Comparing Machine Learning Algorithms in Land Use Land Cover Classification of Landsat 8 (OLI) imagery.

ABSTRACT

In recent times, there have been increased rates at which researchers are searching for advanced ways of carrying out land-use land-cover (LULC) mapping, especially in developing countries. Four machine-learning algorithms, namely Random Forest (RF), Support Vector Machine (SVM), K-Nearest Neighbour (K-NN), and Gaussian Mixture Models (GMM) were examined. This study also attempted to validate the various models using the index-based validation method. Accuracy assessment was performed by using the Kappa coefficient. The results of the LULC showed that RF classified 23% of the study area as bare land, SVM has 24% of the study area classified as bare land, K-NN also allotted 24% to bare land, while that of GMM classifier was 30%. The overall accuracy of RF, SVM, K-NN and GMM were 0.9840, 0.9780, 0.9641 and 0.9421 respectively. The Kappa Coefficient of the various classifiers were RF (0.9695), SVM (0.9580), K-NN (0.9319) and GMM (0.8916). This study showed that though all the algorithms performed relatively very well, RF performed better than the other classifiers. It suffices to state that, there is a need for further studies since other extraneous environmental variables may be underpinning these conclusions.

Keywords: Supervised machine learning, Algorithm, Kappa Coefficient, classification

1. INTRODUCTION

Image classification defines phenomena in an image based on their spectral signatures, considered as a function wavelength. Mapping of land use land cover (LULC) dynamics has been identified as an integral part of a wide range of geospatial activities and applications [1]. Rapid and uncontrolled population growth with associated economic and industrial development, especially in developing countries with intensified LULC have become underpinning reasons for evaluating changes in LULC [2,3]. Changes in LULC have a series of impacts on the environment in many ways such as increased flood, drought vulnerability, soil degradation, loss of ecosystem services, groundwater depletion, landslide hazards, soil erosion and others [4,5,6]. Over the years, researchers had deployed conventional and direct ways of mapping at various scales integrating spatial information with different levels of precision, which were laborious, time-consuming and expensive in mapping large areas [7].

On the other hand, the satellite-based mapping of LULC has proven to be more cost-effective, spatially extensive, multi-temporal, and time-saving [8]. With the advancement in remote-sensing (RS) techniques satellites now provide data at various spatial and temporal scales [9,10]. Satellite images also have the advantages of multi-temporal availability as well as large spatial coverage for the

LULC mapping [11,12]. In recent times, machine-learning algorithms on remotely-sensed imageries for LULC mapping have been attracting considerable attention [13,14]. Remote sensing techniques have advanced in recent decades, and numerous approaches for land use land cover (LULC) change research, such as support vector machines (SVMs), random forests (RFs), and convolutional neural networks, have been created, including machine learning (CNNs). In LULC classification applications, nonparametric machine learning methods such as SVM and RF are well-known for their ideal classification accuracies [60-62]. These algorithms have similar advantages and abilities in classifying multitemporal and multi-sensor data, such as high-dimensional datasets, and enhanced overall accuracy [63]. In LULC classification applications, nonparametric machine learning methods such as SVM and RF are well-known for their ideal classification accuracies [60-62]. These algorithms have similar advantages and abilities in classifying multitemporal and multi-sensor data, such as high-dimensional datasets, and enhanced overall accuracy [63]. Therefore, researchers have been deploying various classification algorithms in the fields of Remote Sensing and Geographic Information systems (GIS). They include parametric algorithms such as maximum likelihood [15], machine learning algorithms such as Random Forest RF) Artificial Neural Networks (ANNs) and Support Vector Machine (SVM) [16,17]. Machine-learning algorithms have been grouped into two categories; supervised and unsupervised techniques [18].Therefore, researchers have been deploying various classification algorithms in the fields of Remote Sensing and Geographic Information systems (GIS). They include parametric algorithms such as maximum likelihood [15], machine learning algorithms such as Random Forest RF) Artificial Neural Networks (ANNs) and Support Vector Machine (SVM) [16,17]. Machine-learning algorithms have been grouped into two categories; supervised and unsupervised techniques [18]. Examples of the supervised classification techniques include Spectral Angle Mapper (SAM), Support Vector Machine (SVM), Random Forest (RF), Mahalanobis Distance (MD), Fuzzy Adaptive Resonance Theory-Supervised Predictive Mapping (Fuzzy ARTMAP), Radial Basis Function (RBF), Naive Bayes (NB), Decision Tree (DT), K-Nearest Neighbour (K-NN), Gaussian Mixture Models (GMM), Multilayer Perception (MLP), Maximum likelihood classifier (MLC), and Fuzzy Logic [19,20].

Conversely, the unsupervised classification techniques include Affinity Propagation (AP) Cluster Algorithm, Fuzzy C-Means algorithms, K-Means algorithm, ISODATA (iterative self-organizing data) etc. [21,14]. Thus, numerous studies on the LULC modelling have been carried out using different machine-learning algorithms [22,23,24] as well as comparing the machine-learning algorithms [25,26,27,28]. Other factors aside from the type of machine learning algorithm used for lulc classification, can affect its accuracy. It must be mentioned firmly. LULC classification utilizing medium-resolution and low-resolution satellites does not have numerous spectral and spatial constraints that impair its accuracy, according to several studies. While there have been some mirrors research on land use classification using a machine learning algorithm [1,31] There hasn't been much work done on a comparative examination of the various models. As a result, the goal of this research is to determine whether machine learning languages can build a

higher precision LULC map based on accurate statistics

2. MATERIALS AND METHODS

Ileiloju/Okeigbo Local Government Area (study area) in Ondo state lies between Longitudes 6^{o} 40' and 7^{o} 14' N and Latitudes 4^{o} 38' E and 4^{o} 53' E [32]. It shares boundaries with Ondo town, Idanre and Ipetu Ijesha. In the study area, towns and villages such as Agunla, Akinsulure, Oloronba, Awopeju, Oloruntele, Bamkemo, Lisamikan and Ileoluji are notable. It covers a total area of about 698 km^2 with an average temperature of 26^{o}C. The topography is inundated with hills such as the Ikeji and Otasun hills. The average temperature is 26^{o}C with a relative humidity of about 66%. The study area has rivers such as Oni, Okurughu and Awo rivers flowing across the local government area in terms of the drainage system. The economy of the study area is based on the cultivation of crops such as cassava, yam, and cash crops such as oil palm, cocoa, rubber, and kola nut (https://www.manpower.com.ng>lga). It must be stated that this study did not cover the entire local government area but mainly the northern part of the local government. The selection of this part of the map was premised on the fact that the focus of this study is on the performance of different machine learning classification algorithms and not on a land use land cover change detection analysis (Figure 1).

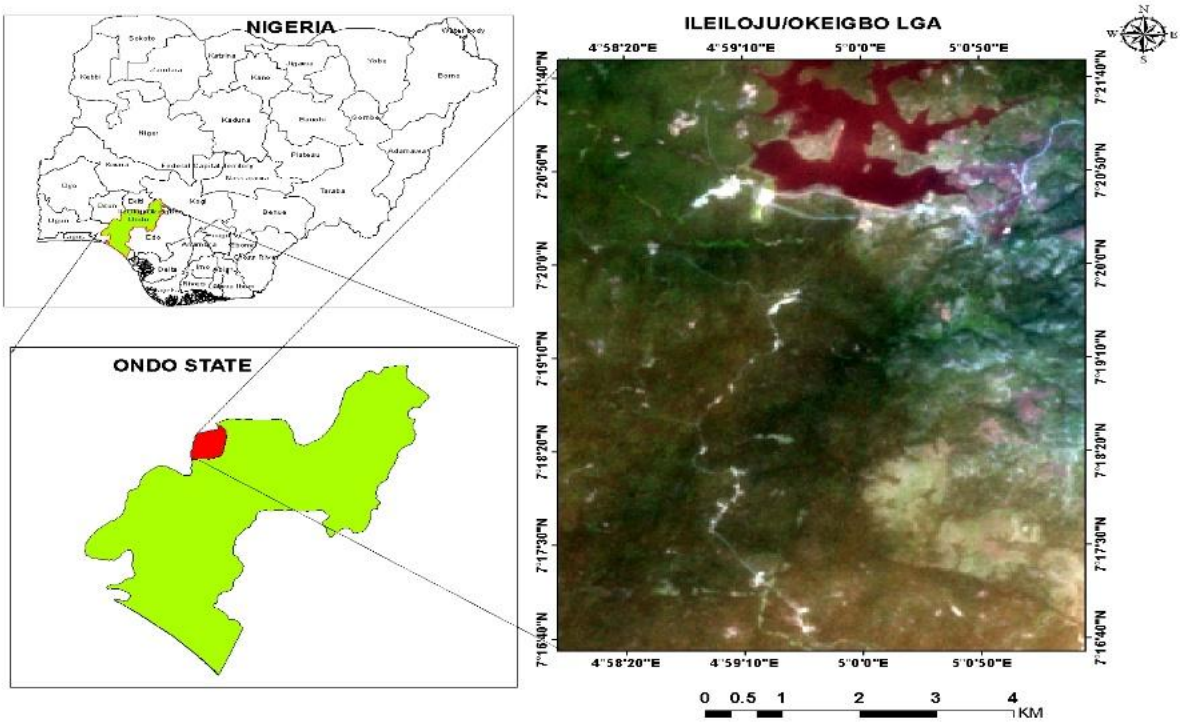

Figure 1: Map of the study area (Source: Political Map of Nigeria: https//www.worldometers.info)

2.1. Materials

The Landsat 8 Operational Land Imager (OLI) image of November 25th, 2021 (path/row 190/055) was downloaded from the United States Geological Survey (USGS) website (https://earthexplorer.usgs.gov). The Google Earth image coupled with some ground control points (GCP) was used to assess the classified LULC maps accurately.

2.2. Preprocessing

An atmospheric correction is a prerequisite for image pre-processing. In this study, the Dark Object Subtraction (DOS) Algorithm in QGIS 3.22 using the (SCP plugin) was deployed for the image correction. The dark object subtraction method operates by removing the effects of scattering from the image data. It is unique because it derives the corrected DN (Digital Number) values majorly from the digital data without relying on outside information [33]. Dark-object subtraction (DOS) is one of the most widely used methods when it comes to reducing haze within an image. Most dark object subtraction technique assumes that there is a high probability that at least a few pixels within an image should be black (0% Reflectance) [33]. The (DOS) method assumes that within a satellite image, there exist features that have near-zero per cent reflectance (i.e., water, dense forest, shadow), such that the signal recorded by the sensor from these features is solely a result of atmospheric scattering (path radiance), which must be removed [34,35]. Like similar research [36], this study utilized seven atmospherically corrected L8 OLI/TIRS spectral bands (Table 1).

Table 1. Landsat 8 (OLI) bands and their wavelength

Bands	Wavelength (micrometres)
Band 1 - Coastal aerosol	0.43 - 0.45
Band 2 - Blue	0.45 - 0.51
Band 3 - Green	0.53 - 0.59
Band 4 - Red	0.64 - 0.67
Band 5 - Near Infrared (NIR)	0.85 - 0.88
Band 6 - SWIR 1	1.57 - 1.65
Band 7 - SWIR 2	2.11 - 2.29

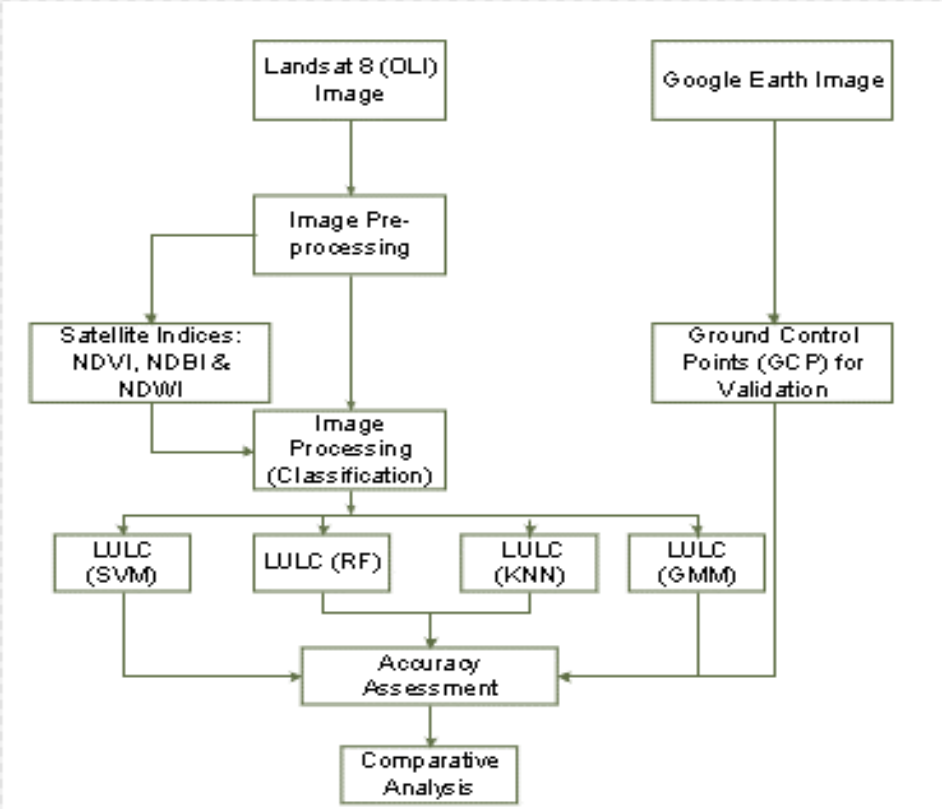

Figure 2: Flowchart showing the methodology deployed for this study

2.3. Random Forest

Random Forest (RF) is a new non-parametric ensemble machine-learning algorithm developed by Breiman [37]. It is unique because it can handle a variety of data, such as satellite imageries and numerical data [38]. RF is an ensemble learning algorithm premised on a decision tree, which integrates massive ensemble regression and classification trees. Several studies have shown a satisfactory performance for LULC classification using RF in the field of remote-sensing applications [13,19,27]. The higher the number of trees involved in this method the better the accuracy in the image classification and land use modelling [39,40].

2.4. Support Vector Machine

Support Vector Machine (SVM) is a non-parametric supervised machine learning method to solve binary classification problems [14]. In remote sensing, the polynomial and radial basis function (RBF) kernel has been used most commonly. However, for LULC classification, RBF is the most popular technique, and it produces better accuracy than the other traditional methods [14]. The objective of the original SVM technique was to find the hyper-plane that can separate datasets into several classes and find the optimal separating hyper-plane from the available hyper-planes [41]. In this process, the vectors ensure that the width of the margin will be maximized [42]. The training

samples or bordering samples that delineate the margin or hyper-plane of SVM are known as support vectors [20]. The operational capacity of the SVM is a function of the kernel size and density. Therefore, the differential between the simulated and the actual satellite data shows the best performance using the support vectors [43]. The SVM was performed in QGIS 3.22 using the dzetsaka plugin.

2.4. K-Nearest Neighbour classifier

K-nearest neighbour (KNN) algorithm [37] is a method for classifying objects based on closest training examples in the feature space. K-nearest neighbour algorithm is among the simplest of all machine learning algorithms. In the classification process, the unlabeled query point is simply assigned to the label of its k-nearest neighbours. K-NN uses k-nearest neighbours from a subset of all of the training samples in determining a pixel's class or the degree of membership of a class. The selection of different values for 'K' can generate different classification results for the same sample object. KNN is a simple classification technique. KNN is used to classify the objects based on their similarity or closest training samples in the feature space [44].

Table 2: Relevant spectral indices and their formulas

INDEX	FORMULA	REFERENCE
Normalized Difference Vegetation Index	$\mathrm{NDVI} = \frac{\mathrm{Green}-\mathrm{NIR}}{\mathrm{Green}+\mathrm{NIR}}$	[55]
Normalized Difference Built-up Index	$\mathrm{NDBI} = \frac{\mathrm{Green}-\mathrm{NIR}}{\mathrm{Green}+\mathrm{NIR}}$	[56]
Normalized Difference Water Index	$\mathrm{NDWI} = \frac{\mathrm{Green}-\mathrm{NIR}}{\mathrm{Green}+\mathrm{NIR}}$	[54]

2.5. Gaussian Mixture Models

A Gaussian mixture model (GMM) is useful for modelling data that comes from one of several groups. The groups might be different from each other, but data points within the same group can be well-modelled by a Gaussian distribution.

2.6. Validation of machine learning classifiers

To validate the results derivable from this study, the "index-based technique" has been chosen to select the best performing machine-learning technique for LULC mapping. For this purpose, three satellite-based indices; Normalized Difference Vegetation Index (NDVI), Normalized Differential Water Index (NDWI) and Normalized Difference Built-up Index (NDBI), have been classified using different thresholds (Table 2). In the end, the area extent of the classifier-derived LULC will be statistically compared to the index-derived area extent.

2.6. Accuracy Assessment

The post-classification accuracy assessment of the LULC generated using various models has become an integral part of the classification process [45]. The Kappa coefficient statistical technique was deployed in this study to assess accuracy. Monserud and Leemans [46] suggested five levels

of accuracy results: very poor (< 0.4), fair (0.4 to 0.55), good (0.55 to 0.70), very good (0.70 to 0.85) and excellent (> 0.85). Thus, the Kappa coefficient was calculated using 501 randomly selected sample points to evaluate the accuracy of LULC maps generated using different algorithms. The reference data was downloaded using Google Earth Pro.

3. RESULT AND DISCUSSION

3.1. LULC Classification

In this regard, image classification is based on the four advanced mathematical and machine learning algorithms including Random Forest, Support Vector Machine, K-Nearest Neighbour and the Gaussian Mixture Models. Landsat 8 (OLI/TIR) image was classified into four thematic classes: The Settlement, Bare land, Vegetation, and Waterbody. The study area is about 9,031 ha. From Table 2, out of the total area under study, the RF classifier classified 392 ha (4%) as Settlement area, 2015 ha (23%) as Bare land, 6264 ha (69%) as Vegetation and 360 ha (4%) as Waterbody. The SVM classifier classified 286 ha (3%) as Settlement, 2136 ha (24%) as Bare land, 6242 ha (69%) as Vegetation and 367 ha (4%) as Waterbody. Also, 359 ha (4%) were classified as Settlement, 2153 (24%) as Bare land,6142 (68%) as Vegetation, and 378 (4%) as Waterbody by K-NN classifier. GMM classifier had 949 ha (10%) classified as Settlement, 2732 ha (30%) as Bare land,5019 ha (56%) as Vegetation and 331 ha (4%) as Waterbody. The LULC maps in Figure 3 showed that the settlement area, as classified by RF (4%), SVM (3%) and K-NN (4%) are very similar. GMM, using the same image and training samples classified 10% of the study area as settlement. With a sharp difference of about 6%, the GMM classifier tends to differ in algorithmic operations when compared to other classifiers. RF classified 23% of the study area as bare land, SVM has 24% of the study area classified as bare land. K-NN also allotted 24% to bare land, while that of the GMM classifier was 30%. Vegetation thematic class has almost the same classified area extends across the four different classifiers i.e. RF (69%), SVM (69%), K-NN (68%) and GMM (56%) which is the least coverage when compared to other classifiers. Waterbody was classified as 4% by all the classifiers (Table 2).

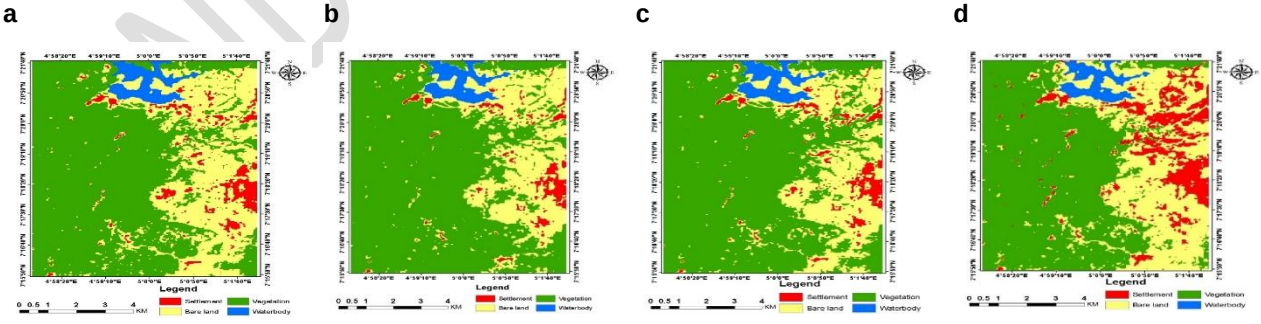

Figure 3: LULC of the study area using different classifiers: (a) with Random Forest (RF) (b) with Support Vector Machines (SVM) (c) with K-Nearest Neighbour (KNN) (d) with Gaussian Mixture Model (GMM).

Table 2 shows the percentage share of each LULC class concerning each classifier's total land coverage in the study area.

	Random Forest (RF)		Support Vector Machine (SVM)		K-Nearest Neighbour (KNN)		Gaussian Mixture Model (GMM)	
Classes	Area(ha)	%	Area(ha)	%	Area(ha)	%	Area(ha)	%
Settlement	392	4	286	3	359	4	949	10
Bare land	2015	23	2136	24	2153	24	2732	30
Vegetation	6264	69	6242	69	6142	68	5019	56
Waterbody	360	4	367	4	378	4	331	4

It is a fact according to [47] that LULC classes cannot be thematically equal amongst the classification techniques, be it machine-learning algorithms or traditional classification techniques. This explained why the area extent of the various LULC classes as shown in Table 2 are different from one classifier to another. Differences in the parameter optimization of the algorithms can also be responsible for the differences in area under LULC classes of different classifiers [48]. Though the studies of [13] and [27] opined that the machine-learning techniques do not have significant differences in the results, this study revealed that there could be significant differences in the LULC results of the different classifiers.

3.2. Validation of models using index-derived techniques

The results in Table 3 show the comparison between the spectral indices-derived area extent and that of the LULC derived from the classifiers. Figure 4 shows the reclassified maps of the NDVI, NDBI and the NDWI. The total area of NDBI-based is 2339 ha compared to settlement/bare land area as classified by RF classifier which is 2407 ha, with a difference of -67 ha. It shows that they are both close when compared to that of SVM (2422 ha), K-NN (2512 ha) and GMM (3681 ha) respectively. The NDVI-based vegetation area remained 6253 ha while that of the RF classifier stood at 6264 ha with a difference of -11 ha. The total vegetation area extent as classified by other classifiers are SVM (6242 ha), K-NN (111 ha) and GMM (1234 ha) respectively. Waterbody area calculated using the NDWI was 365 ha, while that of the RF classifier was 360 ha (Table 3).

Table 3. Area of LULC computed by the spectral indices and the computed areas of the LULC by the Machine Learning (ML) algorithms.

CLASS	Spectral Indices (ha)	Area (ha) computed by algorithms and their differences with spectral indices			
		RF	SVM	K-NN	GMM
Settlement/Bare land	2339	2407 (-67)	2422 (-88)	2512 (-173)	3681 (-1342)
Vegetation	6253	6264 (-11)	6242 (11)	6142 (111)	5019 (1234)

Waterbody	365	360 (5)	367 (-2)	378 (-13)	331 (34)

*Values within parenthesis indicate the difference between the area computed in spectral indices and that of the classification algorithms.

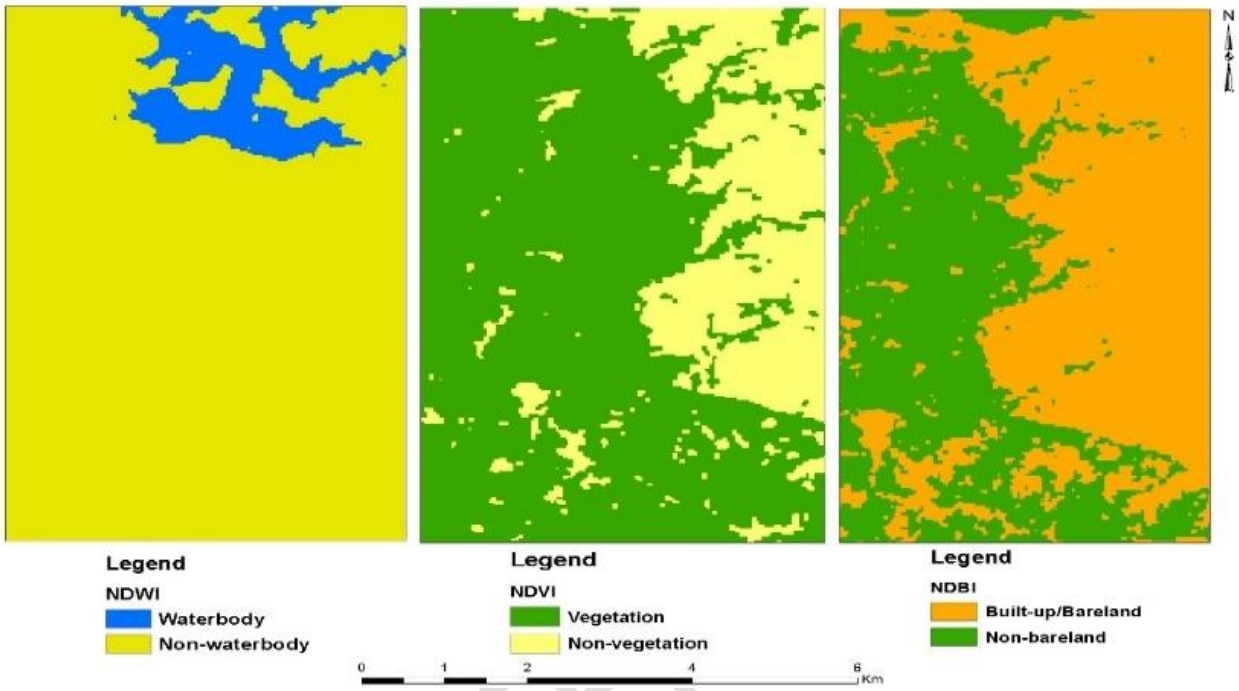

Figure 4: The index-derived maps of NDVI, NDWI and NDBI

3.3. Accuracy Assessment of the classified LULC

To validate these models' accuracy, 501 random points were generated on the classified images which contain classified information. These points were then observed with the actual ground data extracted from google earth historical imagery 2021. The extracted values from the classified image vis-a-vis the reference data (google earth image) were used to calculate the error matrix, overall accuracy and Kappa coefficients of the four classifiers. Tables 4, 5, 6, 7and 8 showed the error matrices of the various classifiers and the Producer Accuracy. The producer accuracy of Settlement as classified by RF (0.9921) is the highest when compared to other classifiers, while the rest of the three classifiers (SVM, K-NN, and GMM) had approximately 0.9545. The user accuracy of the settlement class had RF (0.9167), SVM (0.9130), K-NN (0.8077) and GMM (0.6774). It showed that settlement was accurately classified by RF, but poorly classified by GMM. The results are almost the same pattern as the other classes (Table 8). The Overall Accuracy (OA) and Kappa Coefficient (K) for all the classifiers are shown in Table 9. The overall accuracy of RF, SVM, K-NN and GMM are 0.9840, 0.9780, 0.9641 and 0.9421 respectively. This indicates that the classifier's performance in terms of overall accuracy was very similar. The Kappa Coefficient results of the various classifiers RF (0.9695), SVM (0.9580), K-NN (0.9319) and GMM (0.8916) showed that RF was the most accurate of all the classifiers. It suffices to state that other classifiers also performed very high

when compared to [46] Kappa Coefficient benchmark of 0.85 as excellent performance. Nevertheless, there appeared to be an excellent agreement between the classified LULC map and the reality on the ground. It has been found that SVM and RF generally provide better accuracy when compared to other traditional classifiers. Some researchers have submitted that SVM and RF are the best techniques for the LULC classification compared to all other machine-learning techniques [16,19]. This study revealed that though, all the machine learning classifiers are very good in terms of LULC classification, the Random Forest is still highly recommended.

Table 4. Error matrix for RF

Error Matrix		Observed				Total
Classified	**RF**	Settlement	Bare land	Vegetation	Waterbody	
	Settlement	22	1	0	1	24
	Bare land	0	143	2	0	145
	Vegetation	0	0	310	1	311
	Waterbody	0	1	2	18	21
Total		22	145	314	20	501

Table 5. Error matrix for SVM

Error Matrix		Observed				Total
Classified	**SVM**	Settlement	Bare land	Vegetation	Waterbody	
	Settlement	21	1	0	1	23
	Bare land	0	143	2	1	146
	Vegetation	1	0	309	1	311
	Waterbody	0	1	3	17	21
Total		22	145	314	20	501

Table 6. Error matrix for K-NN

Error Matrix		Observed				Total
Classified	**K-NN**	Settlement	Bare land	Vegetation	Waterbody	
	Settlement	21	3	1	1	26
	Bare land	0	140	5	1	146

	Vegetation	1	0	305	1	307
	Waterbody	0	2	3	17	22
	Total	22	145	314	20	501

Table 7: Error matrix for GMM

Error Matrix		Observed				Total
Classified	**GMM**	Settlement	Bare land	Vegetation	Waterbody	
	Settlement	21	5	4	1	31
	Bare land	0	136	8	1	145
	Vegetation	1	2	298	1	302
	Waterbody	0	2	4	17	23
	Total	22	145	314	20	501

Table 8: LULC Accuracy Assessment statistics of the classifiers

	RF		**SVM**		**K-NN**		**GMM**	
Classes	***Pa***	***Ua***	***Pa***	***Ua***	***Pa***	***Ua***	***Pa***	***Ua***
Settlement	0.9921	0.9167	0.9545	0.9130	0.9545	0.8077	0.9545	0.6774
Bare land	0.9862	0.9862	0.9862	0.9795	0.9655	0.9589	0.9379	0.9379
Vegetation	0.9872	0.9968	0.9841	0.9936	0.9713	0.9935	0.9490	0.9867
Waterbody	0.9	0.8571	0.85	0.8095	0.85	0.772727	0.8501	0.7391

Table 9. Summary of LULC Accuracy Assessment Results

Classifier	Overall Accuracy (OA)	Kappa Coefficient (K)
Random Forest	0.9840	0.9695
Support Vector Machine	0.9780	0.9580
K-Nearest Neighbour	0.9641	0.9319
Gaussian Mixture Model	0.9421	0.8916

The accuracy assessment in this study revealed an insignificant variation among the results of the classifiers. Therefore, comparing this study with some previous studies, the accuracy of LULC classification varied from one classifier to another sequel to variations in methods, techniques, time, and space [49,14,27].

The result of Guanyao Xie and Simona Niculescu, 2021 [57] in consonance with this study showed, that RF and SVM models both performed well for LCLU classification; however, the accuracy assessments show that the SVM is better suited to classification. For urban LCLU classification, Jozdani et al., 2019 [59] deployed machine

learning with Object-based Image Analysis (OBIA) methodologies. Though we did not integrate OBIA in our study, his results also support the high-performance rating of SVM and RF. LeCUN et al., (2015) [58] compared Convolution Neural Network (CNN) with other machine learning algorithms in terms of classification accuracy. In his results, the CNN outperforms other state-of-the-art machine learning classifiers. However, there are certain essential considerations to be made about its effectiveness. In comparison to RF models and SVMs, previous applications of CNN models have tended to stress their complexity. In this situation, cross-validation is frequently used to tune and optimize parameters for CNN algorithms. Variations in the classification outputs could be traceable to the influence of atmospheric, surface and illumination characteristics of the images [26]. It is pertinent to state that some other studies had reported that there are minor to moderate fluctuations in the accuracy of the LULC classification using different classifiers [50,51]. The high accuracy performance of the RF classifier in this study with a Kappa coefficient of 0.97 is furthersupported with previous studies such as that of [13] and [19] with accuracy levels 0.93 and 0.90, respectively, for the RF classifier. A small difference is found between the previous study and this study on the accuracy levels of SVM [52,53]. Furthermore, [26] noted that the accuracy of SVM and RF has very little difference, but the difference increases between either SVM and K-NN.

4. CONCLUSIONS

This study examined the accuracy of four different machine-learning classifiers for LULC classification using Landsat 8 (OLI/TIR) satellite image to elicit the best of all the classifiers. Settlement, bare terrain, vegetation, and waterbody were recognized as four distinct classes. The results revealed that under different classifiers, the area coverage of each LULC class varied. An accuracy assessment of the LULC classification was performed, with the total accuracy and Kappa coefficient as statistical metrics for comparison. In the end, the Kappa coefficient and overall coefficient showed changes in the accuracy of each LULC classifier. Both Kappa coefficient and Overall accuracy analysis showed that RF has the highest accuracy of all classifiers applied to LULC modelling in the study area.

REFERENCES

1. Dutta D, Rahman A, Paul SK, Kundu A. Changing pattern of the urban landscape and its effect on land surface temperature in and around Delhi. Environ. Monit. Assess. 2019; 191: 551. [CrossRef]

2. Hoan NT, Liou YA, Nguyen KA, Sharma RC, Tran DP, Liou CL, Cham DD. Assessing the Effects of Land-Use Types in Surface Urban Heat Islands for Developing Comfortable Living in Hanoi City. Remote Sens. 2018; 10: 1965. [CrossRef]

3. Kumari B, Tayyab M, Hang HT, Khan MF, Rahman A. Assessment of public open spaces (POS) and landscape quality based on per capita POS index in Delhi, India. SN Appl. Sci. 2019; 1: 368.

4. Cerovski-Darriau C, Roering JJ. Influence of anthropogenic land-use change on hillslope erosion in the Waipaoa River Basin, New Zealand. Earth Surf. Process. Landf. 2016; 41: 2167–2176.

5. Pal S, Kundu S, Mahato S. Groundwater potential zones for sustainable management plans in a river basin of India and Bangladesh. J. Clean. Prod. 2020; 257: 120311. [CrossRef]

6. Mahato S, Pal S. Groundwater potential mapping in a rural river basin by union (OR) and intersection (AND) of four multi-criteria decision-making models. Nat. Resour. Res. 2019; 28: 523–545.

7. Langat PK, Kumar L, Koech R, Ghosh MK. Monitoring of land use/land-cover dynamics using

remote sensing: A case of Tana River Basin, Kenya. Geocarto Int. 2019.

8. Hoffmann J. The future of satellite remote sensing in hydrogeology. Hydrogeol. J. 2005; 13: 247–250. [CrossRef]

9. Scaioni M, Longoni L, Melillo V, Papini M. Remote Sensing for Landslide Investigations: An Overview of Recent Achievements and Perspectives. Remote Sens. 2014; 6: 9600–9652. [CrossRef]

10. Liou YA, Wu AM, Lin HY. FORMOSAT-2 Quick Imaging. In Optical Payloads for Space Missions; Qian SE., Ed.; Wiley: Oxford, UK, 2016; 1008, ISBN 9781118945148.

11. Wittke S, Yu X, Karjalainen M, Hyyppä J, Puttonen E. Comparison of two-dimensional multitemporal Sentinel-2 data with three-dimensional remote sensing data sources for forest inventory parameter estimation over a boreal forest. Int. J. Appl. Earth Obs. Geoinf. 2019; 76: 167–178.

12. Viana CM, Girão I, Rocha J. Long-Term Satellite Image Time-Series for Land Use/Land Cover Change Detection Using Refined Open Source Data in a Rural Region. Remote Sens. 2019; 11: 1104. [CrossRef]

13. Adam E, Mutanga O, Odindi J, Abdel-Rahman EM. Land-use/cover classification in a heterogeneous coastal landscape using Rapid Eye imagery: Evaluating the performance of random forest and support vector machines classifiers. Int. J. Remote Sens. 2014; 35: 3440–3458.

14. Maxwell AE, Warner TA, Fang F. Implementation of machine-learning classification in remote sensing: An applied review. Int. J. Remote Sens. 2018; 39: 2784–2817.

15. Otukei JR, Blaschke T. Land cover change assessment using decision trees, support vector machines and maximum likelihood classification algorithms. Int J Appl Earth Obs Geoinf 2010; 12: 27–31

16. Mountrakis G, Im J, Ogole C. Support vector machines in remote sensing: a review. ISPRS J Photogramm Remote Sens 2011; 66(3):247–259

17. Duro DC, Franklin SE, Dubé MG. A comparison of pixel-based and object-based image analysis with selected machine learning algorithms for the classification of agricultural landscapes using SPOT-5 HRG imagery. Remote Sens Environ 2012; 118: 259–272

18. Wu L, Zhu X, Lawes R, Dunkerley D, Zhang H. Comparison of machine learning algorithms for classification of LiDAR points for characterization of canola canopy structure. Int. J. Remote Sens. 2019; 40: 5973–5991.

19. Ma L, Liu Y, Zhang X, Ye Y, Yin G, Johnson BA. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 2019; 152: 166–177. [CrossRef]

20. Shih HC, Stow DA, Tsai YH. Guidance on and comparison of machine learning classifiers for Landsat-based land cover and land use mapping. Int. J. Remote Sens. 2019; 40: 1248–1274.

21. Camps-Valls G, Benediktsson JA, Bruzzone L, Chanussot J. Introduction to the issue on advances in remote sensing image processing. IEEE J. Sel. Top. Signal Process. 2011; 5: 365–369. [CrossRef]

22. Pal M. Random forest classifier for remote sensing classification. Int. J. Remote Sens. 2005; 26: 217–222. [CrossRef]

23. Teluguntla, P.; Thenkabail, P.S.; Oliphant, A.; Xiong, J.; Gumma, M.K.; Congalton, R.G.; Huete, A. A 30-m Landsat-derived cropland extent product of Australia and China using random forest machine learning algorithm on Google Earth Engine cloud computing platform. ISPRS J. Photogramm. Remote Sens. 2018, 144, 325–340.

24. Zhang, C.; Sargent, I.; Pan, X.; Li, H.; Gardiner, A.; Hare, J.; Atkinson, P.M. Joint Deep Learning for land cover and land use classification. Remote Sens. Environ. 2019, 221, 173–187. [CrossRef]

25. Rogan, J.; Franklin, J.; Stow, D.; Miller, J.; Woodcock, C.; Roberts, D. Mapping land-cover modifications over large areas: A comparison of machine learning algorithms. Remote Sens. Environ. 2008, 112, 2272–2283. [CrossRef]

26. Li, X.; Chen, W.; Cheng, X.; Wang, L. A comparison of machine learning algorithms for mapping of complex surface-mined and agricultural landscapes using ZiYuan-3 stereo satellite imagery. Remote Sens. 2016, 8, 514. [CrossRef]

27. Camargo, F.F.; Sano, E.E.; Almeida, C.M.; Mura, J.C.; Almeida, T. A comparative assessment of machine-learning techniques for land use and land cover classification of the Brazilian tropical savanna using ALOS-2/PALSAR-2 polarimetric images. Remote Sens. 2019, 11, 1600. [CrossRef]

28. Jamali, A. Evaluation and comparison of eight machine learning models in land use/land cover mapping using Landsat 8 OLI: A case study of the

northern region of Iran. SN Appl. Sci. 2019, 1, 1448. [CrossRef]

29. Manandhar, R.; Odeh, I.O.; Ancev, T. Improving the accuracy of land use and land cover classification of Landsat data using post-classification enhancement. Remote Sens. 2009, 1, 330–344. [CrossRef]

30. Yang, C.; Wu, G.; Ding, K.; Shi, T.; Li, Q.; Wang, J. Improving land use/land cover classification by integrating pixel unmixing and decision tree methods. Remote Sens. 2017, 9, 1222. [CrossRef]

31. Pal, S.; Ziaul, S.K. Detection of land use and land cover change and land surface temperature in English Bazar urban centre. Egypt. J. Remote Sens. Space Sci. 2017, 20, 125–145. [CrossRef]

32. Adeyeye O.R, Adefusisoye A.A, Awokunmi E.E and Olanipekun E.O.Evaluation of polycyclic aromatic hydrocarbons in water from hand-dug wells at Ile-Oluji, Nigeria*. IOSR Journal of Environmental Science, Toxicology and Food Technology*. 2016 10 (9), 112-119

33. Chavez Jr, P.S. (1988). An improved dark-object subtraction technique for atmospheric scattering correction for multispectral data. Remote Sensing of Environment 24, 459-479

34. Paul, M. Mather; and Brandt, Tso; "Classification Methods for Remotely Sensed Data", Second Edt., CRC Press, Taylor and Francis Group, LLC, pp(5-7, 53-54), (2009).

35. Michael, Köhl; Steen, S. Magnussen; and Marco, Marchetti; "Sampling Methods, Remote Sensing and GIS Multi resource Forest Inventory", Springer, New York, pp (205-207), (2011).

36. Karakizi C, Vakalopoulou M, Karantzalos K (2017) Annual crop type classification from multitemporal Landsat-8 and sentinel-2 data based on deep-learning. In: Proceedings of the 37th international symposium on remote sensing of the environment (ISRSE37), 2017, Tshwane, South Africa

37. Breiman L (2001) Random forests. Mach Learn 45(1):5–32

38. Abdullah, A.Y.M.; Masrur, A.; Adnan, M.S.G.; Baky, M.; Al, A.; Hassan, Q.K.; Dewan, A. Spatio-temporal patterns of land use/land cover change in the heterogeneous coastal region of Bangladesh between 1990 and 2017. Remote Sens. 2019, 11, 790.

39. Liaw, A.; Wiener, M. Classification and Regression by randomForest. R News 2002, 2, 18–22. 85.

40. Feng, Q.; Gong, J.; Liu, J.; Li, Y. Flood mapping based on multiple endmember spectral mixture analysis and random forest classifier—The case of Yuyao, China. Remote Sens. 2015, 7, 12539–12562. [CrossRef]

41. Srivastava, P.K.; Han, D.; Rico-Ramirez, M.A.; Bray, M.; Islam, T. Selection of classification techniques for land use/land cover change investigation. Adv. Space Res. 2012, 50, 1250–1265. [CrossRef]

42. Bouaziz, M.; Eisold, S.; Guermazi, E. Semiautomatic approach for land cover classification: A remote sensing study for the arid climate in Southeastern Tunisia. Euro Mediterr. J. Environ. Integr. 2017, 2, 24. [CrossRef]

43. Mathur, A.; Foody, G.M. Multiclass and binary SVM classification: Implications for training and classification users. IEEE Geosci. Remote Sens. Lett. 2008, 5, 241–245. [CrossRef]

44. Olmanson.L.G.; Bauer.M.E.; Brezonik.P.L. A 20 Year Landsat Water Clarity Census of Minnesota's 10,000 Lakes. Remote Sens. Environ. 2008, 112, 4086–4097.

45. Hurskainen, P.; Adhikari, H.; Siljander, M.; Pellikka, P.K.E.; Hemp, A. Auxiliary datasets improve the accuracy of object-based land use/land cover classification in heterogeneous savanna landscapes. Remote Sens. Environ. 2019, 233, 111354. [CrossRef]

46. Monserud, R.A.; Leemans, R. Comparing global vegetation maps with the Kappa statistic. Ecol. Model. 1992, 62, 275–293. [CrossRef]

47. Abdi, A.M. Land cover and land use classification performance of machine learning algorithms in a boreal landscape using Sentinel-2 data. GISci. Remote Sens. 2019, 1–20. [CrossRef]

48. Erbek, F.S.; Özkan, C.; Taberner, M. Comparison of maximum likelihood classification method with supervised artificial neural network algorithms for land use activities. Int. J. Remote Sens. 2004, 25, 1733–1748. [CrossRef]

49. Rodriguez-Galiano, V.F.; Chica-Rivas, M. Evaluation of different machine learning methods for land cover mapping of a Mediterranean area using multi-seasonal Landsat images and Digital Terrain Models. Int. J. Digit. Earth 2014, 7, 492–509. [CrossRef]

50. Islam, K.; Jashimuddin, M.; Nath, B.; Nath, T.K. Land use classification and change detection by using multi-temporal remotely sensed imagery: The case of Chunati wildlife sanctuary, Bangladesh. Egypt. J. Remote Sens. Space Sci. 2018, 21, 37–47. [CrossRef]

51. Leyk, S.; Uhl, J.H.; Balk, D.; Jones, B. Assessing the accuracy of multi-temporal built-up land layers across rural-urban trajectories in the United States. Remote Sens. Environ. 2018, 204, 898–917. [CrossRef] [PubMed]

52. Szuster, B.W.; Chen, Q.; Borger, M. A comparison of classification techniques to support land cover and land use analysis in tropical coastal zones. Appl. Geogr. 2011, 31, 525–532. [CrossRef]

53. Qian, Y.; Zhou, W.; Yan, J.; Li, W.; Han, L. Comparing machine learning classifiers for object-based land cover classification using very high-resolution imagery. Remote Sens. 2015, 7, 153–168. [CrossRef].

54. Duque, J.C.; Patino, J.E.; Betancourt, A. Exploring the potential of machine learning for automatic slum identification from VHR imagery. Remote Sens. 2017, 9, 895. [CrossRef]

55. Huang Sha, Lina Tang, Joseph P. Hupy Yang Wang, Guofan Shao. A commentary review on the use of normalized difference vegetation index (NDVI) in the era of popular remote sensing. Journal of Forestry Research.2020, 32:45. https://doi.org/10.1007/s11676-020-01155-1

56. Loranty M, Davydov S, Kropp H, Alexander H, Mack M, Natali S, Zimov N. Vegetation indices do not capture forest cover variation in Upland Siberian larch forests. Remote Sens. 2018, 10(11):1686–1700

57. Guanyao Xie and Simona Niculescu (2021) Mapping and Monitoring of Land Cover/Land Use (LCLU) Changes in the Crozon Peninsula (Brittany, France) from 2007 to 2018 by Machine Learning Algorithms (Support Vector Machine, Random Forest, and Convolutional Neural Network) and by Post-classification Comparison). *Remote Sens.* 2021, 13, 3899. https://doi.org/10.3390/rs13193899

58. LeCun, Y.; Bengio, Y.; Hinton, G. Deep learning. Nature 2015, 521, 436–444. [CrossRef]

59. Jozdani, S.E.; Johnson, B.A.; Chen, D. Comparing Deep Neural Networks, Ensemble Classifiers, and Support Vector Machine Algorithms for Object-Based Urban Land Use/Land Cover Classification. Remote Sens. 2019, 11, 1713. [CrossRef]

60. Zafari, A.; Zurita-Milla, R.; Izquierdo-Verdiguier, E. Evaluating the Performance of a Random Forest Kernel for Land Cover Classification. Remote Sens. 2019, 11, 575. [CrossRef]

61. Dee, S.; Yang, X. Support Vector Machines for Land Cover Mapping from Remote Sensor Imagery. In Monitoring and Modeling of Global Changes: A Geomatics Perspective; Li, J., Yang, X., Eds.; Springer Remote Sensing/Photogrammetry; Springer: Dordrecht, The Netherlands, 2015; pp. 265–279. [CrossRef]

62. Sheykhmousa, M.; Mahdianpari, M.; Ghanbari, H.; Mohammadimanesh, F.; Ghamisi, P.; Homayouni, S. Support Vector Machine Versus Random Forest for Remote Sensing Image Classification: A Meta-Analysis and Systematic Review. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2020, 13, 6308–6325. [CrossRef]

63. Song, X.; Duan, Z.; Jiang, X. Comparison of artificial neural networks and support vector machine classifiers for land cover classification in Northern China using a SPOT-5 HRG image. Int. J. Remote Sens. 2012, 33, 3301–3320. [CrossRef]