Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

MONITORING LAND USE AND LAND COVER CHANGE USING ENSEMBLE MACHINE LEARNING CLASSIFIERS

SUBHRA SWETANISHA[1], AMIYA RANJAN PANDA[2], DAYAL KUMAR BEHERA[3]

SHREELA DASH[4],

[1]Research Scholar, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India
[2] Faculty, School of Computer Engineering, KIIT Deemed to be University, Bhubaneswar, India
[3]Department of Computer Science and Engineering, Silicon Institute of Technology, Bhubaneswar, India
[4]Department of Computer Science and Engineering, Centurion University of Technology and Management, Bhubaneswar, India

E-mail: [1]sswetanisha@gmail.com, [3]dayalbehera@gmail.com, [4]shreelamamadash@gmail.com

ABSTRACT

Land Use and Land Cover (LULC) classification is a better tool for change detection and other remote sensing applications. Changes in LULC have become a vital aspect of conventional strategies for land cover monitoring. The aim of this work is to monitor the changes using classified images of the machine learning models. A change detection exercise was conducted in the research region, namely Kendrapara District, Odisha. Landsat7/8 imagery was used in 1999 and 2020 to track potential changes, particularly in agricultural land and urban or built-up land, to detect urban growth. This work uses three machine learning models to determine the best classifier from SVM, XGBoost and Ensemble Model of SVM and XGBoost. The ensemble classifier outperforms the other two machine learning models. Then the best model is used to analyze the change. This change detection may enable the government to implement legal measures & norms and develop the city.

Keywords: *Land Use and Land Cover, Change Detection, SVM, XGBoost, East Odisha*

1. INTRODUCTION

Large-scale land cover monitoring is one of the fundamental goals of remote sensing satellite systems. Identifying when and where the land cover has altered is a critical part of the research. When studying regional environmental change, it's crucial to look at changes in land use and land cover (LULC). As a part of the study of global environmental change, researchers are focusing on land use and land cover. Land-use change analysis has never been easier because of recent developments in remote sensing and GIS technology [1]. By gaining a deeper understanding of the land use process and changes in legislation and trends, the ability to predict future land use and land cover is a valuable tool for land-use planners and managers. Scientists from all over the world have been working hard to come up with creative algorithms and models to predict land cover change [2][3].

Multiple causes at the local, regional, and global stages are responsible for land-use changes, which are dynamic in nature. The land use pattern has been constantly transformed by rapid and unregulated population growth, industrialization, and commercial expansion. LULC transformations could be sparked by a variety of factors, including urbanization and industrialization, climate change [4] and population growth [5], urban development [6], and regulatory measures [7]. Due to natural calamities, land cover changes play a vital role in sudden changes of various physical structures of an area on Earth. For example, The Super cyclone that struck Odisha on 29-30 October 1999 was the most severe tropical cyclone ever recorded in the North Indian Ocean and one of the most destructive in the region. In coastal areas such as Balasore, Bhadrak, Kendrapara, Jagatsinghpur, Puri, and Ganjam, lives and property were lost or damaged due to the storm. The State Government's 2000 white paper on the destruction profile of the 1999 super cyclone stated that there were 9,885 fatalities [8]. Nearly

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

10,000 people lost their lives as a result of the storm's devastating storm surge and projected top wind speeds of 260-270 kph in the storm's core area. It was accompanied by torrential rains, which caused devastating floods and blocked off the state from the rest of the country. Images from satellites, as well as machine learning models[9], are used to characterize land-use changes throughout a certain time period [10].

This research article looks at how remote sensing and GIS technology can be used to track long-term changes in land use and land cover in the Kendrapara district of Odisha, India. This area we took for our study purpose since this is a place where huge vegetation takes place and is also prone to natural calamities. This can be proved suitable for LULC classification [11] and change detection. Landsat Image was the primary source of information. A land-use classification of the study area was done for two different years, 1999 and 2020. Land-use changes from 1999 to 2020 were looked at using the land use transfer matrix, which took changes in elevation and the landscape index. For the change analysis, much attention is paid to the satellite data[5].

Odisha is a primarily rural state. The majority of the state's population depends on agriculture. The agriculture industry generates only approximately 26% of the Gross State Domestic Product (GSDP), and its dependence on more than 70% of the population results in a low per capita income. The 482 kilometres of Odisha's coastline make the state vulnerable to flooding, cyclones, and storm surges. Every year in the middle of the year, the state is ravaged by natural disasters. This is a major contributing factor to Odisha's sluggish economic growth. Kendrapara is one of Odisha's districts most impacted by flooding. Consequently, LULC is essential for all these reasons. Changes to the LULC have demonstrated both direct and indirect effects on many environmental factors.

The problem statement for this work is "Given satellite images, classify the Land Use and Land Cover(LULC) of the Kendrapara region and finding the changes using the best classifier to facilitate adaptability and better policy making".

The following are the work's objectives:

- To classify land use and land cover using various machine learning methods.
- To generate the feature set from the raster using the training and test data's shape file.
- To develop various ML classification models for the Kendrapara District.
- To analyze the effectiveness of machine learning models in terms of user, producer, and overall accuracy.
- To use an ensemble model (XGBoost+SVM) to perform change detection on Kendrapara between 1999 and 2020.

2. LITERATURE SURVEY

Land cover change detection(LCCD) utilizing bi-temporal remote sensing pictures has been a significant topic in remote sensing applications. LCCD may provide timely and large-scale land cover change information to aid urban development plans, such as urban expansion, urban build-up and research of city surface temperature change. In practice, a variety of change detection approaches have been used [12]. These approaches can be classified as "post-classification change detection methods" or "binary change detection methods" based on their detection outcomes [13]. In this section, relevant works in the same field are discussed.

In [14], the authors used data from three Landsat TM and ETM+ pictures of Beijing taken between April 9, 1995, and April 30, 2000. The land surface temperature (LST) and land use and land cover (LULC) classifications were extracted from the dataset. TVX space was designed to evaluate the impact of land changes on the LST index (TVX). Since dense vegetation is associated with lower temperatures than sparse vegetation, this suggests that land-use change is a substantial factor in LST rise.

The authors in [15] described a novel approach for updating land-cover maps by identifying satellite image time series.

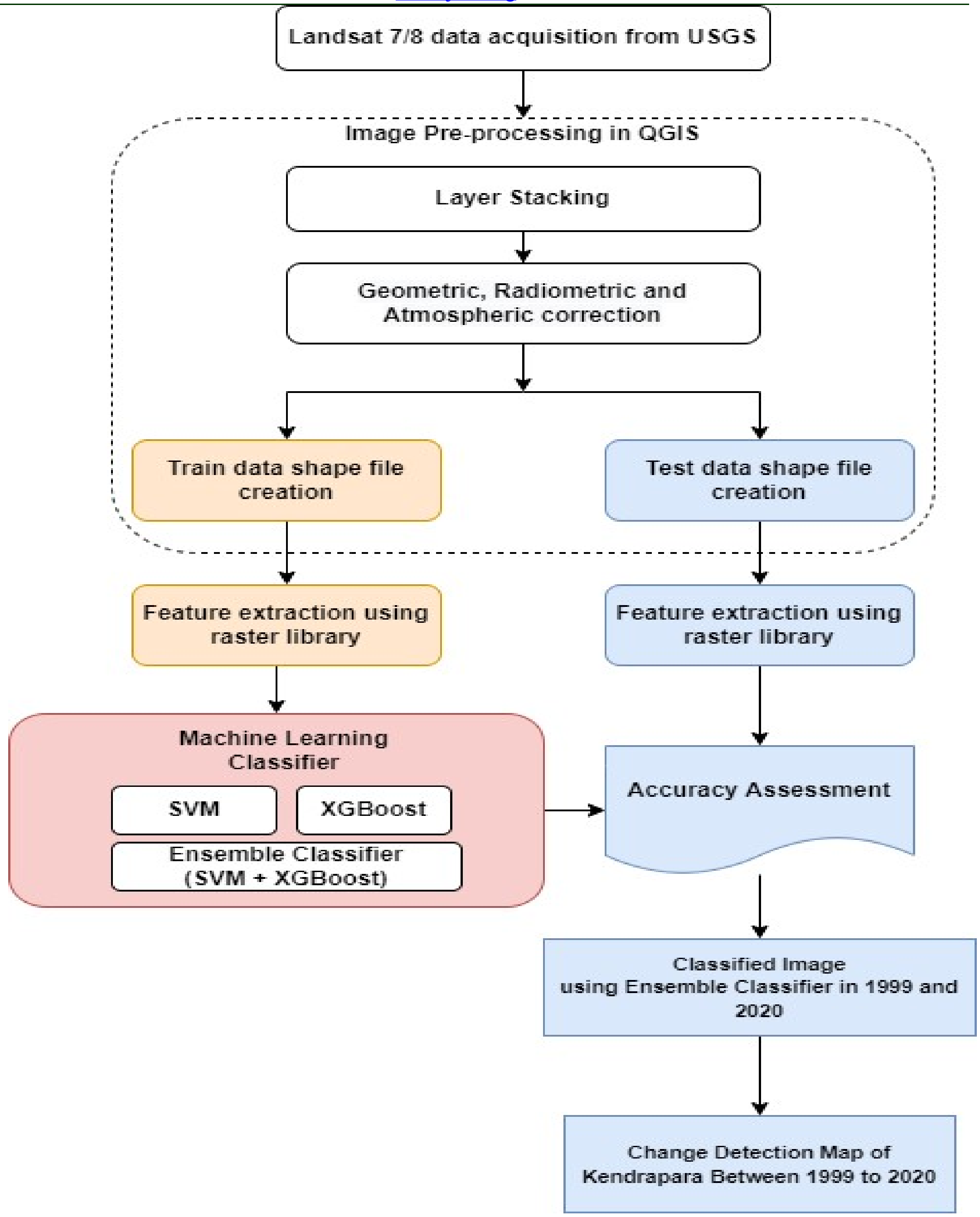


Figure 1: Proposed Model

The suggested method identifies an image for which no ground truth information is available by using knowledge from an image taken at a different time in the same area of interest. This method overcomes the two major flaws of prior approaches. 1) It is unaffected by possible significant discrepancies in the land-cover class distributions of the source and target domains, and 2) it can handle circumstances in which the two domains have different sets of land-cover classes.

In [16], multi-source remote sensing data were utilized to track changes in land use in the Zigui region from 2008 to 2014. During the analysis phase, the transfer matrix, various altitudes, and the patch landscape index were considered. The entire area of vegetation dropped by 6.1 per cent from 2008 to 2014, with the majority of the reduced vegetation changed into cultivated land.

Landsat thermal data and a field study in India's Lucknow city have examined how urbanization has harmed the city's environment and contributed to a rising temperature trend [17]. It was decided to employ the Mono-window approach to estimate land surface temperature (LST), the Normalized Difference Vegetation Index (NDVI), and the Urban Thermal Field Variance Index (UTFVI) for evaluating the city's ecological health.

LULC and built-up surfaces in Delhi's suburbs [18] were examined in relation to population increase and migration between 1990 and 2018. Landsat 5 (TM) and Landsat 8 (OLI/TIRS) data were used for the LU/LC classification of Delhi NCR. The Landsat data was processed using the K-means clustering technique, followed by a change detection technique to measure the LULC change. As a result of increasing built-up areas and open/fallow land and decreasing farmland and vegetation, considerable changes in LULC have been seen across the study period.

In [19], the authors tried to solve challenges like settlement expansion in South Africa's Limpopo Province and detecting deforestation in Australia's New South Wales. Compared to the other tested approaches, the method exhibited a much shorter median detection delay (DD) with equal rates of false alarms.

The rate of deforestation in Northwestern Paraguay was studied using satellite pictures from LandSAT5 and LandSAT7 [20]. The rate of deforestation is graphically observed from 1986 to 2011, and future predictions are produced. LandSAT8 pictures are utilized to verify the prediction given until 2018. An extrapolation of the graph shows that the deforestation process is already three years ahead of schedule.

From 1988 to 2017, the authors used Landsat-5 TM and Sentinel-2 data to assess and monitor changes in LULC patterns in the Rani Khola watershed of the Sikkim Himalaya. The findings show that forestry is the most common land use in the watershed [21].

3. PROPOSED MODEL

Figure 1 depicts the proposed method. A collection of raw satellite images has been made available and utilized for pre-processing. Stacked images and shape files are used to extract the features using the raster library of R. Then, the dataset is used for training and testing, with 70% of the data used for training and 30% for testing. The LULC classification of the Kendrapara dataset is then carried out using machine learning algorithms. The following sections provide more in-depth explanations.

2.1 Data and Study Area

The research area is Kendrapara district of state Odisha, India. The co-ordinate of Kendrapara in the form of degrees and minutes as 20.4969° N and 86.4289°E. District boundaries include Bhadrak District to the north, Jajpur to the west, Jagatsinghpur and Cuttack Districts to the south, and the Bay of Bengal to the east. Cuttack District was divided into Kendrapara District on April 1, 1993. We used the USGS Earth Explorer to gather satellite photographs of the Kendrapara district in 2020. Figure 2 depicts the general layout of our research space. In future, data collected from field surveys can be incorporated into preparing the ground truth for classification.

2.2 Data Pre-Processing

Here we are using the multispectral data of Landsat satellite-7/8, which includes seven bands. Landsat 7/8 can measure colour in the electromagnetic spectrum, but not necessarily a colour that the human eye can see. It's called a "band" for each range. Figure 3 depicts Kendrapara in seven Landsat 8 bands.

The pre-processing of image data includes layer stacking, radiometric correction, geometric correction, atmospheric correction, cloud-moving, etc. The QGIS tool is used to complete all the pre-processing work. The pre-processed image is presented in Figure 4.

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

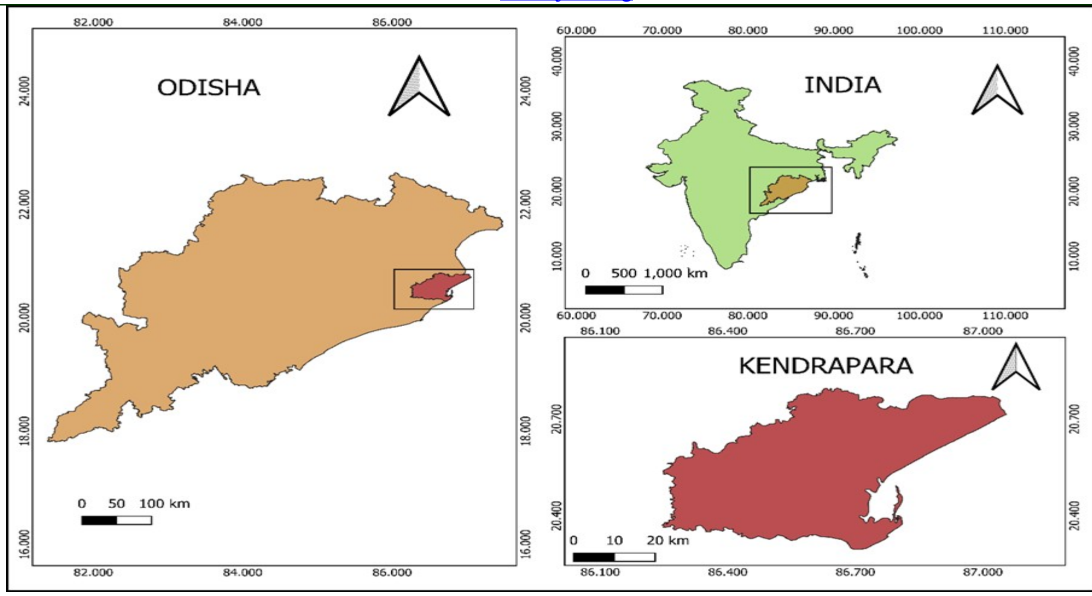


Figure 2: Location of Study Area

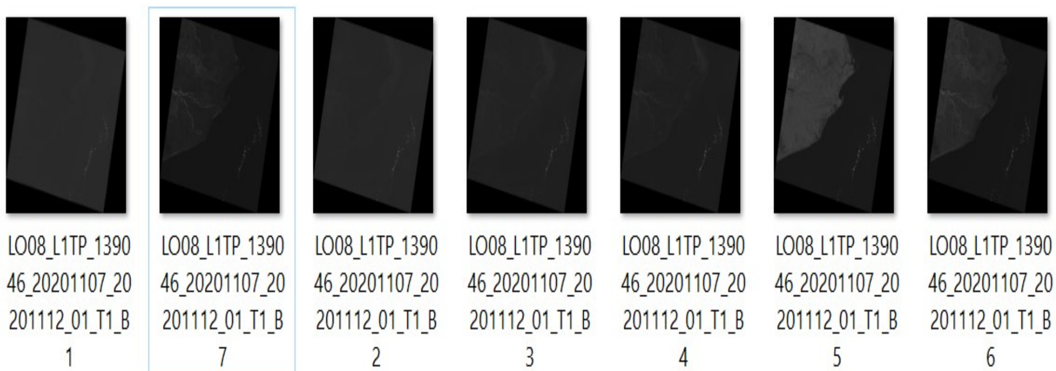


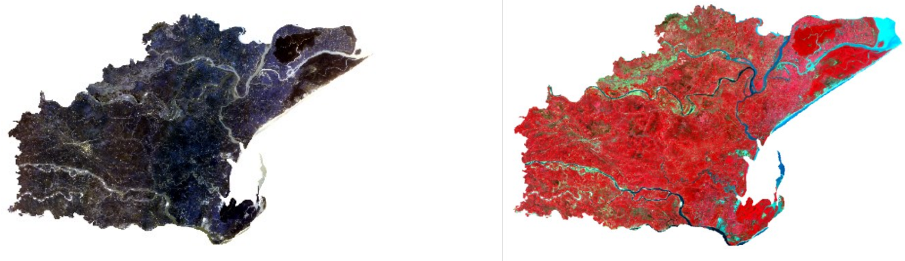
Figure 3: 7 Bands of Kendrapara by Landsat 8

(a) Layer Stacking of 7-bands *(b) After Atmospheric Correction*

Figure 4: Pre-processing (a) Layer Stacking (b) Atmospheric Correction

Journal of Theoretical and Applied Information Technology
30th September 2022. Vol.100. No 18

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195



(a) True Color Composite (b) False Color Composite

Figure 5: Color Composite of Kendrapara District

If a multispectral image contains the red, green, and blue bands that make up the main colours of the human visual system, then those three bands can be combined to create a "true colour" image. The visible light bands red (B04), green (B03), and blue (B02) are used in the respective red, green, and blue colour channels in the true colour composite. This results in a natural-coloured image that is a good portrayal of the Earth as it would appear to people in its natural state. On the other hand, the human eye cannot see some wavelengths, but false colour composites enable humans to perceive them (near the infrared range).

Increasing spectral separation and improving the readability of data can both be accomplished by the application of bands, such as the near-infrared. Any band in a multispectral image can have its display colour assigned in a totally random manner. This can be done with any band. When referring to a series of colour rendering technologies that are used to display images in colour that were captured in the visible or non-visible sections of the electromagnetic spectrum, the term "false colour" (also known as "pseudocolour") is commonly used. An image that represents an object in colours that are different from those that a photograph (which would provide a true-colour image) would reveal is referred to as a false-colour image. Figure 5 shows the true colour and false colour composition in the Kendrapara District.

4. METHOD AND ANALYSIS

The various machine learning model classifiers were used to categorize the land use and land cover data collected during each year of the study. In addition, the land use classification for the years 2008 and 2011 was determined using landsat7 ETM, while the land use classification for the year 2020 was determined using landsat8 OLI.

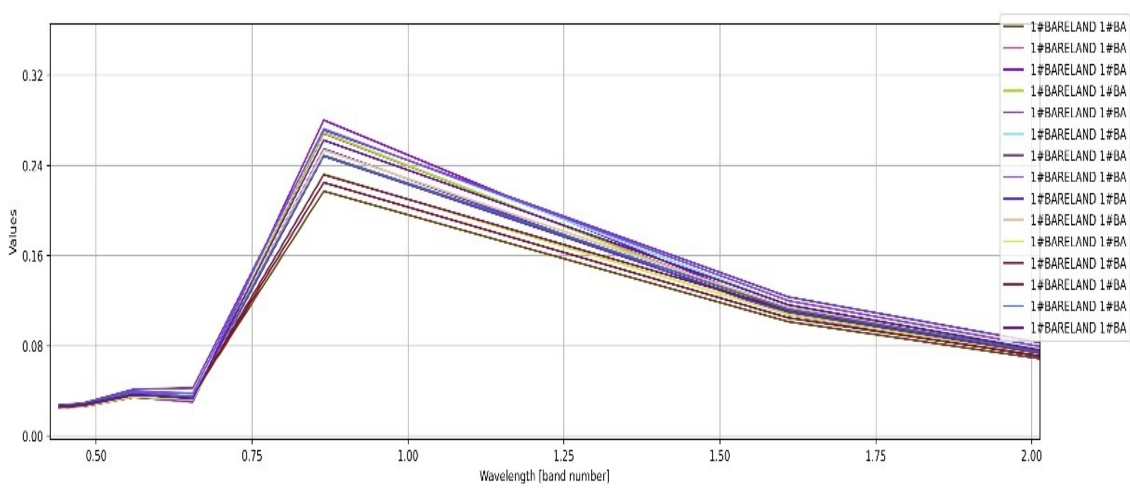
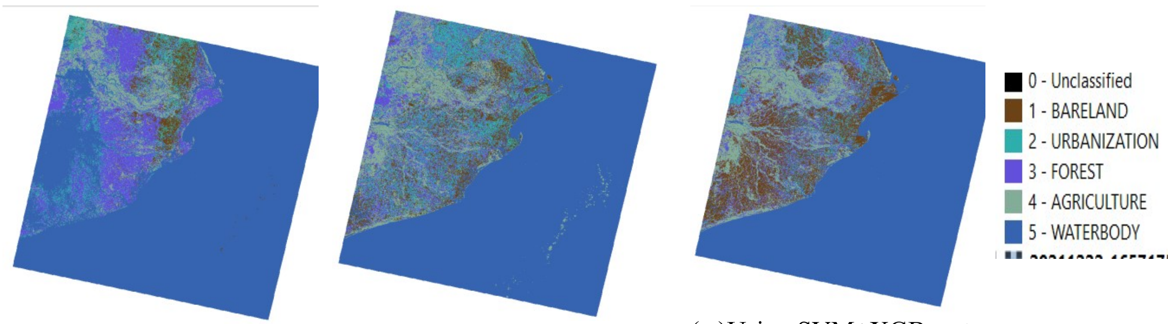


Figure 6: Spectral Signature of Bareland Class



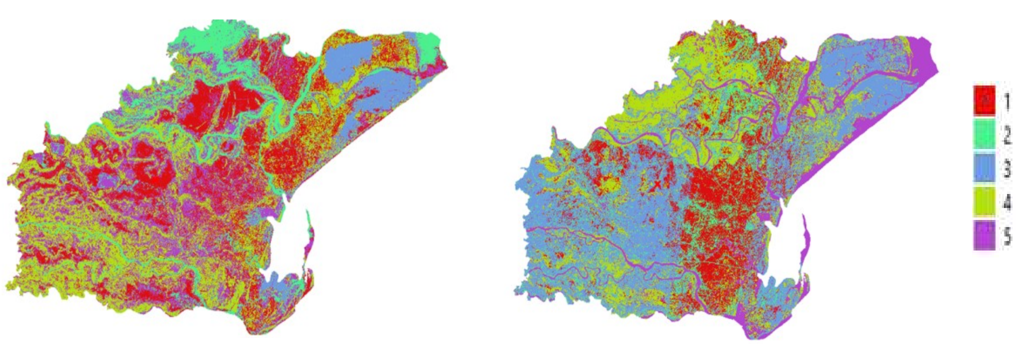
(a) Using XGBoost (b) Using SVM (c)Using SVM+XGBoost

Figure 7: Classified Map of Kendrapara Using ML Models

Taking into account the resolution of the image as well as the particular circumstances of the research region, we divided the land use into the following five categories: Bare-land, Agri-land, Waterbody, Forest and Urbanization.

In order to record the EMR that is emitted and reflected by the various components of the test locations, a variety of remote sensors are utilized. The term "signature" refers to any parameter that may be remotely sensed and that, either directly or indirectly, characterizes the nature and/or state of the object that is the focus of the observation. When conducting remote sensing, spectral signatures are utilized, and this spectral observation is the process of measuring the spectral reflectance of certain objects. These parameters are highly helpful in supplying important data that can be used to differentiate between the objects. The Spectral signature of Bareland class of our study area is depicted in Figure 6.

Samples of Kendrapara districts were extracted to train the model. The samples were separated into five categories: agricultural land, bare land, water bodies, urbanization, and forest. Following the conversion of the geographic polygon data set to a CSV file, ML models are implemented. The XGBoost [22][23], SVM , and ensemble model (SVM + XGBoost) are applied to the dataset and checked based on how our training samples classify the whole Kendrapara district. Classified images of the Kendrapada district by all the used ML models are shown in Figure 7. Accuracy assessment with error metrics is crucial to any classification job, including LULC. Table 1 displays the accuracy computed from the confusion matrix. As SVM + XGBoost is more accurate, Figure 8 shows the classified image of the ensemble model. The colour code from 1 to 5 represents Bareland, Urbanization, Forest, Agriculture, and Waterbody, respectively. Figure 9 depicts the change detection map.



(a) (b)

Figure 8: (a) Classified Image using SVM+XGboost in 1999 (b) Classified Image using SVM+XGboost in 2020

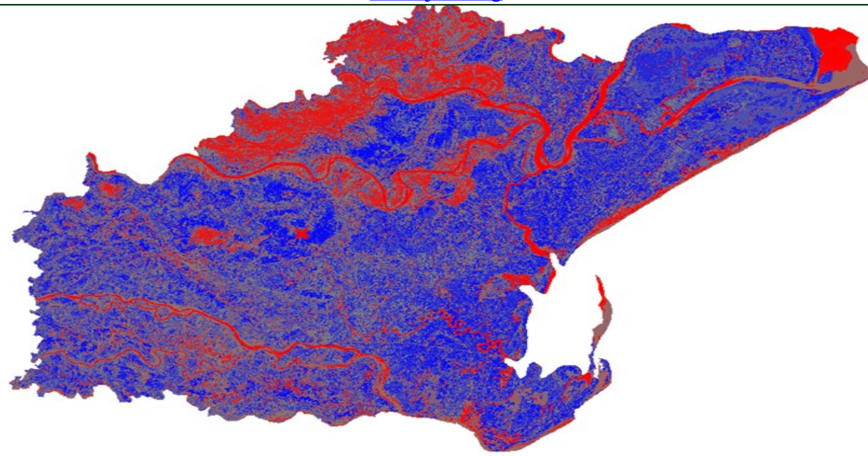


Figure 9: Change Detection Map of Kendrapara Between 1999 to 2020

Table 1: Evaluation Metrics Using Different ML Classifiers

Model	Train Accuracy (%)	Test Accuracy (OA)
XGBoost	68.86	72.09
SVM	69.09	73.39
SVM + XGBoost	**90**	**86.20%**

The changes occur between 1999 to 2020 for five different classes are shown in Table 2. The rows represent the reference class, and the column name represents the new class. The values in the table show the area in square kilometres changes from the corresponding reference class to the new class.

Table 2: Land Cover Change Matrix [KM^2]

Ref. Class	New Class				
	Bare land	**Urban ization**	**Forest**	**Agric ulture**	**Water body**
Bare land	161.09	54.63	182.53	129.39	47.27
Urbani zation	82.12	123.94	348.57	150.08	47.61
Forest	16.14	12.41	150.66	28.44	12.2
Agricul ture	131.94	52.68	226.43	162.26	68.96
Water-body	22.98	15.25	36.83	123.65	142.43
Total	414.26	258.91	945.02	593.84	318.47

5. CONCLUSION

The classification of land use and land cover can help us better understand the dynamics of the city. Maximum Likelihood classifiers are commonly employed, although they fall short of achieving the level of accuracy required for reliable classifying. In this study, multiple ML models were used to classify LULC based on pixel value. The researcher will be able to identify the best classifiers and other evaluation measures as a result of this work. The combination of SVM and XGBoost outperforms competing models. Landsat 7/8 geographical data and atmospheric correction substantially improve the accuracy of LULC classification. When the classified map is used for change detection, we are able to gain knowledge regarding the changes that took place between the years 1999 and 2020. Land cover change metrics represent the class-wise changes in square kilometres. The change detection tool will be extremely helpful in preparing for changes that are to come in the future.

REFERENCES:

[1] S. P. Maurya and A. K. Yadav, "Evaluation of course change detection of Ramganga river using remote sensing and GIS, India," *Weather Clim. Extrem.*, vol. 13, pp. 68–72, 2016, doi: 10.1016/j.wace.2016.08.001.

[2] S. Kaliraj, N. Chandrasekar, and K. K. Ramachandran, "Mapping of coastal landforms and volumetric change analysis in the south west coast of Kanyakumari, South India using remote sensing and GIS techniques," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 20, no. 2, pp. 265–282, 2017, doi: 10.1016/j.ejrs.2016.12.006.

[3] S. Swetanisha, A. R. Panda, and D. K. Behera, "Change Detection using Machine Learning Models: A Case Study on the Puri

District of Odisha, India," in *2021 19th OITS International Conference on Information Technology (OCIT)*, 2022, pp. 100–104, doi: 10.1109/ocit53463.2021.00030.

[4] S. Ahmed, "Assessment of urban heat islands and impact of climate change on socioeconomic over Suez Governorate using remote sensing and GIS techniques," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 21, no. 1, pp. 15–25, 2018, doi: 10.1016/j.ejrs.2017.08.001.

[5] A. Jamali, "Land use land cover modeling using optimized machine learning classifiers: a case study of Shiraz, Iran," *Model. Earth Syst. Environ.*, vol. 7, no. 3, pp. 1539–1550, Sep. 2021, doi: 10.1007/s40808-020-00859-x.

[6] I. R. Hegazy and M. R. Kaloop, "Monitoring urban growth and land use change detection with GIS and remote sensing techniques in Daqahlia governorate Egypt," *Int. J. Sustain. Built Environ.*, vol. 4, no. 1, pp. 117–124, 2015, doi: 10.1016/j.ijsbe.2015.02.005.

[7] C. Voute, "Remote sensing.," *ITC J.*, vol. 1982–1, pp. 37–44, 1982, doi: 10.4324/9781315610139-13.

[8] S. R. KALSI, "Orissa super cyclone – A Synopsis," *Mausam*, vol. 57, no. 1, pp. 1–20, 2006, doi: 10.54302/mausam.v57i1.449.

[9] M. Panigrahi, D. K. Behera, and K. C. Patra, "Epileptic seizure classification of electroencephalogram signals using extreme gradient boosting classifier," *Indones. J. Electr. Eng. Comput. Sci.*, vol. 25, no. 2, p. 884, 2022, doi: 10.11591/ijeecs.v25.i2.pp884-891.

[10] A. V. Rajeswari, S. Saritha, and G. S. Kumar, "Classification based land use/land cover change detection through Landsat images," in *Proceedings - 2014 International Conference on Data Science and Engineering, ICDSE 2014*, 2014, pp. 232–237, doi: 10.1109/ICDSE.2014.6974644.

[11] S. Swetanisha, A. R. Panda, and D. K. Behera, "Land use/land cover classification using machine learning models," *Int. J. Electr. Comput. Eng.*, vol. 12, no. 2, pp. 2040–2046, 2022, doi: 10.11591/ijece.v12i2.pp2040-2046.

[12] D. R. Panuju, D. J. Paull, and A. L. Griffin, "Change detection techniques based on multispectral images for investigating land cover dynamics," *Remote Sens.*, vol. 12, no. 11, pp. 1–36, 2020, doi: 10.3390/rs12111781.

[13] Z. Lv, T. Liu, C. Shi, J. A. Benediktsson, and H. Du, "Novel Land Cover Change Detection Method Based on k-Means Clustering and Adaptive Majority Voting Using Bitemporal Remote Sensing Images," *IEEE Access*, vol. 7, pp. 34425–34437, 2019, doi: 10.1109/ACCESS.2019.2892648.

[14] J. Jiang and G. Tian, "Analysis of the impact of Land use/Land cover change on Land Surface Temperature with Remote Sensing," *Procedia Environ. Sci.*, vol. 2, no. 5, pp. 571–575, 2010, doi: 10.1016/j.proenv.2010.10.062.

[15] B. Demir, F. Bovolo, and L. Bruzzone, "Updating land-cover maps by classification of image time series: A novel change-detection-driven transfer learning approach," *IEEE Trans. Geosci. Remote Sens.*, vol. 51, no. 1, pp. 300–312, 2013, doi: 10.1109/TGRS.2012.2195727.

[16] T. Yumin, B. Bingxin, and M. S. Mohammad, "Time series remote sensing based dynamic monitoring of land use and land cover change," *4th Int. Work. Earth Obs. Remote Sens. Appl. EORSA 2016 - Proc.*, pp. 202–206, 2016, doi: 10.1109/EORSA.2016.7552797.

[17] P. Singh, N. Kikon, and P. Verma, "Impact of land use change and urbanization on urban heat island in Lucknow city, Central India. A remote sensing based estimate," *Sustain. Cities Soc.*, vol. 32, pp. 100–114, 2017, doi: 10.1016/j.scs.2017.02.018.

[18] M. W. Naikoo, M. Rihan, M. Ishtiaque, and Shahfahad, "Analyses of land use land cover (LULC) change and built-up expansion in the suburb of a metropolitan city: Spatio-temporal analysis of Delhi NCR using landsat datasets," *J. Urban Manag.*, vol. 9, no. 3, pp. 347–359, Sep. 2020, doi: 10.1016/j.jum.2020.05.004.

[19] W. C. Olding, J. C. Olivier, B. P. Salmon, and W. Kleynhans, "A Forecasting Approach to Online Change Detection in Land Cover Time Series," *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.*, vol. 12, no. 5, pp. 1451–1460, 2019, doi: 10.1109/JSTARS.2019.2905594.

[20] M. Muller, S. Vincent, O. P. Kumar, and S. Vincent, "Prediction of land-change using machine learning for the deforestation in Paraguay," *Bull. Electr. Eng. Informatics*, vol. 9, no. 5, pp. 1774–1782, 2020, doi: 10.11591/eei.v9i5.2532.

[21] P. K. Mishra, A. Rai, and S. C. Rai, "Land use and land cover change detection using geospatial techniques in the Sikkim Himalaya, India," *Egypt. J. Remote Sens. Sp. Sci.*, vol. 23, no. 2, pp. 133–143, 2020, doi:

ISSN: 1992-8645 www.jatit.org E-ISSN: 1817-3195

10.1016/j.ejrs.2019.02.001.

[22] D. K. Behera, M. Das, S. Swetanisha, and J. Nayak, "XGBoost regression model-based electricity tariff plan recommendation in smart grid environment," *Int. J. Innov. Comput. Appl.*, vol. 13, no. 2, pp. 79–87, 2022, doi: 10.1504/IJICA.2022.123223.

[23] D. K. Behera, M. Das, S. Swetanisha, J. Nayak, S. Vimal, and B. Naik, "Follower Link Prediction Using the XGBoost Classification Model with Multiple Graph Features," *Wirel. Pers. Commun.*, pp. 1–20, 2021, doi: 10.1007/s11277-021-08399-y.