maps in S2 are convolved with 5× 5 kernels to compute one map in C3. Hence, 10 maps are used in C3 since S2 has 5 maps. By doing so, this layer has 10 of 15×15 maps and 2×5×5×10+10 = 510 parameters to be trained. Similarly, S4 has 10 5 × 5 maps and 20 parameters. C5 uses 5 × 5 kernels to compute 100 1×1 maps. Each map in C5 is convolved with all maps in S4. Therefore, C5 has 100×10×5×5 = 25000 parameters. The final layer is the out layer, which is fully connected with C5, and it produces a vector with a same dimension as the number of classes.

2.4 Classification of satellite images using perceptron neural network

Abstract:

Image classification is an important part of digital image analysis and is defined as a process of categorizing the pixels into one of the object classes present in the image. As a prerequisite to image classification, a number of processes such as image enhancement, segmentation and feature extraction are required. This paper presents the classification of satellite images using Perceptron Neural Network with the transfer function hardlim and the learning rule learnpn. Before classification, the enhanced images are divided into a number of blocks and feature extraction is carried out using Principal Components Analysis (PCA)[13]. As color plays an important role in differentiating the objects in the satellite images, color information is used in extracting significant features. Fifty images from Landsat are used for training and testing of the results. The objects in the categories of water, land and vegetation are identified based on RGB components. Accuracy assessment and comparison is carried out using confusion matrix. It is concluded that choosing an appropriate block size affects the classification accuracy.

Introduction

Image classification has always been an important issue in any image processing system.The process of classification is mainly dependent on the type of learning approach that can be either supervised or unsupervised. In supervised approach, the labels for the target data needs to be specified as data to the classifier whereas in unsupervised approach, the targets need not be supplied. A number of classification techniques are available for identifying object classes in digital images. The statistical methods such as maximum likelihood classification may not produce highly accurate results. On the other hand, the techniques based on machine learning such as Artificial Neural Network (ANN)[9] and Support Vector Machine (SVM)[13] produce high accuracy even from fewer training examples as compared to their statistical counterparts. Neural Networks is one of the popular techniques in computational intelligence in the area of machine learning and artificial intelligence. It is based on working of biological neurons in the human body that are capable of complex computations. A neural network comprises of a number of layers, one each for input and output and one or multiple hidden layers. The input layer nodes correspond to data sources, the output layer nodes refer to number of object classes to be recognized and hidden layer nodes carry out the task of computation. Perceptron, Multilayer Perceptron (MLP), Radial Basis Function (RBF)[13] are few types of neural networks.

Neural networks are non-parametric in nature and they do not assume any prior knowledge about data distribution. In the supervised approach used for training a neural network, the network must be taught the characteristics of dataset being used in the study. The neural networks are designed in such a way that they learn to adapt the weights of nodes in input and hidden layers. With each iteration, the weights are used by learning functions in order to compute the error rate. Based on error rates, the Neural Network progresses towards producing better classification results. The selection of suitable techniques for image classification may differ from one application to another. Choosing a suitable classification technique for better interpretation of object classes in remotely sensed images is a difficult task. In case of remotely sensed images, images are captured without physical contact with ground surface. A number of different sensors are used to capture the images over various regions. Such images include satellite and aerial images and these images prove useful in the fields such as determining land use patterns, environmental analysis, weather forecasting, vegetation monitoring and other related areas. Image enhancement is often required for satellite images in order to identify the objects and extract features and their coordinates from images. For satellite images, a number of databases are available such as Landsat, Google Earth, Imageseer and so on. The purpose of Landsat is to archive images of earth and gather facts about natural resources of our planet. Though individual small elements may not be visible in these images but large structures are clearly visible for analysis and interpretation. Landsat data has been used to support wide range of applications and is used by government, commercial, industrial and educational communities throughout the world. The classification of satellite images is useful for research, agriculture, change detection, forestry, mining and many more to mention. Satellite images consist of multiple bands and are called multispectral images[13]. The objects present in these images have unique spectral signatures that are useful in identification of different classes present in an image. For this research, the visible RGB bands have been used. It has been observed that even with RGB bands it is possible to obtain a lot of information that can be very useful in analyzing the object categories. The color features in the images have been used as the basis for feature extraction and classification. A color is usually assigned for every object class that needs to be identified in the image.

Methodology

2.4.1 Dataset:

Some of the sample images of earth taken from Landsat satellite containing these classes are shown in the figure.Different sets of enhanced images were used for generating training and test datasets. The training and test datasets are created by extracting parts of these images from the areas indicating the presence of Water, Land and Vegetation which are also termed as Regions of Interest (ROIs) corresponding to classes. For better interpretation of object classes, a classification color is assigned to each object class that needs to be recognized by a classification algorithm. Usually, vegetation is represented by green color, water by blue color, soil using brown or grey color, snow by white color and so on. The sizes of these blocks from images are taken as 16X16 and 8X8 pixel based on size of objects in the satellite images. Total 150 blocks (50 in every class) each for training and test dataset for both block sizes are collected.

Figure 2.14: Sample Images Of Earth

2.4.2 Feature Extraction:

Since the satellite images contain huge amount of information in multiple bands, there arises the need of feature extraction that reduces the input in the form of features that acts as input to a classification algorithm. The features generated may be large in number but only required number may be used for classification. Principal Component Analysis (PCA) has been widely used as a technique for dimensionality reduction for data analysis. It finds its importance as the most popular technique for reducing data dimensions in case of remotely sensed images. For this paper, feature extraction is performed using Principal Component Analysis (PCA)[4] on a set of blocks/subimages obtained from training and testing image. This version of PCA has been referred as Multiblock PCA but a variation has been implemented on the R, G and B color bands[13] separately for the three classes of land, vegetation and water. Also, different block sizes (16X16 and 8X8 pixels) are chosen for experiments. Focus is given on color feature as it can easily be used for discriminating various object categories in satellite images. For 16X16 pixel block size, for three RGB components, 50 training images for each object class are converted into column vectors to create matrices of size 256X50. This matrix is used to compute 64 principal components corresponding to each RGB component image for land, vegetation and water categories resulting in a combined feature matrix of size 192X150 for training dataset. The same process of feature extraction is repeated for 50 test images for each object class resulting in combined feature matrix of size 192X150 for test dataset. Similarly, for

8X8 pixel block size, for three RGB components, 50 training images for each object class are converted into column vectors to create matrices of size 64X50 which is used for obtaining 16 principal components corresponding to RGB component images for land, vegetation and water categories resulting in a combined feature matrix of size 48X150 for training dataset. The same process of feature extraction is repeated for 50 test images for each object class resulting in combined feature matrix of size 48X150 for test dataset. 1/4th of the total number of features obtained from covariance matrices are used for classifying objects.

2.4.3 Image Classification Using Perceptron Neural Network

The principal components for three classes are used as inputs for training the Neural Networks. Perceptron Neural Networks (PNN) is used for classification with hardlim as transfer function and learnpn as learning rule. The hard limit (hardlim) transfer function generates either 0 or 1 as output and learnpn is a weight and bias learning function. When input vectors have different magnitudes, it can result in faster learning rate.

Training image features and targets are input for training the perceptron whereas test images features are used for simulating the network for predicting the classes. One-vsall approach has been followed for classification of more than two object classes. A group of more than one classifier are used for building a multi-class classifier model using perceptron. Here, three classifiers are created with two classes each in both 16X16 and 8X8 pixel block sizes. The first classifier, Classifier1 is created by taking land as one class and combined vegetation and water as other class (not land). Similarly, Classifier2 is designed with vegetation as one class and combined land and water as other class (not vegetation) and Classifier3 considering water as one class and combined vegetation and land as other class (not water).

Target patterns are defined for each classifier based on block images of water, land and vegetation classes as a part of supervised learning for Neural Network. Also, one-to-one strategy is followed for defining target patterns. For both 16X16 and 8X8 pixel block sizes, three classifiers each for land, vegetation and water are trained with training images and simulated using test images.

Results

S.No	Block Size	Feature Dimensions	Classification Accuracy with Training Images	Classifier1 (for land)	Classifier2 (for vegetation)	Classifier3 (for water)	Overall Classification Accuracy with Test Images
1	16X16	192	100%	66.7%	98.0%	97.3%	87.3%
2	8X8	48	100%	98.0%	94.0%	98.7%	96.9%

Figure 2.15: Classification results obtained using Perceptron

	Classifier				Accuracy	Misclassification Rate
(a)	No. of Samples = 150		Predicted		0.98	0.02
			Land	Not Land		
	Actual	Land(50)	TP=47	FN=3		
		Not Land(100)	FP=0	TN=100		
(b)	No. of Samples = 150		Predicted		0.94	0.06
			Vegetation	Not Vegetation		
	Actual	Vegetation(50)	TP=41	FN=9		
		Not Vegetation(100)	FP=0	TN=100		
(c)	No. of Samples = 150		Predicted		0.987	0.0133
			Water	Not Water		
	Actual	Water(50)	TP=48	FN=2		
		Not Water (100)	FP=0	TN=100		

Figure 2.16: Confusion Matrix

Confusion matrices are generated for evaluating the classification results. First column shows confusion matrices from three binary classifiers for 8X8 pixel block- images from test images for land, vegetation and water classes showing the correct and incorrect prediction for every class. TP, FP, FN and TN denote True Positive, False Positive, False Negative and True Negative respectively. Second and third columns depict the accuracy and misclassification rates. For example, the first confusion matrix (a) for binary classification of land as one class and (vegetation+water) i.e, not land as second class shows that out of 50 instances of land class, 47 have been correctly identified and 3 have not been identified as land. The accuracy and misclassification rates are calculated as

Accuracy = (TP+TN) / No. of Samples

Misclassification Rate = (FP+FN) / No. of Samples

Conclusion

Based on the experimental results obtained by using supervised training of Perceptron Neural Networks, it can be observed that perceptron produces satisfying results for linear features. Also, one-vs-all approach can be used for multiclass classifier by combining multiple binary classifiers. The evaluation of classification accuracy is based on confusion matrices. The comparison is done performing classification based on combined RGB features for three object classes of land, vegetation and water. It is concluded that choosing an appropriate block size affects the classification results.Using very large block size is unable to discriminate between classes. Therefore, the optimal sized blocks of 16X16 and 8X8 pixels generate acceptable results. Classification can also be performed using support vector machines[13]

CHAPTER 3

SATELLITE IMAGE CLASSIFICATION USING SVM

Image classification is a supervised learning problem:define a set of target classes(objects to identify images),and train a model to recognize them using labelled example photos.It is defined as the task of classifying an image from a fixed set of categories.

Several machine learning techniques are used for image classification such as support vector machine,Decision tree,KNN(K-Nearest Neighbor,etc..SVM plays a important role in image classification.

3.1 Machine Learning

Machine learning is an application of artificial intelligence (AI) that provides systems the ability to automatically learn and improve from experience without being explicitly programmed. Machine learning focuses on the development of computer programs that can access data and use it learn for themselves.

We feed in DATA(Input) + Output, run it on machine during training and the machine creates its own program(logic), which can be evaluated while testing.

Machine learning categories:

Machine learning is generally categorized into three types: Supervised learning,Unsupervised learning,Reinforcement learning.

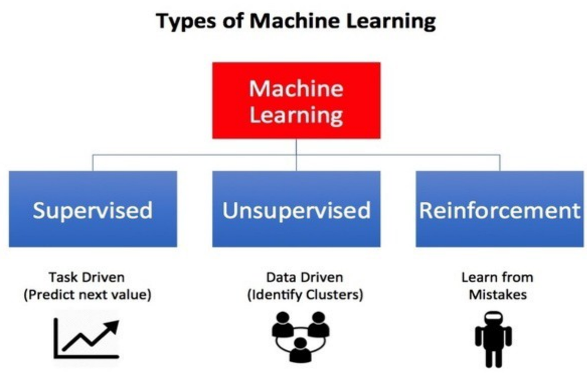

Figure 3.1: Types of Machine learning

3.2 Supervised Learning

Supervised learning is a learning in which we teach or train the machine using data which is well labeled that means some data is already tagged with the correct answer. After that, the machine is provided with a new set of examples(data) so that supervised learning algorithm analyses the training data(set of training examples) and produces a correct outcome from labeled data.Supervised machine learning helps to solve various types of real-world computation problems.Two of the most common supervised machine learning tasks are classification and regression.

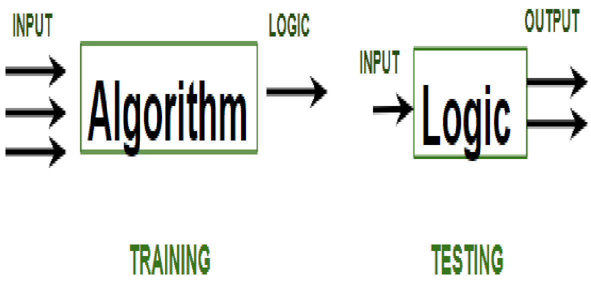

Figure 3.2: Supervised learning

Techniques of Supervised Learning

Classification is one of the most important aspects of supervised learning.Supervised learning problems can be further grouped into Regression and Classification problems. Both problems have as goal the construction of a succinct model that can predict the value of the dependent attribute from the attribute variables. The difference between the two tasks is the fact that the dependent attribute is numerical for regression and categorical for classification. Most commonly used classification techniques are: Support vector machine,Multi-class classification.

3.3 Unsupervised Learning

In the case of unsupervised learning algorithm, the data is not explicitly labeled into different classes, that is, there are no labels. The model is able to learn from the data by finding implicit patterns. Unsupervised Learning algorithms identify the data based on their densities, structures, similar segments, and other similar features.Cluster analysis is one of the most widely used techniques in supervised learning.

3.4 Reinforcement Learning

Reinforcement learning refers to goal-oriented algorithms, which learn how to attain a complex objective (goal) or maximize along a particular dimension over many steps. This method allows machines and software agents to automatically determine the ideal behavior within a specific context in order to maximize its performance. Simple reward feedback is required for the agent to learn which action is best; this is known as the reinforcement signal.

3.5 Major Contribution

:In this paper,emphasize an efficient image representation in classification task using support vector machine(SVM) with the Radial Basis Function(RBF) Kernel as it is an efficient supervised classification technique.Major contributions of this paper includes:

1. Explains the two main stages for image classification and evaluating the effectiveness of supervised SVM for image classification purposes.
2. Comparing the results found to those reported in the literature.

3.6 Proposed Framework

This is an study for the classification of satellite images into Landsat Multi-Spectral Scanner(MSS) to identify the areas of land use.The image is pre-processed and classified using support vector machine with the Gaussian RadiaL Basis Function(RBF) kernel.

3.6.1 Support Vector Machine

SVM are based on the concept of decision planes that define decision boundaries . A decision plane is one that separates between a set of objects having different class memberships. It builds a hyperplane from the training data which separates pixels with different class memberships this is called a linear classifier. The hyperplane gives the minimum distance to the training samples. Larger the margin is good separation achieved by the hyperplane that has the largest distance to the nearest training data.

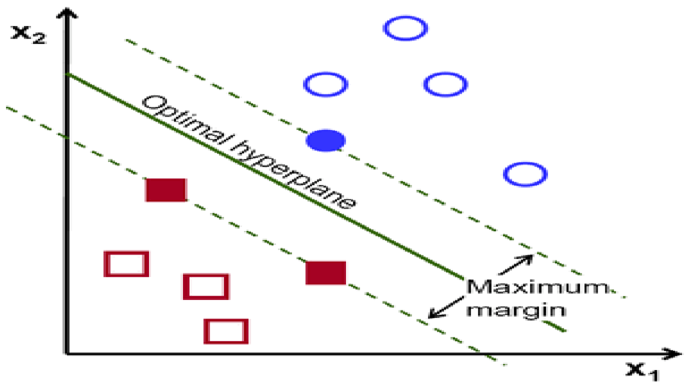

Figure 3.3: SVM

Hyperplanes are decision boundaries that help classify the data points. Data points falling on either side of the hyperplane can be attributed to different classes. Also, the dimension of the hyperplane depends upon the number of features. If the number of input features is 2, then the hyperplane is just a line. If the number of input features is 3, then the hyperplane becomes a two-dimensional plane. It becomes difficult to imagine when the number of features exceeds 3.

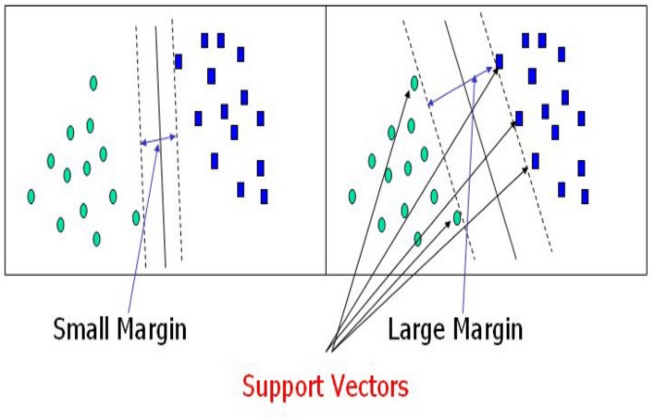

Figure 3.4: Support Vectors

Support vectors are data points that are closer to the hyperplane and influence the position and orientation of the hyperplane. Using these support vectors, we maximize the margin of the classifier. Deleting the support vectors will change the position of the hyperplane. These are the points that help us build our SVM. The hyperplane is normally represented by weights w and bias b, b an element in real numbers, x corresponds to the data values. SVM is usually maximizes the margin between data values of opposite classes.

$$w.x + b = 0$$

In the case of nonlinear classifier,kernel functions are used.

3.6.2 Kernel Function

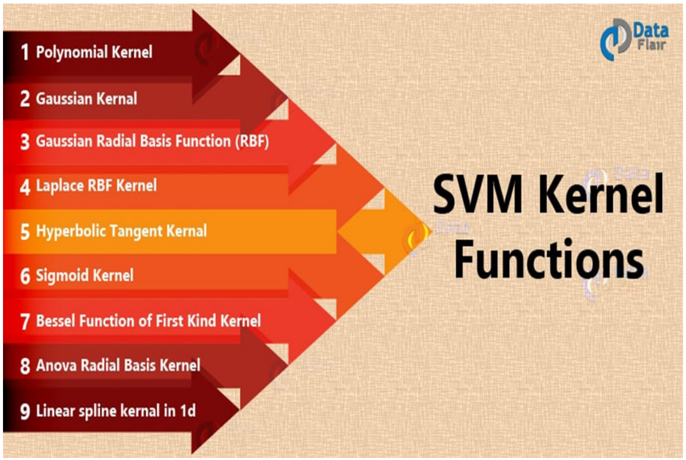

Figure 3.5: Kernel Functions

The function which transforms these data points to higher dimensional space from lower dimensional space.

SVM algorithms use a set of mathematical functions that are defined as the kernel. The function of kernel is to take data as input and transform it into the required form. Different SVM algorithms use different types of kernel functions. These functions can be different types. For example linear, nonlinear, polynomial, radial basis function (RBF), and sigmoid. Introduce Kernel functions for sequence data, graphs, text, images, as well as vectors. The most used type of kernel function is RBF. Because it has localized and finite response along the entire x-axis. The kernel functions return the inner product between two points in a suitable feature space. Thus by defining a notion of similarity, with little computational cost even in very high-dimensional spaces.

3.6.3 SVM kernels:

1) Polynomial Kernel

It is popular in image processing.

$$k(xi,xj) = (xi * xj + 1)^d \tag{3.1}$$

2) Gaussian radial basis function (RBF)

It is a general-purpose kernel; used when there is no prior knowledge about the data.

$$k(xi,xj) = exp(-\gamma||xi - xj||^2) \tag{3.2}$$

3) Sigmoid Kernel

We can use it as the proxy for neural networks.

$$K(X,Y) = tanh(\alpha x^T y + c) \tag{3.3}$$

3.7 Methodology

Satellite image classification process involves grouping the image pixel values into meaningful categories Several satellite image classification methods and techniques are available [1].

Supervised Classification methods require a target solution for each training data from an analyst.Input data can be classified into pixel base, local base, global base. This is an study of a supervised Learning for satellite image classification based on pixel base and local base inputs. Classification is done by SVM.

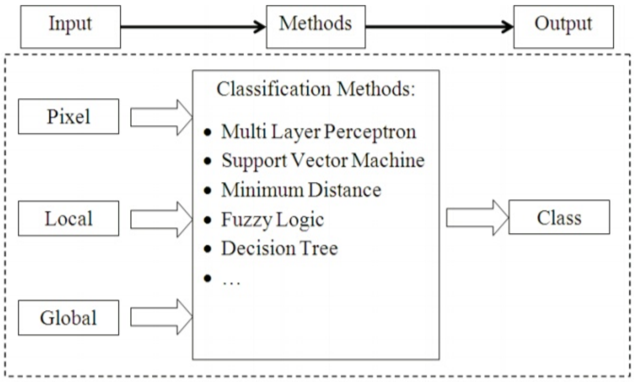

Figure 3.6: Supervised Classification

Data:

The data from LANDSAT Multispectral Scanner System.Characteristics of data can be divided into seven defined classes.One set of imagery from LANDSAT MSS consist of four spectral bands.

Band 4-Green

Band 5- Red region of visible spectrum

Band 6-near infra-red

Band 7-near infra-red spectrum

Each sample consisted of the pixel values in the four spectral bands of each of the 9 pixels in the 3*3 neighborhoods.

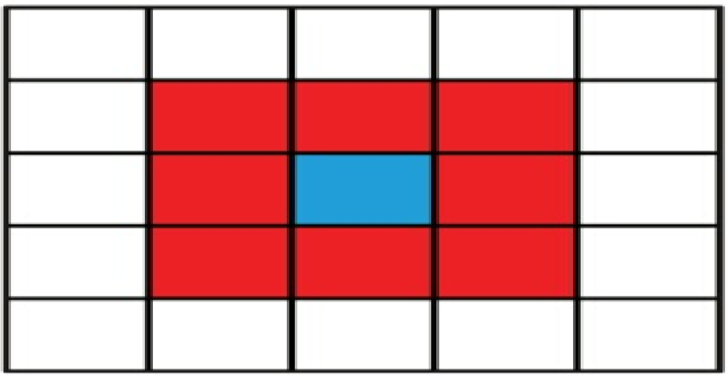

Figure 3.7: Neighborhood pixels

Characteristics of MSS data can be dividied into seven classes which overlap with each other as shown in below figure:

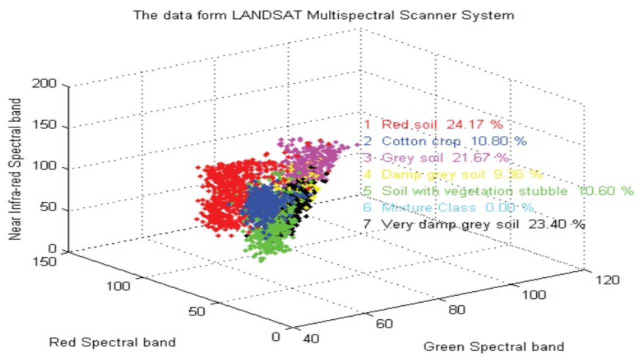

Figure 3.8: Characteristics of Data

Classification Method:

Support Vector Machine classification requires a process to train its model. Training requires some data points, known as training data, so sample data has been collected and divided into training data set and testing data set. Each record in the training dataset composes of 36 attributes, 9 pixels x 4 attributes/pixel. Characteristics of MSS data can be divided into 7 classes. The classification base on multiclass support vector machine.multiclass classification is the problem of classifying instances into one of three or more classes.

The basic SVM supports only binary classification, but extensions have been proposed to handle the multiclass classification case as well. In these extensions, additional parameters and constraints are added to the optimization problem to handle the separation of the different classes.Here, Radial Basis Function(RBF) kernel is used.Radial Basis Function (RBF) kernel has a parameter called sigma which control the shape of the function in higher dimensional space. In this experiment, value of sigma is varied between 1 and 2. The sample data is divided into a 2218 records training set and 2217 records test sets. In this study, 2 SVMs classification method are used:

1) SVM Standard:

In this work, there are 7 classes therfore 7 SVMs are needed for each class . For each SVM, a 1-label will correspond to output associate with that class and a 0-label is otherwise. In case more than one SVMs are set to 1-label, the final output for that sample is a random class among the 1-label SVM.

Target output	Output	Class						
		1	2	3	4	5	6	7
1	1	1	0	0	0	0	-	0
2	2	0	1	0	0	0	-	0
3	3	0	0	1	0	0	-	0
4	5	0	0	0	1	1	-	0
5	7	0	0	0	0	1	-	1
7	7	0	0	0	0	0	-	1

Figure 3.9: Classification with SVM Standard

2) SVM Max: This method solve an uncertainty in case the output of SVMs with 1-label has more than one machine. The raw classfication value are used to determine the final class rather than a random among the 1-label SVM. The SVM with the highest raw classification value is selected as the winning SVM and thus the output will be set accordingly.

Target output	Output	Class						
		1	2	3	4	5	6	7
1	1	1.17	-1.13	-1.66	-1.47	-0.96	-	-1.45
2	2	-1.09	1.02	-0.96	-1.01	-1.04	-	-0.98
3	3	-1.78	-1.20	1.41	-0.79	-1.31	-	-0.94
4	4	-1.13	-1.27	-0.85	0.65	-1.13	-	-0.08
5	5	-0.98	-1.36	-1.20	-1.58	1.75	-	-1.82
7	7	-1.23	-0.89	0.41	-1.51	-0.86	-	0.69

Figure 3.10: Classification with SVM Max

3.8 Results

Support Vector Machine classification is performed over the satellite image with various kernel functions. Here tried 4 variations of kernel functions (Linear, Quadratic, Polynomial, Radial Basis Function (RBF) kernel). The preliminary result shows that the RBF kernel has the highest accuracy.The main aim is to continually experiment to find the optimal sigma parameter of RBF kernel, value between 1.0 and 2.0 is varied to test their accuracy.

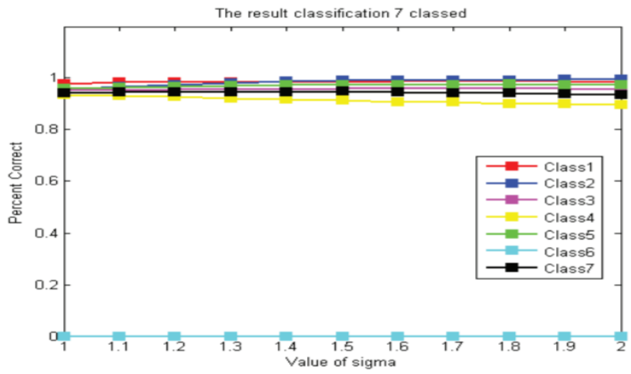

Figure 3.11: Classification Result

Class	Description	Percentage in the image	Classification Correct Rate
1	Red soil	24.17 %	98.21 %
2	Cotton crop	10.80 %	98.19 %
3	Grey soil	21.67 %	96.43 %
4	Damp grey soil	9.36 %	91.27 %
5	Soil with vegetation stubble	10.60 %	97.16 %
6	Mix Class	0.00 %	-
7	Very damp grey soil	23.40 %	94.32 %

Figure 3.12: Classification Correct Rate

This research focuses on Radial Basis Function (RBF) kernel determining the different parameters sigma value (1.0, 1.1, 1.2, 1.3, 1.4, 1.5, 1.6, 1.7, 1.8, 1.9, 2.0) and set parameters to optimize in each experiment were encoded in a vector, bound to maximum values. The sigma value at 1. 7 of multiclass maximum value gives the highest correct classification rate of 90.89case of the multiclass standard as shown in below figure.The accuracy is higher when the sigma value is approaching 1.7.

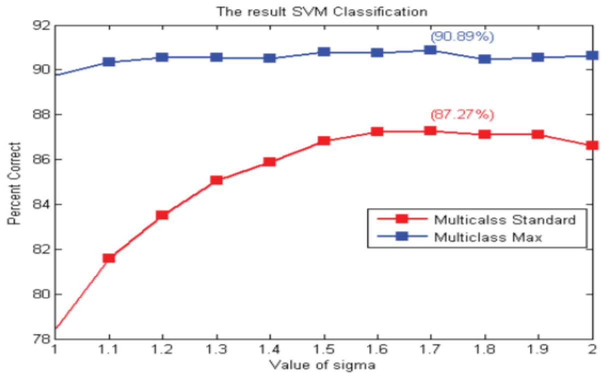

Figure 3.13: SVM Classification

Obtained classification result is compared with other classification methods like MLP 4 attribute ,MLP 36 attribute, PCA, SVM Standard, SVM Max.

The MLP 4 attribute is a Multilayer Perceptron with 4 values.The MLP 36 attribute is a Multilayer Perceptron with a set of 3*3 neighborhoods each pixel has 4 attribute [12].SVM Max and standard use 36 attribute datasets.

The result of MLP 4 attribute gives a correct classification rate of 87.25% but MLP 36 attribute of 89.32%

The result of SVM standard and Max of 87.27% and 90.89%.

The result from this SVM method has a higher accuracy compared to other methods

CHAPTER 4
ADVANTAGES AND DISADVANTAGES

4.1 ADVANTAGES

1. The technique provides better search accuracy for image classification.
2. It is relatively memory efficient
3. SVM is effective in cases where the number of dimensions is greater than the number of samples.
4. SVM works relatively well when there is a clear margin of separation between classes.
5. Automatically classify the land use information from satellite images using support vector machine.
6. improves speed
7. improves accuracy
8. Reduces satellite classification cost

4.2 DISADVANTAGES

1. SVM algorithm is not suitable for large data sets because the required training time is higher.
2. it also doesn't perform very well,when the data set has more noise ie.target classes are overlapping.

CHAPTER 5
REAL WORLD APPLICATIONS

5.1 APPLICATIONS

1. Bioinformatics - it includes protein and cancer classification.it uses SVM for identifying the classification of genes,patients on the basis of genes and other biological problems.
2. Face detection - SVM classify parts of the image as a face and non-face create a square boundary around the face.
3. Text and Hypertext categorization - it uses training data to classify documents into different categories such as news artcles,e-mails,and webpages.

CHAPTER 6
CONCLUSION

Developed an integrated model of Support Vector Machine for the land use classification of satellite images. SVM technique is an efficient image representation in the classification task.In this paper there are two main stages,the first is the training set and the second is the testing stage,every stage has number of steps.In general,this paper presents a method for the classification of Landsat Multi-Spectral Scanner(MSS) satellite images to identify the areas of land use.The image is preprocessed and classified using SVM with the Radial Basis Function(RBF)Kernel.The work consists of two main stages.At the first stage,the optimal parameter,sigma value of RBF kernel,for SVM is studied.At the second stage,the obtained classification result is compared with other classification methods.In this study,SVM is applied to classify a spatial data from satellite image. The result from simulation shows that a RBF kernel with a sigma value of 1.7 gives a higher correct classification rate.The system resulted in an accuracy of 90% in image classification.The comparison among many classification methods shows that multiclass SVM classification method gives a higher correct classification rate than Multilayer Perceptron and PCA[12].hhh Satellite image classification is a field which has great significance for different socioeconomic, environmental applications. Through classification of satellite imagery, the information as cadastral information land cover and land use type, vegetation type, soil properties could be obtained.

REFERENCES

[1] Sunitha Abburu and Suresh Babu Golla. Satellite image classification methods and techniques: A review. *International journal of computer applications*, 119(8), 2015.

[2] R Ahmed, Z Mourad, BH Ahmed, and B Mohamed. An optimal unsupervised satellite image segmentation approach based on pearson system and k-means clustering algorithm initialization. *Int. Sci. Index*, 3(11):948–955, 2009.

[3] FS Al-Ahmadi, AS Hames, et al. Comparison of four classification methods to extract land use and land cover from raw satellite images for some remote arid areas, kingdom of saudi arabia. *Earth*, 20(1):167–191, 2009.

[4] Wanessa da Silva, Matheus Habermann, Elcio Hideiti Shiguemori, Leidiane do Livramento Andrade, and Ruy Morgado de Castro. Multispectral image classification using multilayer perceptron and principal components analysis. In *2013 BRICS Congress on Computational Intelligence and 11th Brazilian Congress on Computational Intelligence*, pages 557–562. IEEE, 2013.

[5] Anita Dixit, Nagaratna Hedge, and B Eswar Reddy. Texture feature based satellite image classification scheme using svm. *International Journal of Applied Engineering Research*, 12(13):3996–4003, 2017.

[6] Moisés Espínola, José A Piedra-Fernández, Rosa Ayala, Luis Iribarne, and James Z Wang. Contextual and hierarchical classification of satellite images based on cellular automata. *IEEE Transactions on Geoscience and Remote Sensing*, 53(2):795–809, 2014.

[7] James D Hurd, Daniel L Civco, Martha S Gilmore, Sandy Prisloe, and Emily H Wilson. Tidal wetland classification from landsat imagery using an integrated pixel-based and object-based classification approach. In *American Society for Photogrammetry and Remote Sensing, 2006 Annual Conference, Reno, Nevada*, 2006.

[8] Shabnam Jabari and Yun Zhang. Very high resolution satellite image classification using fuzzy rule-based systems. *Algorithms*, 6(4):762–781, 2013.

[9] Nur Anis Mahmon and Norsuzila Ya'acob. A review on classification of satellite image using artificial neural network (ann). In *2014 IEEE 5th Control and System Graduate Research Colloquium*, pages 153–157. IEEE, 2014.

[10] Thao Nguyen, Jiho Han, and Dong-Chul Park. Satellite image classification using convolutional learning. In *AIP Conference Proceedings*, volume 1558, pages 2237–2240. American Institute of Physics, 2013.

[11] Muhammad Shahbaz, Aziz Guergachi, Aneela Noreen, and Muhamad Shaheen. Classification by object recognition in satellite images by using data mining. In *Proceedings of the World Congress on Engineering*, volume 1, pages 4–6, 2012.

[12] Kanita Tangthaikwan and Narongdech Keeratipranon. Multilayer perceptron neural net-

work for classification spatial data from satellite image. In *2015 38th Electrical Engineering Conference*, pages 842–845, 2015.

[13] Kanita Tangthaikwan, Narongdech Keeratipranon, and Adshariya Agsornintara. Multi-class support vector machine for classification spatial data from satellite image. In *2017 9th International Conference on Knowledge and Smart Technology (KST)*, pages 111–115. IEEE, 2017.