Land Use Land Cover Mapping with change detection of Mumbai City, using Spatial-Temporal Analysis and Machine Learning Techniques

Apoorv Panse
MT2020013
apoorv.panse@iiitb.ac.in

Deepti Chawda
MT2020050
deepti.chawda@iiitb.ac.in

Sounak Dey
MT2020081
Sounak.Dey@iiitb.ac.in

***Abstract*—This is the detailed project report on our work on GIS Project based on Land Used Land Cover of Mumbai city. Prime goal of this project was to learn how to use machine learning techniques to analyze the urban growth and better planning of future growth. We used Supervised, Unsupervised ML techniques following the result and accuracy assessments.**

Keywords: ML, Supervised, Unsupervised, Training Data, Pixels, Landsat, Sentinel, LULC.

I. Introduction and Problem Statement

Mumbai is India's one of the most populous city and its urban growth has been unprecedented in recent years. In order to see how it has expanded over the years, and to better facilitate the future areas of expansion, there was need to use GIS techniques to analyze it. The main background research was to see the changes over the last two decades. Main problem statement includes using ML techniques to classify areas around Mumbai city and identify rapid urbanization areas. Also to find out which classification techniques work best among supervised and unsupervised. The main objective of this project was to reduce harmful and potential hazardous impact of urbanization and to use analysis data in better planning of future developments.

II. Data Set

Primary data is the satellite images that were procured from USGS official website and EO Browser. Following images were downloaded:

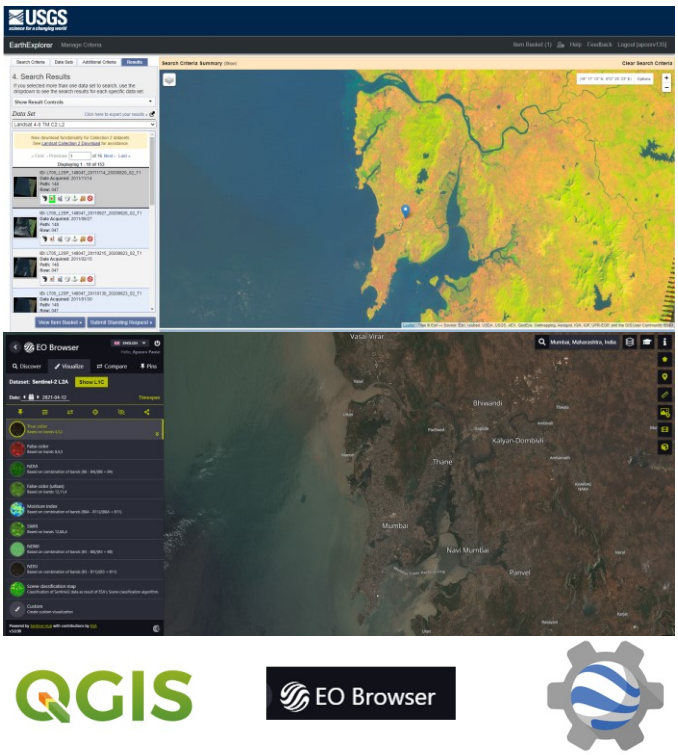
LANDSAT 5 – Spatial Resolution 30m – Year 2001
LANDSAT 8 – Spatial Resolution 30m – Year 2021
SENTINEL 2 – Spatial Resolution 10m – Year 2013
SENTINEL 2 – Spatial Resolution 10m – Year 2021

The data bundle included all the RAW band sets from respective satellite data.

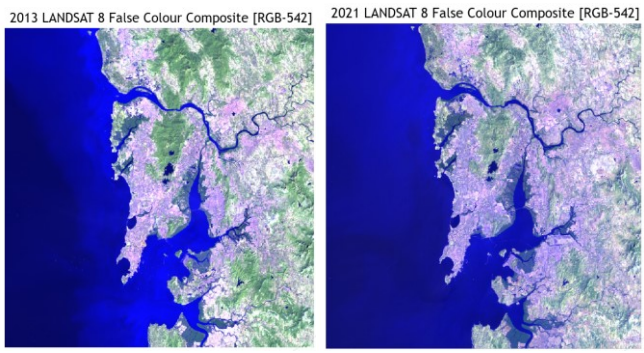
III. Methodology

1. Data Collection from USGS official website for LANDSAT images.

LANDSAT 5 (2001), LANDSAT 8 (2021), SENTINEL 2 (2013, 2021)

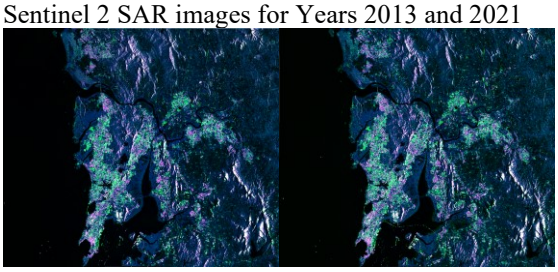


2. Pre-Processing of acquired images. Cropping of images, creating virtual band sets for classification inputs.



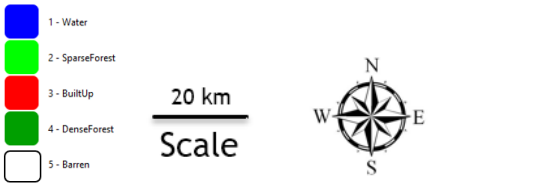
False colour composite images helps us better understand the urban areas, as shown above, highlighted in pink colour, while retaining vegetation as green and water bodies as blue colour.

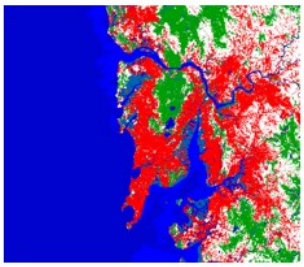
Sentinel 2 SAR images for Years 2013 and 2021



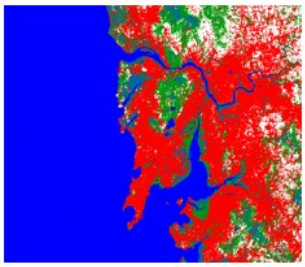
The SAR (Synthetic Aperture Radar Images) were procured from Sentinel Hub website which clearly highlights the urban regions for our initial analysis.

3. Unsupervised and Supervised classification using QGIS Semi-automatic classification plugin. Creation of training inputs for supervised classification, tuning various parameters available to get optimal result.



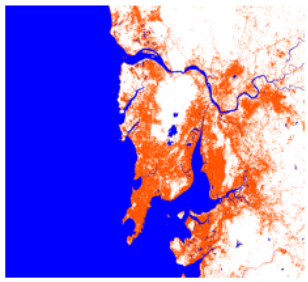


2013 Un-supervised Classification

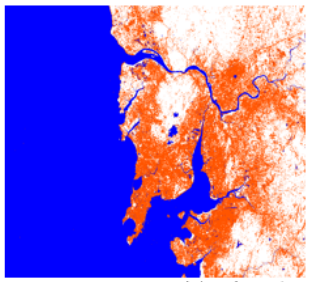


2021 Un-supervised Classification

Here we can see that the Supervised classification did not provide accurate result because the Built-Up area is significantly larger than what we saw visually earlier. For this method we used K-Means Clustering algorithm and tuning parameters were set to include 10 classes which were later merged to show 5 Major classes as shown in the legends above.



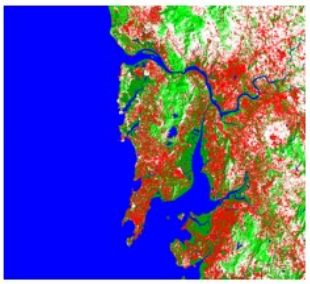
2001 Supervised (Built Up)



2021 Supervised (Built Up)

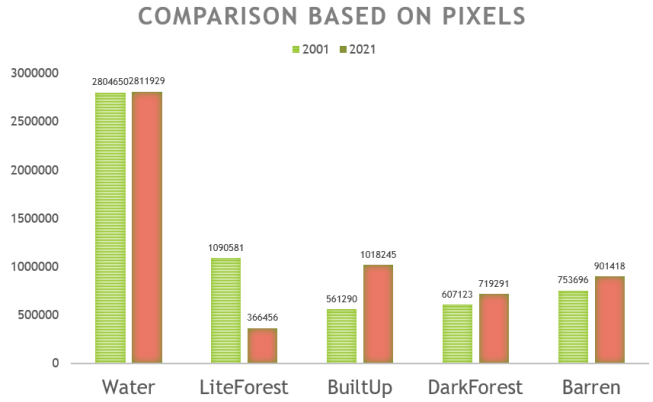
For initial analysis of Built-Up area using supervised classification gave us fair result and near accurate representation of the growth around the city. We can clearly see the expansion happening around the riverbanks.

2001 Supervised Classification



2021 Supervised Classification

Supervised classification gives us satisfactory result and help us analyze the spread of urban area over the years very well. As shown above, there has been drastic growth both in the prior urban area as well as the area surrounding the rivers.



If we see the changes based on classes, Water is almost unchanged as major portion is ocean which will not change much. Sparse forest has slightly reduced, and Dense Forest have increased a bit. Built Up area is significantly increased.

4. Accuracy measurements of classified images. Users Accuracy, Producers Accuracy, Overall Accuracy

2021 Accuracy Assessment (Unsupervised classification)

% classified as →	Water	Sparse Forest	BuiltUp	Dense Forest	Unused	User's Accuracy
Water	**90**	5	0	0	5	90
Sparse Forest	0	**80**	0	15	5	80
BuiltUp	0	5	**85**	10	0	85
Dense Forest	0	10	0	**90**	0	90
Unused	0	20	20	0	**60**	60
Producer's Accuracy	100	66.6	80.9	85.7	85.7	**Overall 81 %**

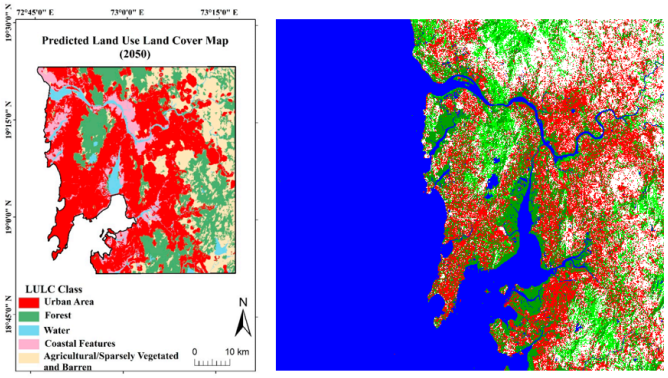
2001 Accuracy Assessment (Supervised classification)

% classified as →	Water	Sparse Forest	BuiltUp	Dense Forest	Unused	User's Accuracy
Water	**100**	0	0	0	0	100
Sparse Forest	0	**100**	0	0	0	100
BuiltUp	0	10	**80**	0	10	80
Dense Forest	10	30	0	**60**	10	60
Unused	0	25	5	0	**70**	70
Producer's Accuracy	90.9	60.6	94	100	77	**Overall 86%**

Users' accuracy is how good the classification is, compared to original image. Producers' accuracy is how accurately each pixel is classified to its correct class. We can see that in Supervised learning, more than 80% overall accuracy was observed.

5. Temporal Change analysis of classified images. Urban Growth and potential expansions

Change Heat Map
(Lighter region is unchanged pixels
Darker regions depict drastic changes)



Now, the urban growth rate has been significant as shown on the left side of the above image, taken from a research paper, it predicts the urban expansion in 2050 and current scenario on the right-hand side shows that the progression is correct. The areas near river side and area which are around the hills are likely to expand more.

Our analysis can help us in better planning in those areas, which may include resource utilization, deforestation, agricultural efficacy, industrial growth etc.

IV. Areas of Improvement and Comments

- Satellite data obtained had lesser resolution than actual. Hence classification produced is not as accurate as we expected it to be.
- Supervised classification gave better result compared to Unsupervised classification.
- Clear indication of areas that were expanded over the years and projected areas where it can further expand, enabling us better planning of urban transport, setting up of new housings and nearby resources.
- Potential impact includes deforestations, land degradation, pollution, population growth resulting in exponential urbanization.

V. Acknowledgements

We would like to thank Prof. Uttam Kumar, IIIT Bangalore, for his tremendous feedback and encouraging system, which helped us not only learn much more about GIS subject, but to implement it in a practical scenario. We would also like to thank all the TA involved in explaining various hands-on during lab sessions. Project was fruitful and we enjoyed the learning path.

VI. Reference

[1] Lecture Slides by Prof. Uttam Kumar, IIITB.

[2] Vinayak, B.; Lee, H.S.; Gedem, S. Prediction of Land Use and Land Cover Changes in Mumbai City, India, Using Remote Sensing Data and a Multilayer Perceptron Neural Network-Based Markov Chain Model. Sustainability ***2021****, 13, 471. https://doi.org/10.3390/su13020471*

[3] *Producing land cover change maps and statistics, ESCAP September 2020, Step by step guide on the use of QGIS and RStudio.*

[4] Land Use Land Cover Mapping with Change Detection: A Spatio-temporal Analysis of NCT of Delhi from 1981 to 2015, INTERNATIONAL JOURNAL OF GEOMATICS AND GEOSCIENCES Volume 8, No 1, 2018