J S S Mahavidyapeetha

Sri Jayachamarajendra College of Engineering (SJCE), Mysore –570006

An Autonomous Institute Affiliated to

Visvesvaraya Technological University, Belgaum

"LUCI : Land use change analysis of Indian subcontinent using ML"

Thesis submitted in partial fulfillment of curriculum prescribed for the award of the degree of

BACHELOR OF ENGINEERING

IN

COMPUTER SCIENCE AND ENGINEERING

by

Anusha Gururaja M
(4JC14CS014)

Avinash Kulkarni
(4JC14CS021)

Sudhanva Channagiri
(4JC14CS031)

Under the Guidance of

Prof S.Brunda
Assistant Professor,
Department of Computer Science & Engineering,
SJCE, Mysore

Under

Indian Space Research Organisation

DEPARTMENT OF COMPUTER SCIENCE & ENGINEERING
April 2018

J S S Mahavidyapeetha

Sri Jayachamarajendra College of Engineering (SJCE), Mysore –570006

An Autonomous Institute Affiliated to
Visvesvaraya Technological University, Belgaum

CERTIFICATE

This is to certify that the work entitled **"LUCI : Land use change analysis of Indian subcontinent using Machine learning"** is a bonafide work carried out by **Anusha Gururaja M, Sudhanva Channagiri , Avinash Kulkarni** in partial fulfillment of the award of the degree of **Bachelor of Engineering in Computer Science and Engineering of Visvesvaraya Technological University, Belgaum during the year 2017-18**. It is certified that all corrections / suggestions indicated during CIE have been incorporated in the report. The project report has been approved as it satisfies the academic requirements in respect of project work prescribed for the Bachelor of Engineering degree.

Guide
Prof S.Brunda
Assistant Professor,
Dept. of CS & E,
S.J.C.E, Mysore

Head of the Department
Dr. H.C.Vijayalakshmi
Head, Dept. of CS & E,
S.J.C.E, Mysore

Place : Mysuru
Date : 27 / 04 / 2018

Company Certificate

16/01/2018 15:37 9108023411474 OSD ISRO HQ PAGE 01/01

भारतीय अन्तरिक्ष अनुसंधान संगठन
अन्तरिक्ष विभाग
भारत सरकार
अन्तरिक्ष भवन
न्यू बी ई एल रोड, बेंगलूर – 560 231, भारत
दूरभाष : +91 80 2341 5474
फैक्स :

इसरो isro

Indian Space Research Organisation
Department of Space
Government of India
Antariksh Bhavan
New BEL Road, Bangalore - 560 231, India
Telphone : +91 80 2341 5474
Fax :

सं/No.मु.HQ:प्रशा.ADMN:GEN:20

जनवरी/ January 13, 2018

Dr H C Vijayalakshmi
Head of the Department
Department of Computer Science & Engineering
Sri Jayachamarajendra College of Engineering
JSS Technical Institutions Campus
Mysore 570 006

महोदय/Madam,

विषय/Subject: **Permission for the students to carry out project work at ISRO HQ - regarding**

****

I am directed to refer to your letter dated 22/8/2017, addressed to Scientific Secretary, ISRO, requesting for permission to carry out project work at ISRO HQ, Bangalore by the following students of your esteemed institution:-

(1) Ms Anusha Gururaja M
(2) Ms Sudhanva Ch··nagiri
(3) Shri Avinash Kulk··i

2. The above students are permitted to carry out the project work at ISRO HQ from 15/1/2018 to April 2018. Shri T Shamsudheen, Scientist/Engineer SE, EOS, ISRO HQ has been identified as the focal point for this purpose.

3. The students may be advised to report to Shri T Shamsudheen at 9.30 AM on 16/1/2018 at ISRO HQ, Antariksh Bhavan, Bangalore. The students may be advised to bring two latest colour photographs for issuing identity cards for smooth movement in the office premises. In case the students carry laptop, mobile phone or any other gadgets, separate permission is required to be obtained for carrying the items inside office premises

भवदीय/Yours sincerely,

(बी अनिल कुमार)/(B Anil Kumar)
व.प्र.का.एवं.सा.प्र/Sr Head, P&GA

Copy to
1) Shri T Shamsudheen, S .. Engineer 'SE', EOS/ISRO HQ

2) Sr. Admin. Officer (GA), ·SRO HQ

Declaration

We, the students of the Department of Computer Science and Engineering from Sri Jayachamarajendra College of Engineering Mysuru, hereby declare that this dissertation entitled "LUCI : Land use change analysis of Indian subcontinent using machine learning" has been independently carried out by us under the guidance of Dr H.C Vijayalakshmi, Head of Department, Department of Computer Science and Engineering, S Brunda, Assistant Professor, Sri Jayachamarajendra College Of Engineering, Mysuru, Dr P G Diwakar, Scientific Secretary, and our mentor, Mr T Shamsudheen, Earth Observation Office, Indian Space Research Organisation, Bengaluru, in partial fulfilment of the requirements for the award of the degree of Bachelor of Engineering in Computer Science and Engineering, Visvesvaraya Technological University, Belgaum. We also declare that we have not submitted this dissertation to any other University for the award of any degree.

Declarants:

Anusha Gururaja M (4JC14CS014)

Avinash Kulkarni (4JC14CS021)

Sudhanva Channagiri (4JC14CS031)

Abstract

The availability of high-resolution remote sensing data has opened up the possibility for interesting applications, such as feature classification of satellite images. This exploration uses National Remote Sensing Centre, India (NRSC), obtained remotely sensed data. An unsupervised approach is considered at first after which various design choices of the neural network architecture are evaluated and analyzed. The ROIs (regions of interest) selected, comprise of various classes which include various terrains. The results show that neural networks are a viable tool for solving segmentation tasks in the area of remote sensing.

Acknowledgement

The satisfaction that accompanies the successful completion of any task would be incomplete without the mention of people who made it possible. Success is the epitome of hard work and perseverance, but steadfast of all encouragement and guidance.

We acknowledge and express sincere thanks to **Dr T.N Nagabhushan**, **Principal**, SJCE , Mysuru for having supported in my academic endeavours and also providing us with all the facilities at the college premises.

We are grateful to **Dr H.C. Vijayalakshmi**, **Professor** and **Head of the Department of Computer Science and Engineering**, SJCE, Mysuru for her support and encouragement in facilitating the successful completion of this work.

We are grateful to our guide,**, S. Brunda**, **Assistant professor, Department of Computer Science and Engineering**, SJCE, Mysuru for her valuable guidance and the keen interest taken throughout the progress of our project.

We thank **Dr P G Diwakar**, **Scientific Secretary** and **Mr Thayada Shamsudheen**, our mentor at Indian Space Organisation Research for their guidance.

We also thank all the teaching and non-teaching staff of Department of Computer Science and Engineering, SJCE, Mysuru for their kind cooperation in the course of this work. We are thankful to all the family members and friends for their support in completion of this work.

Table of Contents

CHAPTER 1 - INTRODUCTION

Chapter 1 - Introduction

1.1 Introduction to Problem Domain

Over the past few decades, there has been a rapid change in land cover due to factors like urbanization and adverse climatic conditions. Changes in land use result from the complex interaction of many factors including policy, management, economics, culture, human behavior, and the environment. An understanding of how land use changes occur is critical since these anthropogenic processes can have broad impacts on the environment, altering hydrologic cycles, biogeochemical dynamics, size and arrangements of natural habitats such as forests and species diversity. Changes to land use can also affect local and regional economies. To maintain a balance in the environment, these effects have to be compensated. Planning and foresight will help predict the change in land cover for the coming years. This requires a land use / land cover map to be developed. The current methods used to map images of the Indian subcontinent using Indian Remote Sensing (IRS) data have the following problems.

- Image classification techniques are not standardized.
- The currently available techniques are mostly manual and obsolete.
- Tools are obsolete and done manually.
- Statistical techniques are used, which is a good starting step but not adequate.
- Lack of advanced and intelligent processing and analysis techniques (for example, machine learning).

This project aims to explore various machine learning techniques to develop a terrain map using the most suited algorithm to address this issue.

1.2 Aim/Statement of the problem

This project uses advanced machine learning techniques to better predict, analyze and track land use and land cover changes in the Indian subcontinent using satellite remote sensing data from Indian Space Research Organization (ISRO) to develop a thematic map depicting land cover.

1.3 Objectives of the project

The objectives are to

- Development of a thematic map indicating land cover in the Indian subcontinent.
- Comparison of various machine learning algorithms to achieve high accuracy in segmentation.
- Analyze change in land cover patterns over the study period.

1.4 Existing methods

The existing methods are briefed below.

1. The images were obtained from Landsat satellite. The unsupervised approach, ISODATA was applied on the Landsat images. The images were labelled. ISODATA algorithm was applied which gives the spectral distance of the pixels as well iteratively clusters the pixels including more number of classes depending on the heterogeneity of the land cover. These labels are used for supervised classifications, which were obtained from the ancillary data. All the segments of the images were classified and then mosaicked to form the whole.

2. Another existing approach uses the image differencing algorithm. The images were obtained from Landsat. The images were geo-coded, terrain corrected as well as resampled to 25-metre pixels. The image differencing algorithm compares two different files and implementation of simple GIS models. A single band difference file was created, which involved the change in pixel value, the line number and reflectance values in both the dates. The difference file was hence analyzed to obtain the changes in vegetation cover between dates.

3. In another approach, the Maximum likelihood algorithm, a common method in remote sensing owing to its robustness, was implemented to classify the images. The Maximum likelihood method is a supervised statistical approach to pattern recognition. It estimates the probability of a pixel belonging to each of a predefined set of labels and then assigns the pixel to the classes with highest probability. Maximum likelihood is based on the Bayesian approach.

1.5 Applications

Applications include :

- This geospatial information on patterns and trends in land cover and land use can be used for understanding, modeling, and projecting land change for urban planning and future development.
- This knowledge will help preserve environmental resources and balance the negative impacts of development such as urbanization.
- Once developed, this model will be able to classify terrain for any piece of land.
- Mapping of agricultural land will help predict the harvest for the coming year.

1.6 Proposed solution

The solution to this problem is evolutionary and is to be developed in a top down fashion, starting from abstract solutions to independent problem specific solutions.

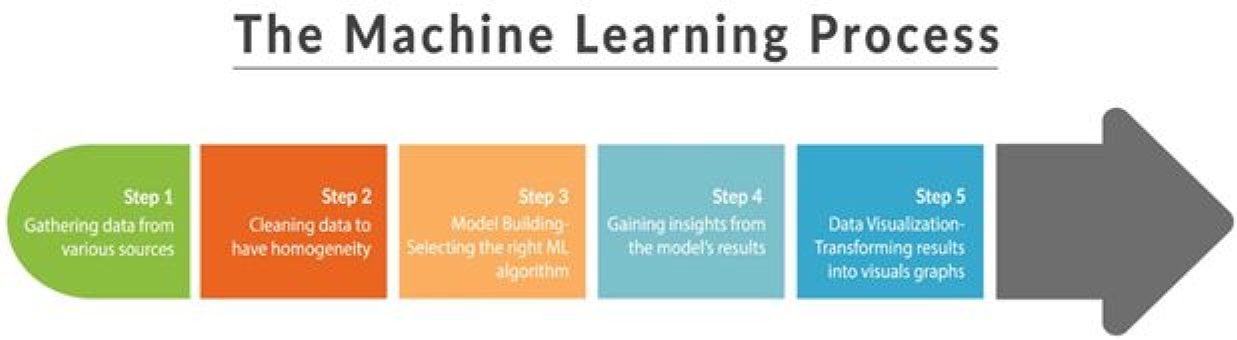


Figure 1 The machine learning process

The broad areas the solution to be addressed are

1.6.1 Data procurement

The dataset to be used is the remote sensing data provided by ISRO. These images are of those captured by the Resourcesat satellite launched by ISRO over the years. The images are raw and unlabeled. A task primary to data procurement is to identify and map the images to their respective geographic locations. This however depends on the availability of images over the span of the nation.

1.6.2 Preprocessing

The images procured are then subjected to an initial processing which makes them more suitable to work with in the further steps involved. This step is crucial since the accuracy and working of the system depends on how well we have understood the data and standardized it. Tasks in preprocessing include trimming the images to required shapes and sizes. Also the images must be normalized in format, number of bands and geotagged.

1.6.3 Labeling

The dataset of images that are preprocessed are unlabeled with respect to the classes the system is aimed at classifying into. While using this dataset fetches the required number of clusters taking the unsupervised approach, the accuracy can take a hit. Manually labeling these can influence the accuracy greatly. Therefore, the images are manually labeled. Since labeled satellite images are rarely available, these labeled images will be an added dataset to the machine learning community.

1.6.4 Segmentation & Classification

This is the critical step in the solution. The unlabeled images are first segmented into different classes using unsupervised clustering algorithms. A land cover/land use terrain map is then obtained by stitching the images together. Further, images of the same area at different points of time are classified and segmented.

1.6.5 Land use change analysis

Now that we have the classified image, we have used this as a prerequisite to analyze the change in land cover. This is important as it will help us keep track of how land cover has changed over the years. These images are compared and the percentage change in land cover is calculated. This can be further used to analyze the change in land cover over a period of time and determine the extent of change. This will help us balance out the impacts of environmental, ecological and urbanization changes. The percentage change in area was calculated.

1.7 Project Schedule

The project schedule is as follows :

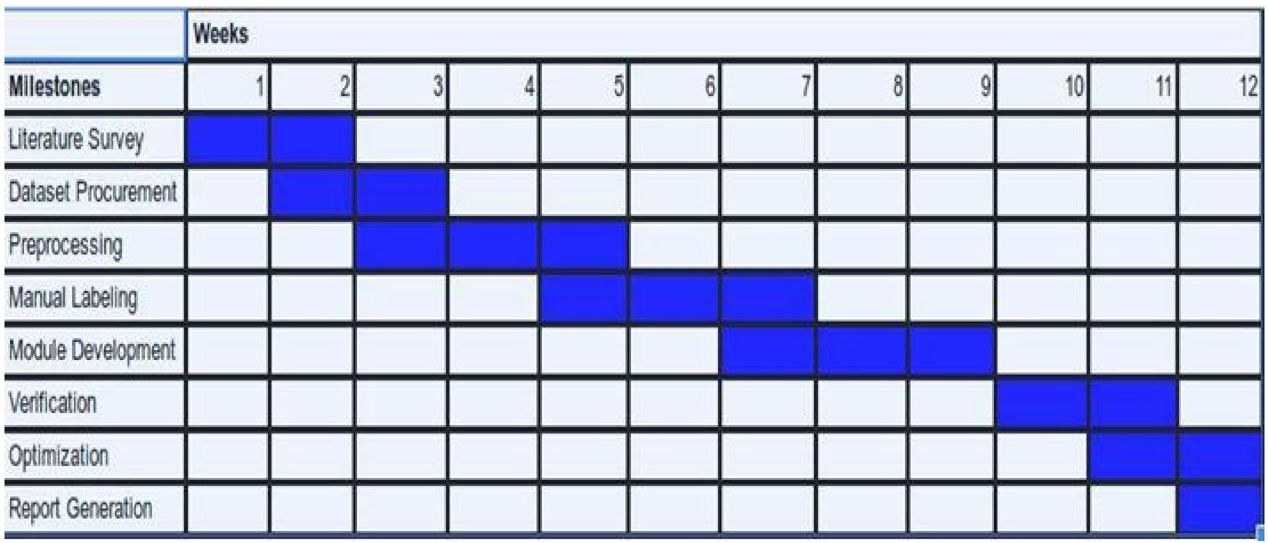


Figure 2 Project schedule

CHAPTER 2 - LITERATURE SURVEY

Chapter 2 - Literature survey

The work currently done in this area are as follows.

1. Land cover changes in Laka Hawassa watersheds. Authors Nigatu , Havard and Oystie.

In this paper we come through another method for change detection and land cover changes in Lake Hawassa Watershed, Southern Ethiopia, were investigated using Landsat MSS image data of 1973, and Landsat TM images of 1985, 1995, and 2011, covering a period of nearly four decades. Each image was partitioned in a GIS environment and classified using an unsupervised algorithm followed by a supervised classification method. A hybrid approach was employed in order to reduce spectral confusion due to high variability of land cover. Each segment of image data was classified using ISODATA allowing a large number of classes (15–30). Supervised classification was performed using a set of classes generated from unsupervised classification and training samples created. This operation was performed repeatedly until the required thematic map was produced.

2. Land change detection. Authors Green , Kempk and Lackey

The change detection was accomplished by comparing two Landsat TM scenes of two different days pertaining to different years(1987 and 1991). Several different types of algorithms are available for detecting change using multi-date imagery. The one used here is image differencing, which involves subtraction of one image date from another. For this project, a comparison was made of two separate different files and implementation of simple GIS models. One difference file was developed from subtraction of Band 7 for each date. A second file was created from subtraction of a vegetation index for each date composed of a ratio of Band 3 to Band 4. The results were compared and combined to take advantage of the individual strengths of each technique. The difference file was visually analysed to determine the relationships between different values and changes in vegetative cover between dates.

3. Satellite Image Classification Methods and Techniques. Author Sunitha Abburu.

The available satellite image classification algorithms were compared and results were obtained. The paper classified the major techniques available into 1) Manual 2) Automated and 3)Hybrid classes. The advantages and disadvantages with each method was discussed and conclusions drawn.

4. Techniques in Image Classification; Authors S.V.S.Prasad, Dr. T. Satya Savithri and Dr. Iyyanki V. Murali Krishna.

A study was performed on the issue of satellite image classification. The many issues addressed in this paper are remote sensing classification process, selection of remotely sensed data, selection of classification system and training samples, data preprocessing, selection of suitable classification method, use of multiple features of remote sensed data. This paper mainly highlights the brilliance of remotely sensed data. Since this data is multi-spectral, and is available in multiple formats, multiple features can be used to classify the raw satellite images into the required classes with high accuracy measures. This paper advises researchers to maximize the width of remotely sensed data, to obtain maximum accuracy in their use of multiple features of remote sensed data classifications.

CHAPTER 3 - SYSTEM REQUIREMENT AND ANALYSIS

Chapter 3 - System requirements and analysis

3.1 Hardware Requirements

The hardware requirements may serve as the basis for the implementation of the system and should therefore be a complete and consistent specification of the whole system. They are used by software engineers as the starting point for the system design. The recommended system specification is given below.

Processor : Intel Core 2 Duo , Pentium IV 2.6 Ghz.
RAM : 8GB HDD.
HDD : 40GB.

3.2 Software Requirements

The software requirements document is the specification of the system. It should include both a definition and a specification of requirements. It is useful in estimating cost, planning team activities, performing tasks and tracking the team's and tracking the team's progress throughout the development activity.

Operating system : Windows 10 / Ubuntu
Language : Python
IDE : Spyder / Ubuntu environment

CHAPTER 4 - TOOLS AND TECHNOLOGY USED

Chapter 4 - Tools and Technologies used

4.1 OpenCV

OpenCV (*Open Source Computer Vision*) is a library of programming functions mainly aimed at real-time computer vision. Originally developed by Intel , it was later supported by Willow Garage then Itseez which was later acquired by Intel. The library is cross-platform and free for use under the open-source BSD license. OpenCV (Open Source Computer Vision Library) is an open source computer vision and machine learning software library. OpenCV was built to provide a common infrastructure for computer vision applications and to accelerate the use of machine perception in the commercial products.

4.2 GIMP (GNU Image Manipulation Program)

GIMP is a multi-platform photo manipulation tool. GIMP is an acronym for GNU Image Manipulation Program. The GIMP is suitable for a variety of image manipulation tasks, including photo retouching, image composition, and image construction. GIMP has many capabilities. It can be used as a simple paint program, an expert quality photo retouching program, an online batch processing system, a mass production image renderer, an image format converter, etc. GIMP is expandable and extensible. It is designed to be augmented with plugins and extensions to do just about anything. The advanced scripting interface allows everything from the simplest task to the most complex image manipulation procedures to be easily scripted.

4.3 ImageMagick

ImageMagick is used to create, edit, compose, or convert bitmap images. It can read and write images in a variety of formats (over 200) including PNG, JPEG, GIF, HEIC, TIFF, DPX, EXR, WebP, Postscript, PDF, and SVG. Use ImageMagick to resize, flip, mirror, rotate, distort, shear and transform images, adjust image colors, apply various special effects, or draw text, lines, polygons, ellipses and Bézier curves. The functionality of ImageMagick is typically utilized from the command-line or features from programs on other languages. With a language interface, we can use ImageMagick to modify or create images dynamically and automatically. ImageMagick utilizes multiple computational threads to increase performance and can read, process, or write mega-, giga-, or tera-pixel image sizes.

4.4 Python Libraries

4.4.1 Keras

Keras is a minimalist Python library for deep learning that can run on top of Theano or TensorFlow. It was developed to make implementing deep learning models as fast and easy as possible for research and development. It runs on Python 2.7 or 3.5 and can seamlessly execute on GPUs and CPUs given the underlying frameworks. It is released under the permissive MIT license.

4.4.2 NumPy

NumPy is an acronym for "Numeric Python" or "Numerical Python". It is an open source extension module for Python, which provides fast precompiled functions for mathematical and numerical routines. Furthermore, NumPy enriches the programming language Python with powerful data structures for efficient computation of multi-dimensional arrays and matrices. The implementation is even aiming at huge matrices and arrays. Besides that the module supplies a large library of high-level mathematical functions to operate on these matrices and arrays.

4.4.3 Pandas

Pandas is a Python package providing fast, flexible, and expressive data structures designed to make working with "relational" or "labeled" data both easy and intuitive. It aims to be the fundamental high-level building block for doing practical, real world data analysis in Python. Additionally, it has the broader goal of becoming the most powerful and flexible open source data analysis / manipulation tool available in any language. It is already well on its way toward this goal. Pandas is well suited for many different kinds of data. The data actually need not be labeled at all to be placed into a pandas data structure.

4.4.4 Scikit

Scikit-learn was initially developed by David Cournapeau as a Google summer of code project in 2007. Scikit-learn provides a range of supervised and unsupervised learning algorithms via a consistent interface in Python.The library is built upon the SciPy (Scientific Python) that must be installed before you can use scikit-learn. This stack that includes

- NumPy
- SciPy
- Matplotlib
- Pandas

CHAPTER 5 - SYSTEM DESIGN

Chapter 5 - System Design

5.1 Dataset

The sample images were obtained from Resourcesat - 1 AWIFS satellite. The details of the sample images obtained are as follows

- The images were taken from the BHUVAN NRSC website. The images are from the Resourcesat satellite and from the AWiFS (Advanced Wide Field Sensor) camera. The images are in grayscale mode with each image having 4 bands. The four bands present are namely Red , Green , NIR (Near Infrared) , SWIR (Short wave infrared). The images are in .tif format. The spatial resolution being 56m and swath being 570km. The size of the image file is upto 16 Mb per band.

- Associated Xml files include the following details
  - Swath.
  - Resolution.
  - Location (With all Coordinates)
  - Image format.
  - Ownership
  - Camera

5.2 Pre-Processing

The pre-processing steps are given below.

- **Noise reduction** Gaussian noise reduction was applied but this leads to loss of valuable pixel information. Since we are using pixel value accuracy for segmentation and classification , this low will lead into loss in accuracy.

- **Normalisation** The images were normalised for clear visual perception of the structure and to differentiate between the terrain. ImageMagick and GIMP tools were used.

- **Adding false color** For initial processing and visual identification of different terrain, false color has to be added since there are no RGB bands available. Since no structure could be seen from the images , false color had to be added.

5.3 Algorithms used

5.3.1 K-Means

K-means clustering is a method of vector quantization that is popular for cluster analysis in data mining. k-means clustering aims to partition *n* observations into *k* clusters in which each observation belongs to the cluster with the nearest mean, serving as a prototype of the cluster. The algorithm works iteratively to assign each data point to one of *K* groups based on the features that are provided. Data points are clustered based on feature similarity. The results of the *K*-means clustering algorithm are:

1. The centroids of the *K* clusters, which can be used to label new data.
2. Labels for the training data (each data point is assigned to a single cluster).

Rather than defining groups before looking at the data, clustering allows you to find and analyze the groups that have formed organically. The "Choosing K" section below describes how the number of groups can be determined. Each centroid of a cluster is a collection of feature values which define the resulting groups. Examining the centroid feature weights can be used to qualitatively interpret what kind of group each cluster represents. The false color added images are processed by k-means clustering.

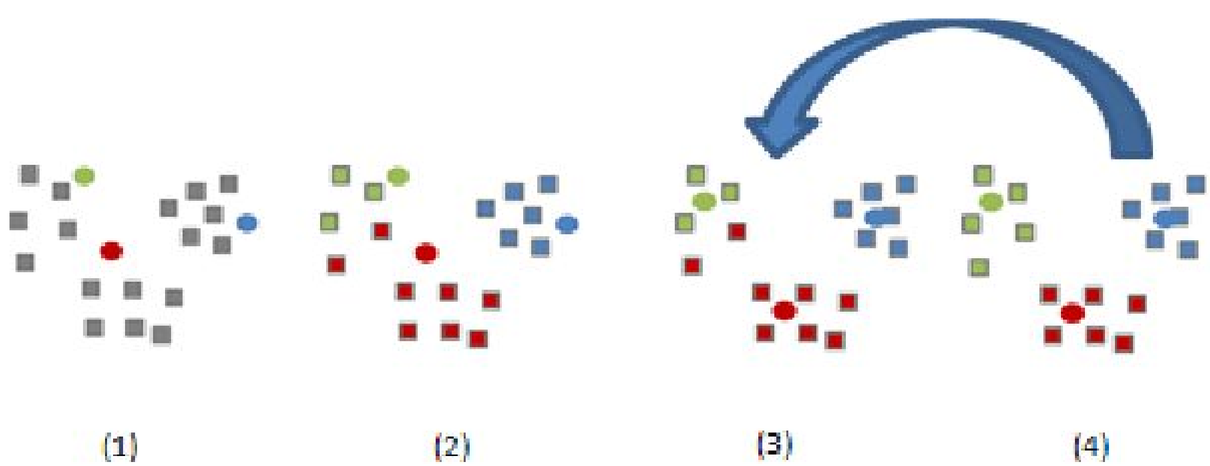


Figure 3 Process of k-means clustering

5.3.2 Neural networks

Neural networks are a set of algorithms modeled loosely after the human brain, that are designed to recognize patterns. They interpret sensory data through a kind of machine perception, labeling or clustering raw input. The patterns they recognize are numerical, contained in vectors, into which all real-world data, be it images, sound, text or time series, must be translated. Neural networks model helps us classify a given piece of land into any of the terrain types correctly.

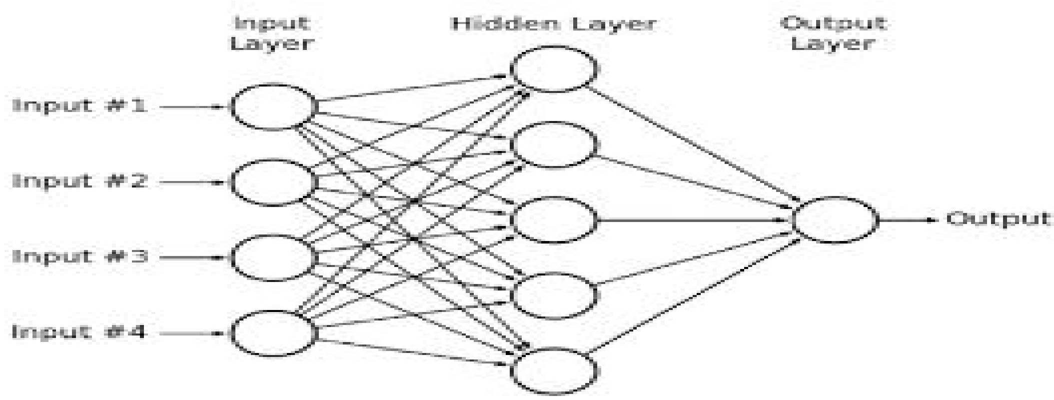


Figure 4 Multilayer perceptron

5.4 Land use change analysis

Once we obtain the classification output, we can use the generated images to analyze the change in land cover across a period of time. This can be accomplished by generating images for consecutive years in order to observe the change in land cover by subtracting images. This will help in developing a predictive model which considers various parameters such as climate change, ecological changes etc, to predict the percentage increase / decrease in area of landcover. Having this knowledge will help us with sustainable development and prepare for the future.

CHAPTER 6 - SYSTEM IMPLEMENTATION

Chapter 6 - System implementation

6.1 Dataset

The dataset was procured directly from ISRO HQ and also from the BHUVAN portal, which is India's gateway to earth observation. The images were orthorectified, but unlabeled.

6.2 Pre-Processing

The images were initially normalised using Imagemagick and GIMP tools to see the structure. For better visual perception, OpenCV along with python was used to add false color to the images. The false color is added to differentiate between terrains since there are no RGB bands available. Among the various color maps available, Rainbow Jet color map seemed to have the largest variation in the color and hence makes it easier to differentiate between terrains. The next step would be to label the images. The region of interest (ROI) is obtained by using the mouse click events to draw the region of interest using the mouse. The same region of interest is obtained for all the 4 bands for each image. The intensity values for each terrain is considered for manual labelling. The intensity values for each terrain is obtained using GIMP tool. The intensity values vary from 0 to 255. T the range of values in the ROI for each of the bands and for multiple images are noted down. The results for each terrain is taken by compiling the ranges obtained for multiple images and a common range is obtained.

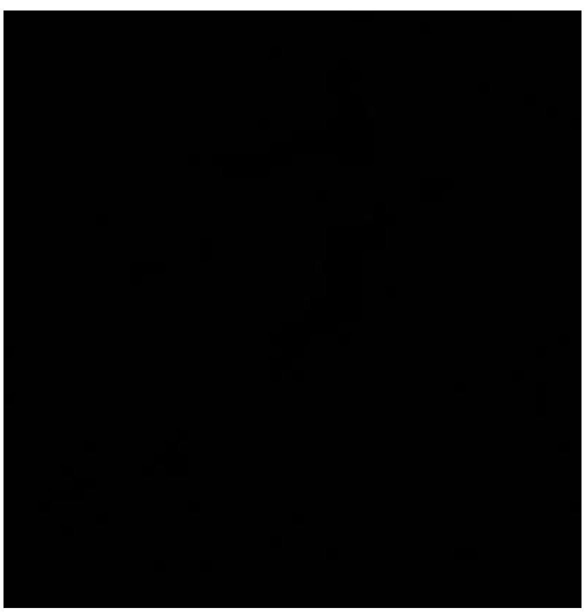


Figure 5 : Obtained raw image

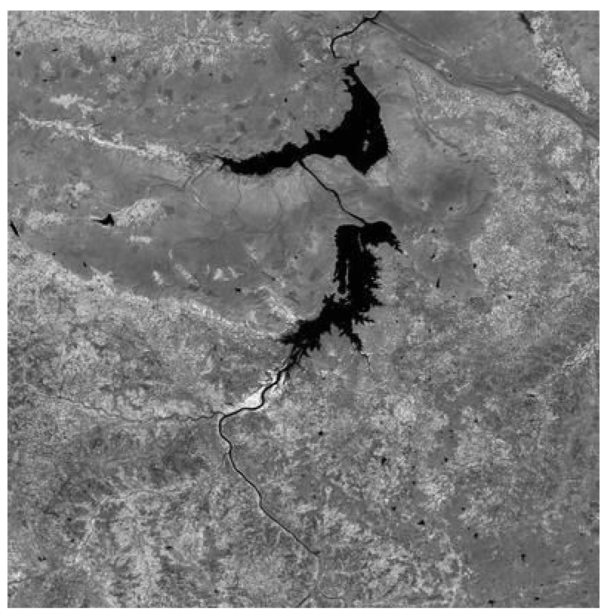


Figure 6 : Normalised image

Figure 7: Before adding false color

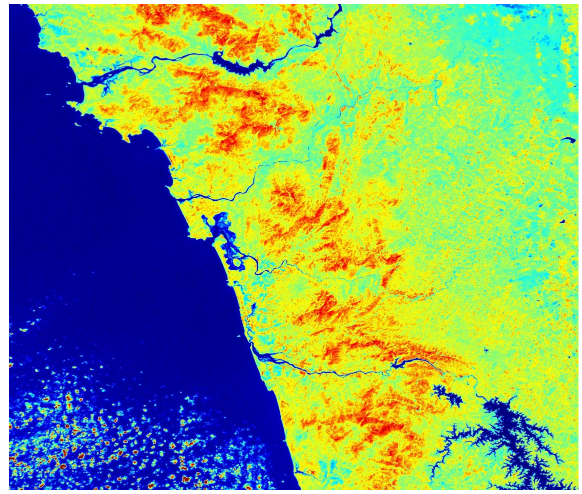


Figure 8 : After adding false color

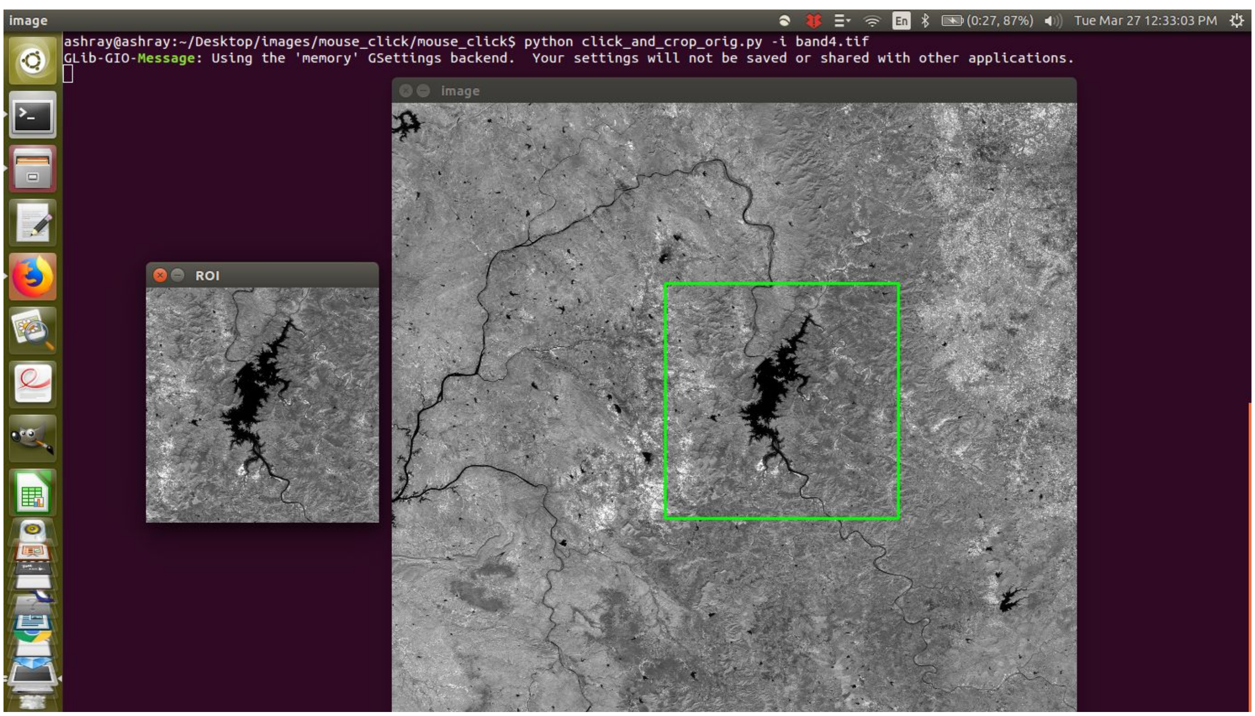


Figure 9 :   Extracting ROI

6.3 Implementation

- Firstly, an unsupervised machine learning algorithm was used to segment the images into different terrains. The K-Means algorithm was considered initially as the images are unlabeled. The algorithm is designed such that the number of clusters can be changed by the user through the terminal and without accessing the code. K-means is fed with the false color added images without any labels and clusters are formed. The images are stitched to obtain a thematic map. The obtained segmented images are compared with the ground truth for testing purposes.

- For better accuracy and control over the output of the model and also to achieve the objective of comparing various machine learning algorithms, we consider a supervised approach as well. Neural networks were used to obtain a pixel level classification of the satellite images.

- The intensity ranges for each chosen terrain over the 4 bands for each image is considered for multiple images. The ranges are studied and compiled. This is further used to label the images pixel-wise. The labels are encoded such that the labels consist of 0s and 1s only. The pixel values and labels to be used to train the neural network model are stored as objects.

- The next step is to design the neural network model and optimize the model to achieve best accuracy. The model is developed and implemented in python using various libraries.

- For the post processing stage, the labels obtained as output from the keras model are used to add create a segmented RGB image.

6.4 Land use change analysis

We use the generated images to analyse the change in land cover over consecutive years

- Firstly, we create the images using the output labels generated by the neural network model.
- The images considered belong to the same grid and are of consecutive years.
- The images were calculate the area of each type of terrain in a single image.
- The percentage increase / decrease was calculated.
- A timelapse of images was created to view the change in land cover over time.

CHAPTER 7 - SYSTEM TESTING AND RESULT ANALYSIS

Chapter - 7 System testing and result analysis

To test the output of the K-Means algorithm, the segmented image is compared with the ground truth.

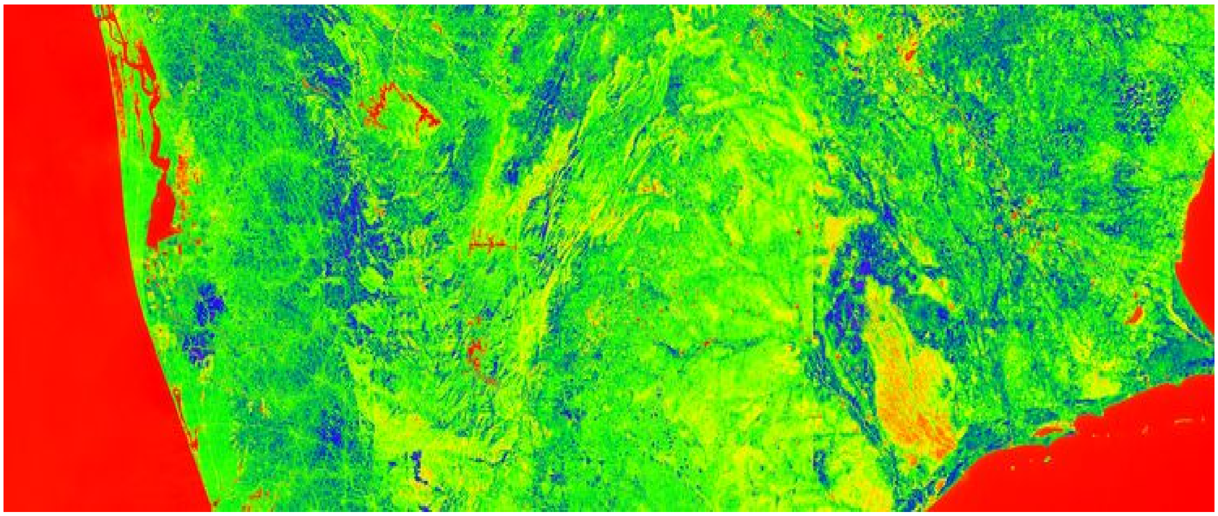


Figure 10 : K-means stitched image.

The labels generated by the neural network model are used to build the classified image which is further compared with the ground truth for testing.

Figure 11 : Normalised image(Ground truth)

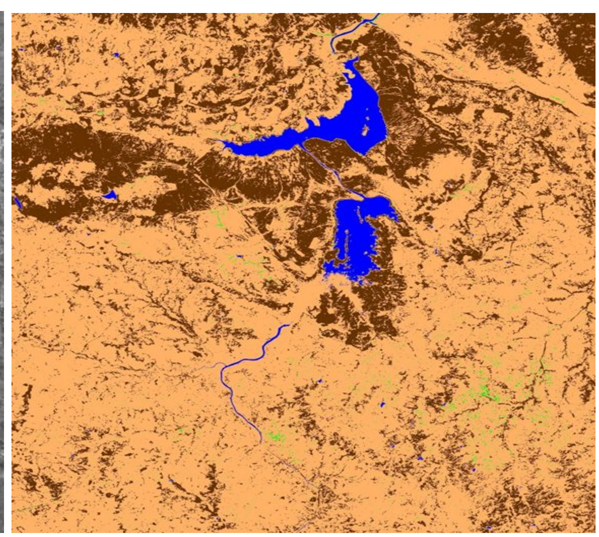


Figure 12 : Image generated by NN model

The output generated by the neural network model for the analysis of land use/land change with the dates. The images refer to KRS dam. The **Blue** color in the images indicates **water ,** **Brown** indicates **Hilly or Desert areas** , **peach** indicates **land** , **Green** indicates **Greenery** , **White** indicates **Snow.**

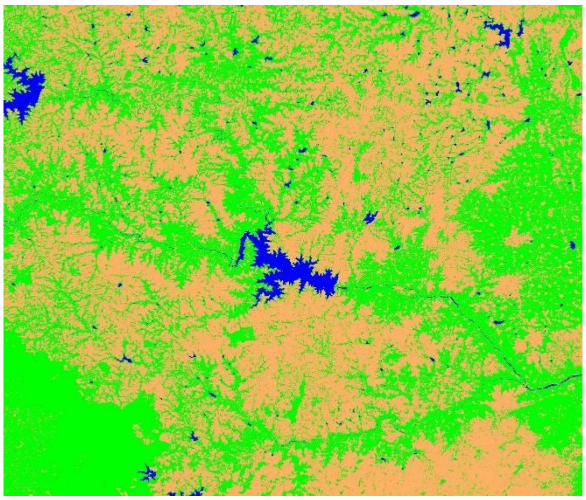


Figure 13 : Dated - 09/2007

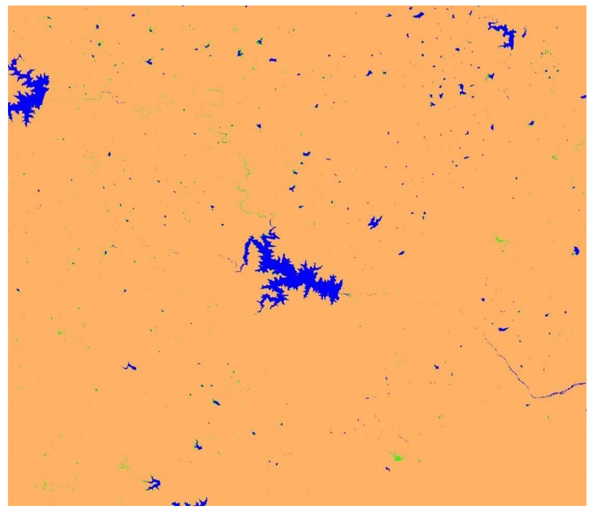


Figure 14: Dated - 09/2009

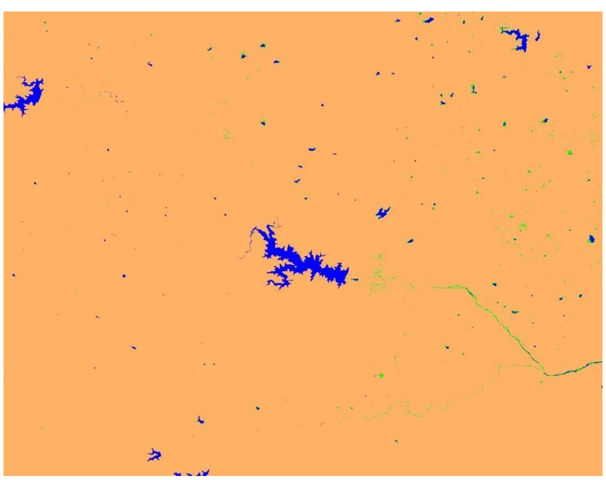


Figure 15 : Dated – 03/2012

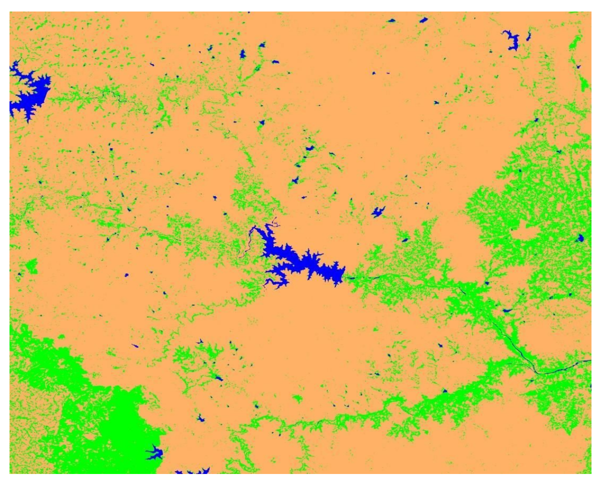


Figure 16 : Dated – 10/2012

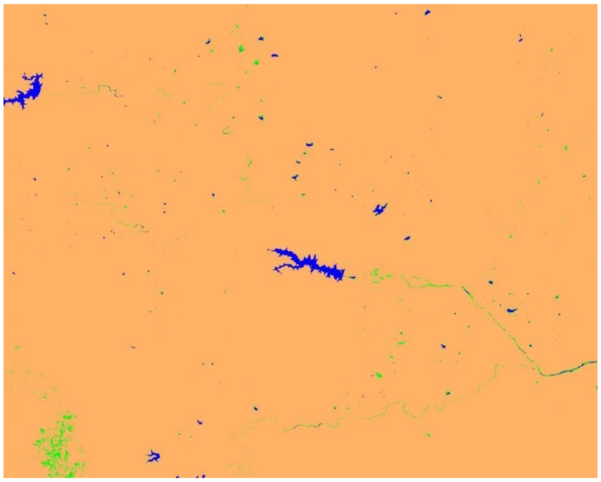


Figure 17 : Dated - 02/2013

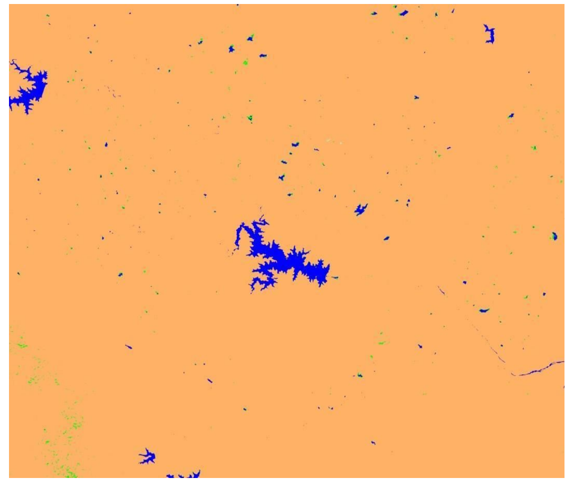


Figure 18 : Dated - 10/2013

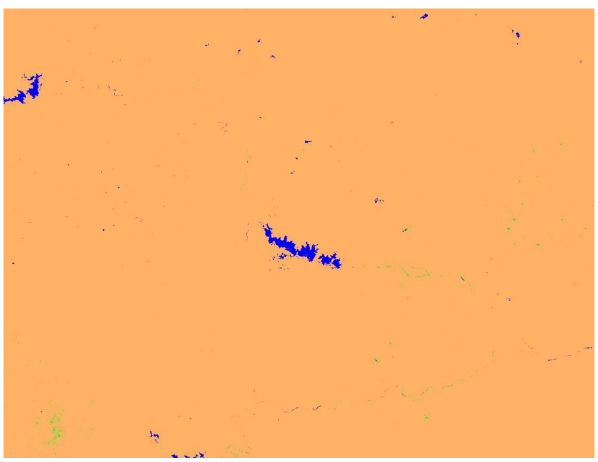


Figure 19 : Dated - 04/2014

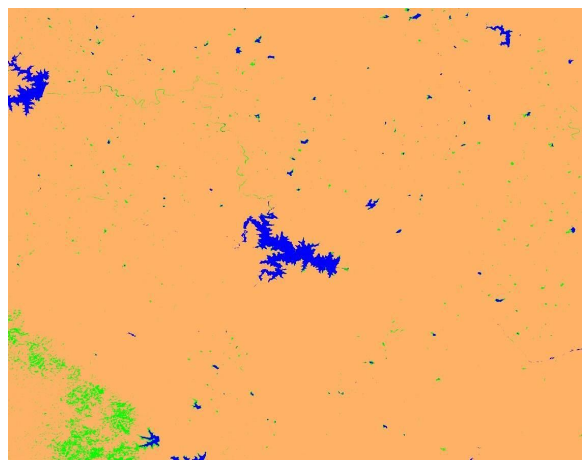


Figure 20 : Dated - 09/2014

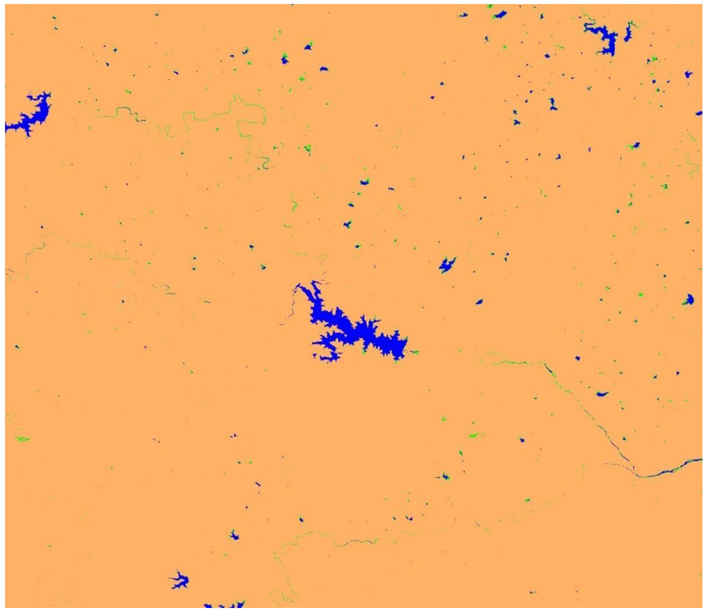


Figure 21 : Dated - 01/2016

CHAPTER 8 - CONCLUSION AND FUTURE WORK

Chapter - 8 Conclusion and future work

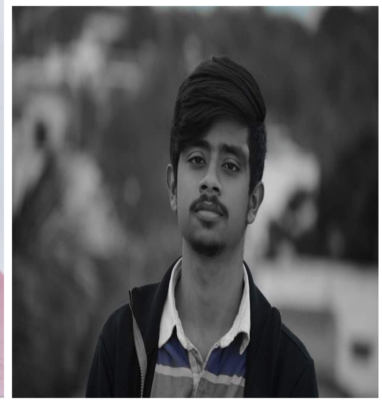
This project applies unsupervised as well as supervised classification technique to produce a thematic map covering the Indian Subcontinent. K-means approach was used to segment the unlabeled images. The labelled sets are then classified using the supervised classification. Neural network model was trained and developed to obtain the pixel level classification, the accuracy of the results were high since we were simultaneously comparing with the output ground truth. The terrain map is hence developed for the area.

The further work in this regard would be to approach the segmentation by considering the structure of the image.

Appendix A : Project Team Details

Project Title	**LUCI Analysis of land use change of Indian subcontinent using ML**		
USN	**Team Members Name & CGPA**	**E-Mail Id**	**Mobile Number**
4JC14CS014	Anusha Gururaja M	anusha.manur@gmail.com	7899842682
4JC14CS031	Sudhanva Channagiri	schannagiri@gmail.com	9538146315
4JC14CS021	Avinash Kulkarni	tammanni.avi@gmail.com	7760762745

Project Team Photograph:

Anusha (4JC14CS014)	Sudhanva (4JC14CS031)	Avinash (4JC14CS021)

Appendix B : COs, POs and PSOs Mapping

Course Outcomes

CO1: Formulate the problem definition, conduct literature review and apply requirements analysis.
CO2: Develop and implement algorithms for solving the problem formulated.
CO3: Comprehend, present and defend the results of exhaustive testing and explain the major findings.

Program Outcomes

PO1: Apply knowledge of computing, mathematics, science, and foundational engineering concepts to solve the computer engineering problems.

PO2: Identify, formulate and analyze complex engineering problems.

PO3: Plan, implement and evaluate a computer-based system to meet desired societal needs such as economic, environmental, political, healthcare and safety within realistic constraints.

PO4: Incorporate research methods to design and conduct experiments to investigate real-time problems, to analyze, interpret and provide feasible conclusion.

PO5: Propose innovative ideas and solutions using modern tools.

PO6 Apply computing knowledge to assess societal, health, safety, legal and cultural issues and the consequent responsibilities relevant to professional engineering practice.

PO7: Analyze the local and global impact of computing on individuals and organizations for sustainable development.

PO8: Adopt ethical principles and uphold the responsibilities and norms of computer engineering practice.

PO9: Work effectively as an individual and as a member or leader in diverse teams and in multidisciplinary domains.

PO10: Effectively communicate and comprehend.

PO11: Demonstrate and apply engineering knowledge and management principles to manage projects in multidisciplinary environments.

PO12: Recognize contemporary issues and adapt to technological changes for lifelong learning.

Program Specific Outcomes

PSO1: Problem Solving Skills Ability to apply standard practices and mathematical methodologies to solve computational tasks, model real world problems in the areas of database systems, system software, web technologies and Networking solutions with an appropriate knowledge of Data structures and Algorithms.

PSO2: Knowledge of Computer Systems An understanding of the structure and working of the computer systems with performance study of various computing architectures.

PSO3: Successful Career and Entrepreneurship The ability to get acquaintance with the state of the art software technologies leading to entrepreneurship and higher studies.

PSO4: Computing and Research Ability Ability to use knowledge in various domains to identify research gaps and to provide solution to new ideas leading to innovations.

SEM	SUBJECT	CODE	CO	PO1	PO2	PO3	PO4	PO5	PO6	PO7	PO8	PO9	PO10	PO11	PO12	PSO1	PSO2	PSO3	PSO4
VIII	Project Work	CS84P	CO1	1	1	3	2	2	3	3	1	2	3	1	3	3	3	2	3
			CO2	2	2	3	3	3	3	3	2	2	3	2	3	3	3	3	3
			CO3	2	2	3	2	2	3	3	3	2	3	1	3	3	3	2	3

We have considered various aspects of the project to match the course outcomes with the project outcomes.

1. Our problem statement falls under the domain of addressing real life problems using technology and we focus on the change on change in climate and its effect on land cover. Hence, this project has high relevance to using the concept of computing to solve real life technological issues.
2. We have also focused on different research methods and various machine learning algorithms to design and investigate real-time problems and to provide feasible conclusion.
3. During the course of the project, we have worked effectively as a team to understand the problem statement thoroughly and to come up with a solution.
4. We have used state of the art software technologies to implement the solution and also been able to identify the lack of an automated and intelligent method to map land cover and also provide a better solution.
5. Keras Deep Learning and SciKit were used to develop neural network models built after reducing the pixels from different bands of images and integrating them into making predictions for each pixel on an entirely new dataset (ResourceSat - AWIFS).
6. Since this was a standard pattern recognition problem, the major concepts of mathematical probabilistic estimation, and approximation was used to arrive at the solution.
7. Since machine learning is an ever growing field with various characteristics of terrains still unharnessed, there is still scope for increasing the accuracy of the predictions.

References

[1] A. W. Abbas , N . Minallh, N. Ahmad , S.A.R. Abid , M.A.A. Khan. "K-Means and ISODATA Clustering Algorithms for Land Cover Classification Using Remote Sensing". Sindh Univ. Res. Jour. (Sci. Ser.) Vol.48 (2) 315-318 (2016).
[2] Clustering - http://scikit-learn.org/stable/modules/clustering.html#k-means
[3] Kass Green, Dick Kempka, Lisa Lackey. "Using remote sensing to detect and monitor landuse and landcover change"
[4] Mr. S.V.S.Prasad, Dr. T. Satya Savithri , Dr. Iyyanki V. Murali Krishna. "Techniques in Image Classification A Survey ".
[5] Nigatu, Havard, Oystie. "Land cover changes in Lake Hawassa Watershed, Southern Ethiopia"
[6] Remote sensing : http://remote-sensing.eu/image-classification-with-python/
[7] Sci-kit documentation : http://scikit-image.org/docs/stable/api/api.html
[8] Sunitha Abburu,"Satellite Image Classification Methods and Techniques A Review". Dept. of Computer Applications Adhiyamaan College of Engineering Hosur, Tamilnadu, India.

List of Figures