LULC Mapping For Identifying Change In Agricultural Land Cover using Satellite Images

Prathamesh Santosh Raje
Department of Computer Science and Engineering
Walchand College Of Engineering, Sangli
Sangli, India
prathamesh.raje@walchandsangli.ac.in

Raj Dattatray Dalvi
Department of Computer Science and Engineering
Walchand College Of Engineering, Sangli
Sangli, India
raj.dalvi@walchandsangli.ac.in

Samrat Vishwas Jadhav
Department of Computer Scienece and Engineering
Walchand College Of Engineering, Sangli
Sangli, India
samrat.jadhav@walchandsangli.ac.in

Nitin Laxmanrao Gavankar
Department of Computer Science and Engineering
Walchand College Of Engineering, Sangli
Sangli, India
nitin.gavankar@walchandsangli.ac.in

***Abstract*—Remotely sensed images or satellite images has opened a new window for researchers to extract and analyze various land cover classes spread over the earth surface. This study proposes an enhanced Machine Learning (ML) based change detection method for the analysis of change in agricultural land cover along with other land cover classes such as, water bodies, built-up area and barren land. Machine learning algorithm employs classification of multi-spectral remote sensing data into different classes created during training of model. The study has been carried out by analyzing Geographical features of Sangli city and its surrounding area by using LISS III multispectral satellite images of year 2014 and year 2019, collected from ISRO Bhuvan web-portal. Satellite images are processed using GIS tools and spatial modelling has been implemented to automate the manual process of classification using ML algorithm. Further, the extracted classes of the classified images have been compared to analyze change in the land cover classes. After analysis it has been observed that vegetation which includes most of the agricultural land has been decreased by 13.40% and Barren land has been decreased by 25.36%. However, built up has been increased by 23.99% due to growing urbanization.**

***Keywords*—classification, satellite image, GIS, machine learning, spatial modelling**

I. INTRODUCTION

Feature extraction from remotely sensed imagery or satellite image is the technique of extracting spatial features such as, vegetation, urban area, land cover and water bodies from satellite images. Satellite images are interpreted manually based on the basis of interpretation elements such as, the image colour, tone, association, texture and pattern information without coming into physical contact of these objects. Multispectral images/data consist of images captured in different bands of electromagnetic spectrum which helps to extract various land cover features.

Geographical Information System (GIS) is an integration of computer hardware and software that store, process, manage, analyze, maps, enhance, edit, output and visualize geographic referenced data [1]. The GIS tools such as, ERDAS IMAGINE [9], ArcGIS [10], etc. are frequently used to visualize and analyze multi-spectral satellite images. Here, in this study, both ERDAS IMAGINE and ArcGIS tools have been used for identifying change in area covered by agricultural land.

The study of LULC (Land Use Land Cover) mapping has been carried out to identify the change in the area covered by the agricultural land. Here, LISS III multispectral satellite images of the year 2014 and 2019 have been collected and analyzed for detecting the change in the area cover of agricultural land. Due to the growing urbanization, it has been observed that the considerable agricultural land has been used for residential and industrial purpose.

In recent years, Artificial Intelligence (AI) and Machine learning (ML) have found very effective almost in all sectors and lot of research is also going on in this field. Similarly, AI and ML are also found effective in various remote sensing and GIS applications where the machine learns from previous experience/data, frequently known as training samples. Depending on the previous learning, the AI and ML model predicts the pattern for unknown or new samples [2]. In this study, Random Forest Algorithm [3] which is one of the supervised learning algorithms has been used. Here, labeled vector data is provided as training samples to ML model, which further learns from these training data and classifies the input data into different classes, such as, vegetation, barren land etc. Further, Land cover mapping corresponding to different classes have been obtained from all the satellite images (of the year 2014 and 2019) and thereafter change in the area cover of each class have been analyzed. The analysis clearly shows decrease in vegetation which is mostly a agricultural land and increase in built up area which is residential or industrial area due to increase in urbanization.

II. Study Area and Data Collection

A. Study Area

In this case study, investigations have been performed on Sangli city and its surrounding area of Maharashtra which is known for its rich agricultural products. Sangli district lies in the southern part of the state, adjoining Karnataka and lies between north latitudes 16° 43' and 17° 38' and east longitudes 73° 41' and 75° 41'. Study area include the satellite image of area 16.75 N to 17.00 N and 74.50 E to 74.75 E which covers the part of the Sangli district. It comprises Sangli city, Miraj and other major villages from the Miraj, Tasgaon, Kavathe Mahankal tehsils which is shown in Fig.1.

Fig.1 Study area in square tile in Sangli District [11]

B. Data Collection and Processing Tools

a. Data Collection

The LISS–III Sensor from series of Earth Observation satellites, i.e. Indian Remote Sensing Satellites (IRS), launched and maintained by Indian Space Research Organization, provides Multispectral data in 4 Bands. Table I shows the various band compositions of the multispectral image used for analysis. These multispectral images have been acquired from ISRO Bhuvan portal [12].

Table I. Spectral bands (in microns)

Band	Spectral Colour	Wavelength(mm)
B2	Green	0.52 - 0.59
B3	Red	0.62 - 0.68
B4	NIR	0.77 - 0.86
B5	SWIR	1.55 – 1.70

b. Data Processing Tools

1. ERDAS IMAGINE: 2020

ERDAS Imagine is an GIS based image processing software package where geospatial as well as vector data can be processed. The software is very useful in remotesensing and GIS application which facilitate data collection, processing, analysis, understanding raw geospatial data, etc. It featured with spatial modellingtools along with machine learning modules which is helpful in carrying out automation applications.

2. ArcGIS Pro 3.0.0

Esri developed a desktop GIS software named ArcGIS. It is mainly used for creating labeled training sample data required for supervised learning classification.

III. Methodology

The overall methodology of the proposed system is shown in Fig. 2. The process begins with collecting the satellite image which is one of the important steps. The result of the application largely depends upon the selection of appropriate satellite image which is suitable to the application in hand.

A. Data Collection

Multi-spectral digital images of LISS-III sensor has been collected from ISRO Bhuvan portal. Four images of four separate bands are collected. Data was of two different time, one from February, 2014 and another from February, 2019.

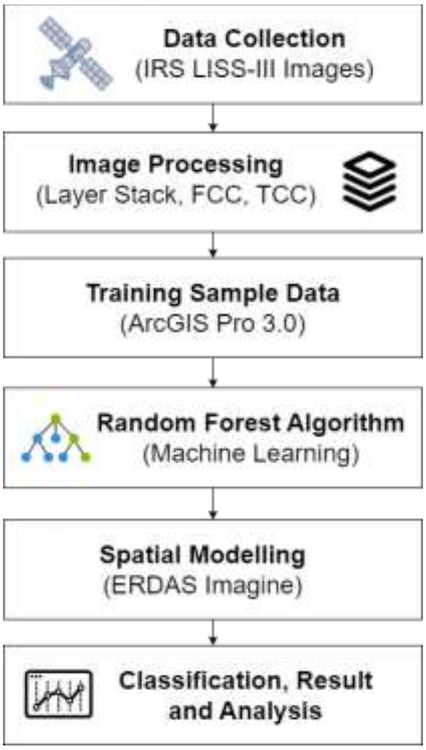


Fig. 2 Flow Diagram of the overall methodology

B. Image Processing

Here, the collected four band images have been processed following the steps given below

1. Layer stack image has been formed by combining four band images using ERDAS Imagine.
2. Spatial, spectral, radiometric and temporal resolutions have been applied to enhance the image quality.
3. Geometric and radiometric corrections have been done on the collected data.
4. True Colour Composite Image (TCC):

   In TCC, there are R, G and B values correspond to true colours of red, green and blue, Fig. 3.

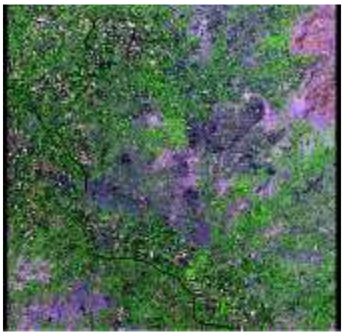


Fig. 3 True Colour Composite

5. False Colour Composite (FCC):

The RGB values used in FCC do not necessarily represent the true colors of red, green, and blue. Specifically, the FCC corresponds very-near infrared as red, red as green, and green as blue. Hence, vegetation which reflect near-infrared appears as red, Fig. 4.

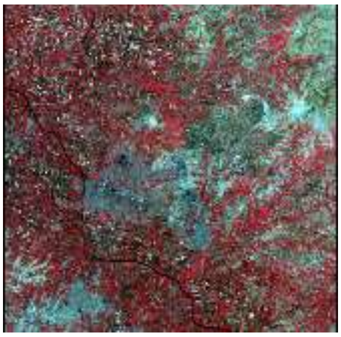


Fig. 4 False Colour Composite

C. *Training Sample Data*

- ArcGIS Pro 3.0.0 provides toolbox with feature of creating training samples in vector format.
- Initially, FCC image has been opened in workspace, in the next step, different classes for classification have been defined and each class has assigned a specific colour in order to differentiate the classes visually after classification.
- A set of circles, polygons has been used to denote samples of particular class on image.
- Multiple samples of a class are merged together as single entity.
- Vector data image have been downloaded after creating samples for all classes.
- More the number of samples, more is accuracy.
- In this case study, four classes were defined.
- Classes along with number of samples taken for training are shown in Table II.

Table II. Class wise sample collected

Class Name	**No. of Samples**
Vegetation	120
Barren Land	112
Built-Up	98
Water	124

D. *Machine Learning Algorithm*

In machine learning, classification algorithm has been used to classify data. It is a supervised learning technique used to identify the category of new observations based on the basis of training samples. Here, a program model learns from the given training dataset and then classifies the new observation into respective category or class. In order to classify land cover into different classes Random Forest algorithm has been used, Fig. 5.

Fig. 5 Machine Learning Workflow [13]

E. *RANDOM FOREST (RF)*

- It is a regression and classification tree technique invented by Breiman.
- A Random Forest is a machine learning algorithm that generates a large group or forest of decision trees for classification and regression tasks, Fig. 6. It achieves this by iteratively and randomly sampling the data. [4]
- Algorithm randomly selects observations, builds a decision tree and the average result has been taken.

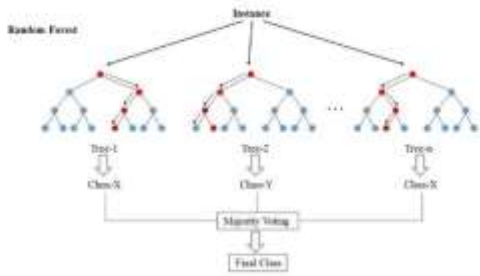


Fig. 6 Random Forest Algorithm [14]

F. *Spatial Modelling*

Spatial modeller is a tool in ERDAS Imagine which allows to build a flow chart that contain logically connected operators to produce output results. It is an effective way to process input raster data and to produce output raster data.

Operators used for this study are as follows:

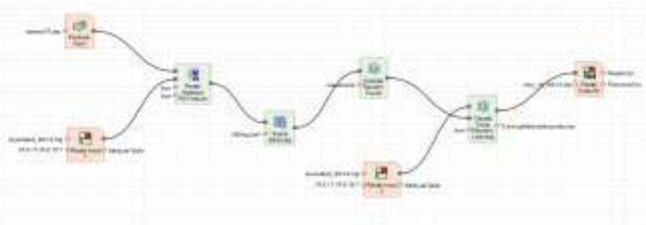


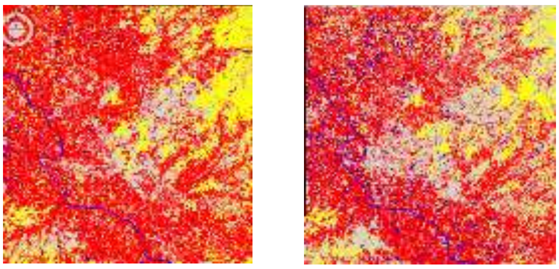
Fig. 7. Machine learning Spatial Model

1. Feature Input: Training data in the form of shape file i.e. vector data has been provided to this operator. Shape file is an image containing manually digitized circular or rectangular (vector) polygons.
2. Raster Input: This operator takes raster image as input which is in ERDAS IMAGINE format.
3. Raster Statistics per Feature: Statistics will be computed on pixel of input raster in this Operator. The computed values of statistics have been added as attributes to the feature stream. Mean computation is default.
4. Select Attribute: In the select attribute operator, required attributes for initializing random forest like id, class or value has been used
5. Initialize Random Forest: Attributes selected from previous operator are provided to this operator which defines and trains a random forest classifier.
6. Machine Learning Input: This operator takes the saved output file which has been used in classifying. Here, direct output of Random Forest is connected to input of machine learning operator.
7. Classify using Machine Learning: Classification on input data using trained classifier specified on the machine intellect has been carried out by this operator.
8. Raster Output: Processed, resultant and analysed data has been presented in the form of raster output
9. Reprojection: Resultant images have been reprojected from georeferenced to UTM zones. Finally, area section has added to the attribute table.

IV. Results and Discussion

The steps mentioned in methodology have been performed on both February, 2014 and February, 2019 data set.

The final processed and machine learning classified images of February, 2014 and February, 2019 are as follows:



Feb, 2014 (a) Feb, 2019 (b)

Fig. 8 ML Classified Images

In Fig. 8 (a) and (b), in both output images, colour representation is as follows:

Table III: Different classes and respective Colour Representation

Color	Class Name
Black	Background
Yellow	Barren Land
Grey	Built-Up
Red	Vegetation
Blue	Water

Attribute tables for each of the output image which contain attribute values like histogram, area (in acres), etc. are as follows:

A. *February, 2014 Attribute Table*

Table IV. Attribute Table of February, 2014 Image

Class Name	Area
Vegetation	106144
Barren Land	32241.2
Built-Up	53153.8
Water	3793.82

B. *February, 2019 Attribute Table*

Table V Attribute Table of February, 2019 Image

Class Name	Area
Vegetation	91921.6
Barren Land	24061.8
Built-Up	69933.7
Water	9415.46

V. Analysis

From Table IV and Table V, based on area attribute, analysis has been carried out by plotting values of area of February, 2014 and February, 2019 images shown in Table VI for each class i.e., vegetation, barren land, built-up area and water. Table VI contains columns showing net increase or decrease in the area for particular class which is given as follows:

Table VI Analysis Table

Class	Area (in acres)		Change	Comment
	Feb 14	Feb 19		
Vegetation	106144	91921.6	14222.4	Decreased

Barren Land	32241.2	24061.8	8179.4	Decreased
Built-Up Area	53153.8	69933.7	16779.9	Increased
Water	3793.82	9415.46	5621.64	Increased

From the above empirical study, it has been observed that vegetation has been decreased by 13.40%, which includes most of the agricultural land in study area. Barren land also shows 25.36% decrease in 5 years. However, there is a rise by approximately 23.99% in built-up area which is due to growing urbanization.

VI. Conclusion and Future Scope

Land use land cover (LULC) change detection analysis using remote sensing data by using machine learning is an important source of extracting information for various land cover classes, such as vegetation, water bodies, built-up, etc. Here, comparison of various classes extracted after classification from two different satellite images captured in 2014 and 2019 has been done. Further changes in various classes such as, vegetation, water body, built-up, etc. have been analyzed to identify the change in vegetation which mostly include agricultural land. The result obtained shows that the vegetation in the study area is decreased by 13.40% which includes most of the agricultural land and barren land also decreased by 25.36% in 5 years from 2014 to 2019. However, there is a rise by approximately 23.99% in built-up area. Such result may be important for government organization for town planning and related decisions. This shows that, the change detection analysis is an efficient way of describing the changes observed in each land use category.

Further, use of normalized difference vegetation index (NDVI) may be used to extract data about vegetation. NDVI technique may also be used to identify and obtain LULC for different types of vegetation like shrubs, forest land, agricultural land, etc

References

[1] Chouhan, Rajlaxmi, and Neeraj Rao. "Vegetation detection in multispectral remote sensing images: Protective role-analysis of vegetation in 2004 indian ocean tsunami." PDPM Indian Institute of Information Technology (2011).

[2] Liu, Zelin, Changhui Peng, Timothy Work, Jean-Noel Candau, Annie DesRochers, and Daniel Kneeshaw. "Application of machine-learning methods in forest ecology: recent progress and future challenges." Environmental Reviews 26, no. 4 (2018): 339-350.

[3] Dobrinić, Dino, Mateo Gašparović, and Damir Medak. "Sentinel-1 and 2 time-series for vegetation mapping using random forest classification: A case study of Northern Croatia." Remote Sensing 13, no. 12 (2021): 2321.

[4] Zagajewski, Bogdan, Marcin Kluczek, Edwin Raczko, Ajda Njegovec, Anca Dabija, and Marlena Kycko. "Comparison of random forest, support vector machines, and neural networks for post-disaster forest species mapping of the krkonoše/karkonosze transboundary biosphere reserve." Remote Sensing 13, no. 13 (2021): 2581.

[5] Svoboda, Jan, Přemysl Štych, Josef Laštovička, Daniel Paluba, and Natalia Kobliuk. "Random Forest Classification of Land Use, Land-Use Change and Forestry (LULUCF) Using Sentinel-2 Data—A Case Study of Czechia." Remote Sensing 14, no. 5 (2022): 1189.

[6] Jiang, Gang, and Quanshun Zheng. "Remote Sensing Recognition and Classification of Forest Vegetation Based on Image Feature Depth Learning." Mobile Information Systems 2022 (2022).

[7] Caruana, Rich, and Alexandru Niculescu-Mizil. "An empirical comparison of supervised learning algorithms." In Proceedings of the 23rd international conference on Machine learning, pp. 161-168. 2006.

[8] Nageswara Rao, P. P., S. V. Shobha, K. S. Ramesh, and R. K. Somashekhar. "Satellite-based assessment of agricultural drought in Karnataka state." Journal of the Indian society of remote sensing 33 (2005): 429-434.

[9] Web Reference: https://hexagon.com/products/erdas-imagine last accessed on 20/03/2023.

[10] Web Reference: https://www.esri.com/en-us/arcgis/products/arcgis-pro/overview last accessed on 20/03/2023.

[11] Web Reference: https://nrhm.maharashtra.gov.in/mapsangli.jpg, https://www.researchgate.net/figure/Location-map-of-study-The-commonly-used-verities-of-grapes-in-Sangli-district-are_fig1_299545115 last accessed on 20/03/2023.

[12] Web Reference: https://bhuvan.nrsc.gov.in/home/index.php last accessed on 20/03/2023.

[13] Web Reference: https://www.researchgate.net/figure/Basic-Classification-Algorithm-kNN-is-one-of-the-best-algorithm-that-can-be-used-for_fig1_328738427 last accessed on 20/03/2023.

[14] Web Reference: https://www.researchgate.net/figure/Classification-process-based-on-the-Random-Forest-algorithm-2_fig1_324517994 last accessed on 20/03/2023.

[15] Web Reference: https://www.esri.com/en-us/what-is-gis/ last accessed on 20/03/2023.

[16] Web Reference: https://bhuvan.nrsc.gov.in/home/index.php last accessed on 20/03/2023.