Kurdistan Journal of Applied Research (KJAR)
Print-ISSN: 2411-7684 | Electronic-ISSN: 2411-7706

Original Article

Evaluation of Different Classification Algorithms for Land Use Land Cover Mapping

Kaifi Chomani [a]*, **Shaki Pshdari** [a]

[a] Civil Engineering Department, University of Raparin, Ranya, Sulaymaniyah, Iraq.

Submitted: 16 April 2024
Revised: 22 May 2024
Accepted: 19 July 2024

***Corresponding Author:**
Kaifi.chomani@gmail.com

Keywords: Image classification, Satellite imagery, LULC; GIS; Google earth engine, Machine learning.

How to cite this paper: K. Chomani, S. Pshdari " Evaluation of Different Classification Algorithms for Land Use Land Cover Mapping", KJAR, Vol. 9, no. 2, pp. 13-23 ,Aug. 2024, doi: 10.24017/science.2024.2.2

Abstract: For efficient sustainable management and monitoring landscape changes over times, reliable land use land cover (LULC) mapping using the most accurate classification algorithms is required. Increasing innovative classification algorithms and satellite data demands finding the most suitable classifier to create accurate maps of different features efficiently. The challenge addressed in this study is to identify the most accurate algorithm for classifying and generating reliable LULC. The objective of this research was to identify the best classification among several algorithms both overall and in each individual class by using ArcGIS Pro and Google Earth Engine with Landsat 8 and Sentinel-2 datasets for Ranya city as the study area. Support vector machine (SVM), maximum likelihood, random tree, classification and regression tree, K-Nearest Neighbor and iterative self organizing cluster algorithms were used to classify the satellite image of the study area. The kappa coefficient matrix was used to assess the performance of each classifier and method. The study showed that the random tree algorithm achieved highest overall accuracy using Sentinel-2 with 83%. Meanwhile, when the specific class accuracy is priority, the result suggests the use of SVM algorithm using Sentinel-2 for building footprint extraction with 92% accuracy. The result also showed that the outcomes of most algorithms were better using Sentinel-2 rather than Landsat 8, making Sentinel-2 more suitable for accurate LULC mapping. The outcomes of the research assessed different classification algorisms to find the best algorithms and methods that can be used to generate accurate and efficient LULC maps.

1. Introduction

Having accurate land use land cover (LULC) maps is critical to monitor environmental changes and hazards such as floods and droughts. As such, efficient evaluation of related factors significantly relies on correct research of LULC maps. Variations in different land features, such as urbanization and deforestation, geomorphological changes, monitoring water quality, managing groundwater resources, effects of human activities and land monitoring are all significantly affected by environmental changes in which LULC maps can be used effectively to address the goals of each evaluation. Creating thematic LULC maps and analyzing changes over time is helpful in analyzing and extracting meaningful information from satellite imagery. However, for accurate LULC maps, huge amounts and accurate data are needed, which also has an impact on the storage and processing capacities and in choosing different classification methods. Because of this, quantitative statistics and visual interpretation of traditional classifications will not give the required accuracy. However, technological innovations and creating user-friendly satellite image classification methods and machine learning algorithms through utilizing aeronautical reconnaissance coverage geographic information system (ArcGIS) and Google

Earth Engine (GEE) significantly helped researchers to extract meaningful information from satellite imagery more accurately and efficiently, because there is no need for coding expertise to use these platforms, which is helpful for a wider range of users to create LULC maps more easily [1-4].

The output of satellite image classification provides valuable information to evaluate changes in the Earth's surface, variation in land features, environmental assessment, and resource management. Researchers, decision-makers, and environmental management organizations can gain a comprehensive understanding of various land cover types, their spatial distribution, and changes over time through processing and extracting meaningful information from satellite imagery [5, 6].

With the use ofgeographic information system platforms, or GEE, we can now easily use various machine learning algorithms to classify different LULC features and get a deeper understanding of spatial and temporal analysis. ArcGIS Pro hasa user friendly interface making it easier for users with limited coding experience and can also be used offline. However, processing larger datasets requires significant computing capabilities. Using various classification algorithms for satellite imagery increases the reliability of classification processes. Each algorithmuses unique approaches, allowing the complementary power and capabilities offering the researchers the ability to assess the performance of LULC classification outputs and enhance the overall accuracy of LULC maps. Meanwhile, Google Earth Engine is a non-commercial cloud-based platformthat does not need to download a large amount of satellite imagery and which provides access to vast computing power and allows for processing large datasets efficiently. The main key of GEE is that it has access to pre-processed satellite imagery available for analysis. The limitation of GEE is that it requires the users to have coding knowledge and lacks auser friendly interface compared to ArcGIS Pro; it also needs internet access for most functionalities. Satellite imagery like Sentinel collection and Landsat series and multiple free available dataset sources can be used in GEE, which help users to process large amounts of data faster and easier [7-9].

Recently, more studies have beenfocused on different image classification techniques such as support vector machine (SVM), maximum likelihood (ML), random forest (RT), classification and regression tree (CART), K-Nearest Neighbor (KNN) and iterative self organizing (ISO) cluster algorithm. A large number of studies used different technique of classification for their research, having different levels of accuracy results. However, limited studies have compared new LULC classification algorithms using different satellite and platforms, particularly regarding individual class [9, 10]. Previous studies showed that SVM [1, 11, 12], RT [10, 13], and ML [14] had better overall performance among other classifiers. Meanwhile, demanding more accurate LULC maps from satellite imagery is growing, especially using user friendly and easier platforms such as GIS and GEE. Assessing the accuracy of different classifiers for specific classes is thus more critical using different imagery, platforms and algorithms [1].

The objective of this research was to compare and evaluate various techniques performance for classifying LULC using Landsat 8 andSentinel-2 satellite imagery with 30 m, 10 m resolutions, respectively, with ArcGIS Pro and GEE platforms, as well as finding the best classifier based on overall accuracy and specifically for each class. The output of the research will be helpful for researchers to use the most efficient and effective classification algorithm for future use of satellite remote sensing imagery.

2. Materials and Methods

2.1. Study Area

This study focuses on Ranya city, in Sulaymaniyah governorate in the Kurdistan region of Iraq. Ranya city has a total size of 27.2 km² and is located in the north of Sulaymaniyah with 123 km farfromErbil, the capital of the Kurdistan region. Ranya is one of the largest citiesin the Sulaymaniyah Governorate and situated at 36° 15' 14" north and 44° 52' 56" east as shown in figure 1. Ranya is surrounded by mountainsand is also known for its agricultural products, having a huge area of crop lands. The built-up areas radically increase from the city center in all directions, and there have recently been noticeable land cover changes and replacement of farmland and bare lands by settlement areas. The study area has different land cover types (e.g., trees, vegetation, bare land, and built-up areas). Having an effective algorithm to classify each land featureaccurately for a study area such as Ranya city is critical, in which the different classes in thestudy area are suitable for classification method assessments and meetthe objectives of our research.

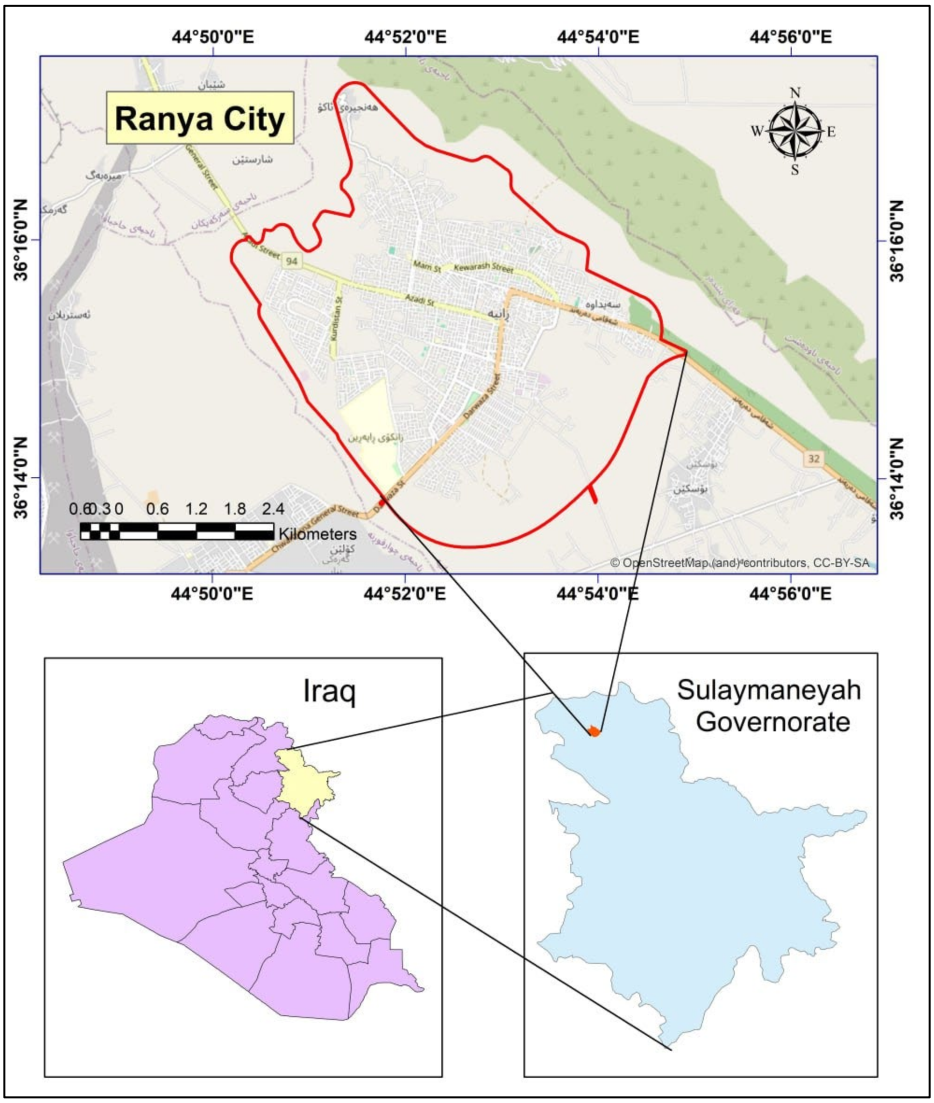

Figure 1: Study area of the research (Ranya city).

2.2. *Dataset*

Two different satellite imagery, Landsat 8 collection 2 level 2 and Sentinel-2, were used on the same dates (March 2023) with zero cloud cover for classification and analyzing the imagery. Landsat 8 was downloaded from USGS website and Sentinel-2 from the Copernicus platform. The ground points used for accuracy assessment for the four classes were collected from high resolution satellite imagery from Google Earth Pro with 30 centimeter resolution to compare it with the classified LULC maps (Table 1).

Table 1: Satellite imagery used for the LULC mapping.

	Landsat 8 Operational Land Imager (OLI)	Sentinel-2
Date	31-May-2022	31-May-2022
Bands	1,2,3,4,5,6,7	1,2,3,4,5,8
Resolution (m)	30	20, 10
Website for download	USGS Earth Explorer https://www.usgs.gov/	Copernicus Open Access Hub https://scihub.copernicus.eu/

2.3. Classification Algorithms

The study focuses on the evaluation of the different in order to evaluate the best classifier for each class by comparing various classification algorithms from GIS and GEE, which are two of the most powerful and user friendly platforms for processing huge data and classifying satellite imagery. Based on the different features, the study area was classified into four distinct classes: built-up area, trees, bare land, and vegetation, covering all Ranya city (Table 2). The methodology workflow of the study illustrated in figure 2 for classifying land use land cover categories usedvarious classification algorithms in ArcGIS Pro, GEE based on Landsat 8 and Sentinel-2. The Sentinel-2 and Landsat 8 imagery used in this study were pre-processed and corrected for geometric and radiometric calibrations. Additionally, both of the satellite imagery was visually processed to get the optimal result for classification using different band combination and enhancement techniques,such as manual contrast stretching in ArcGIS Pro. However, the processing procedure wasdifferent in GEE and suitable bands were selected with cloud masking. Accuracy assessment wasapplied to assess the performance of each algorithm using kappa coefficient. The resulting maps were visualized and prepared by ArcGIS Pro for analyzing,as shown in figure 2.

Table 2: Land use and land cover scheme categories.

Name of the Class	Class Description
Built-up area	Settlement area, roads, utilities, transportation, and industrial area.
Trees	Forest areas, parks.
Bare land	Open fields, rocky areas, barren land.
Vegetation	Agriculture, grass, cropland.

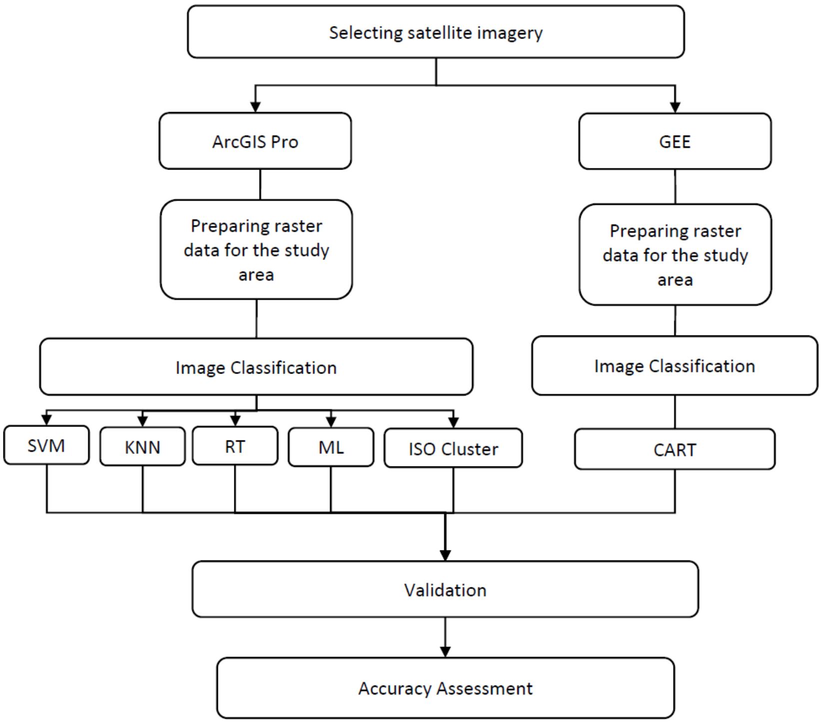

Figure 2: Methodology of the research.

Both supervised and unsupervised algorithms were used to create an LULC map for both Landsat 8 and Sentinel-2. We used ISO Cluster, ML, SVM, KNN, and RT in ArcGIS Pro and CART classifier in a GEE platform:

Random tree classifier is one of the non-parametric machine learning algorithms that creates several independent decision trees, in which each has the power to emphasize input by sequentially creating binary judgments using unique information. The output will be determined by the prediction of identifying the most common class prediction across all trees. The main key of random trees is the non-overfitting characteristics and also it provides significance ratings to each feature, by showing which characteristics the categorizations are most affected. However, the training process is demanding, especially using huge datasets. Meanwhile, despite certain drawbacks, RT classifiers are commonly used for satellite image classification in many machine learning applications due to their user-friendly, efficiency, and accuracy. The main two critical factors in using RT classifiers are size of the data and the number of generated trees in the algorithm. Recently, researchers illustrated the power of the RT algorithm for land use land cover classification meeting the desired requirements. Using more trees usually increases the accuracy of the classification and land cover mapping; however, more trees will cost the computational process and is time consuming [1, 10].

The KNN classifier is one of the unsupervisedmachine learning techniques commonly used in ArcGIS Pro to achivesatellite image classification and regression tasks and is flexible and easy-to-understand. The process of the supervised classifier involves dividing data into training and testing and identifying the k nearest data points to unclassified data points utilizing distance measured by the Euclidean distance statistical formula. However, the selection of the k value and dimensionality have an impact on the performance of the algorithms, resulting in creating low accuracy distances; we used five dimensions for our case using fivebands. The value of k used in this study was set to 5, which was based on cross-validation tests to balance the model's robustness. Value of 5 was found to enhance classification accuracy, providing a more reliable performance of the model for our specific dataset.Moreover, KNN demands high computational costs for large datasets due to its need of the storage of the whole training data for classification [10].

$$d(A,B) = \sqrt{(A1 - B1)^2 + (A2 - B2)^2 + \cdots + (An - Bn)^2} \qquad (1)$$

Where A and B are the coordinates of each band (pixel values) used for classification and d is the Euclidean distance [15].

The CART algorithm is a powerful method that uses decision trees to estimate the different features and classify the satellite image by dividing the data into various similar categories until it reaches the terminal nodes. It's powerful for straightforward decision-making procedures;however, the overfitting occurrence in small datasets is one of the limitations. The GEE platform made the classifier available to be used for different datasets with efficient and accurate results [1, 16].

The SVM is a popular and accurate non-parametric supervised machine learning method that is usually used for classifications by selecting and defining a maximum margin between unique classes. The SVM was introduced in 1979 by Vapnik using the theory of statistical learning. It's a powerful technique for high dimensional with minimum noise and overfitting and can distinguish patterns using pixel-based or object-based classification. However, the training algorithm is computationally demanding for big datasets although its accuracy will increase with the increase of training data. Despite the complexity, SVMs are effective and accurate for anomaly detection and satellite classification and more accurate than other methods like neural network algorithms [1, 10].

The ML uses the normal distribution Byes theorem probability to classify data based on the predefined categories. It's a supervised algorithm that learns from the training data to define and identify new data to the class with greatest likelihood and probability using statistical models. It's a powerful and effective technique especially with well-defined datasets. Moreover, its computation efficiency and user friendliness makes it an easy method to use in a wide range of applications. However, the ML faces challenges with data structures that havenon-normal distribution with unbalanced category size [1, 17].

ISO cluster is an effective unsupervised classification algorithm, especially for users without in-depth remote sensing background, based on iterative algorithms to classify pixels by spectral characteristic. The algorithms automatically identify distinct features without predefined and training data, which makes it easy to use and beneficial for exploration of distinct features in the data. However, the spectral characteristics of data and amount of clusters critically affect the performance of the algorithm. Also the ISO cluster will face challenges with fragmented areas or classes with modest spectral differences [18].

2.4. Data Processing

ArcGIS Pro was used to classify Landsat 8 and Sentinel-2 imagery with zero cloud cover for 30 May 2022 for Ranya city utilizing SVM, ML, KNN, ISO cluster, and RT algorithms. Landsat 8 imagery from collection 2 level 2 and Sentinel-2 level-2A orthorectified atmospherically corrected surface reflectance were used for classification. Satellite images were prepared for the study area by clipping and creating reliable composite bands. For supervised classification algorisms such as SVM, ML, KNN, and RT, a total 570 sample points were used for each four classes for training the algorithms. However, for unsupervised classification ISO cluster algorithm, 40 different classes were created based on their spectral similarities then the classes were reduced to four main categories manually. However, for GEE platform, we used CART classifier for both surface reflectance Landsat 8 and Sentinel-2 satellite image for the study area with 450 sample trainings. The imagery wasselected using the boundary of the study area, cloud masking and filling the gaps utilizing median aggregation approach [1].

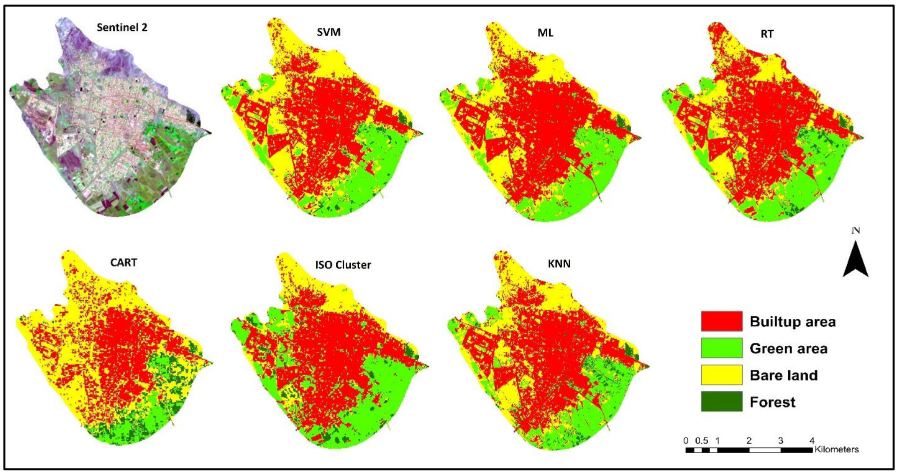

Figure 3: Results of using LULC algorithms for the image of Sentinel-2.

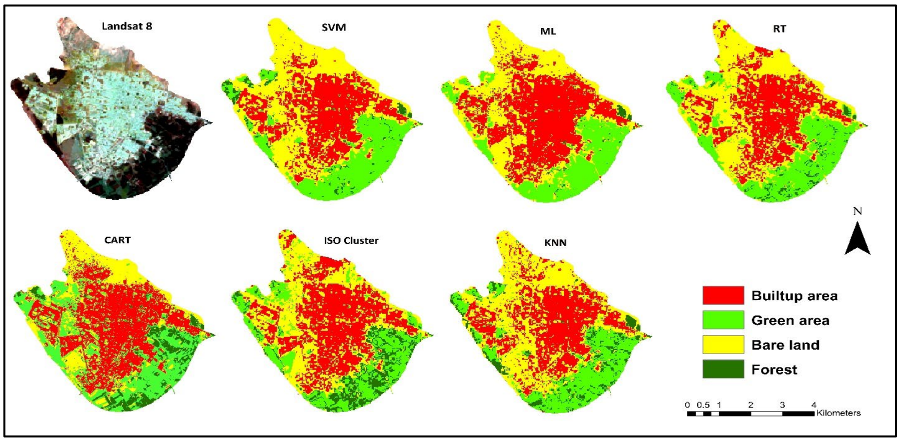

Figure 4: Results of using LULC algorithms for the image of Landsat 8.

2.5. Accuracy Assessment

The aim of this study was to find the best approach for classifying satellite imagery into categories such as built-up areas, vegetation, bare land, and trees. To find the best classifier and evaluate the performance of each classification algorithm for each LULC class, we assessed the accuracy of each classification technique by collecting 100 random points equally distributed over the four classes using equally stratified random sampling. The testing points were evaluated by using satellite imagery of 2022 in Google Earth Pro by comparing the LULC map of 2022 of each algorithm to the actual photo of 2022 of Ranya. The accuracy of each classand overall performance of all algorithms were evaluated, which showshow much of the testing data were correctly identified using kappa coefficient matrix.

However, the accuracy was different, not only by choosing different algorithms, but also by location of the testing samples. The testing samples were used to calculate the accuracy of the classifier utilizing ArcGIS Pro by which high kappa accuracy percentage means the classifier is more accurate, and vice versa (Table 3)[19, 20, 21].

Table 3: Kappa coefficient categorizations [22].

No.	Kappa coefficient (KC)	Categorization
1	< 0.00	Poor (P)
2	0.00-0.20	Slight (S)
3	0.21-0.40	Fair (F)
4	0.41-0.60	Good (G)
5	0.61-0.80	Very good (V.G)
6	>0.81	Excellent (E)

The KC equation (5) and OA equation (4) in the kappa coefficient matrix illustrate that the assessment of the accuracy of each land use land cover classes is reliable (Table 3). These equations are the criteria for evaluating the accuracy and determining the kappa coefficient [23]:

$$User\ Accuracy\ (UA) = \frac{CR}{TR} \quad (2)$$

where CRequation (2)is represented by the proportion of rows with valid categorization. However, the term "TR" equation (2) is the total classified pixels in the same row.

$$Producer's\ accuracy\ (PA) = \frac{CC}{TC} \quad (3)$$

Where CC equation (3) refers to the number of correctly classified pixels in each column. Also the term TC equation (3) represents the accurate pixels that are classified in the column.

$$Overall\ accuracy\ (OA) = \frac{CD}{TR} \quad (4)$$

The "CD" equation (4) parameter is the total number of correctly classified pixels diagonally. Also the sum of all reference points represented by the term of TP.

$$\text{Kappa coefficient } (KC) = \frac{(TP*CD)-\sum(TC*TR)}{TP^2-\sum(TC*TR)} \quad (5)$$

3. Results

Based on the LULC classification using different algorithms from ArcGIS Pro and Google Earth Engine for Landsat 8 and Sentinel-2 imagery, we found that LULC maps created using Landsat 8 data have major misclassification in all classes compared to Sentinel-2 imagery. Comparing both results, considering the overall accuracy the outcome shows that the random tree algorithm achieved highest accuracy using Sentinel-2 with 83%,as illustrated in figure 3. This illustrates RT's ability to handle high dimensional data and classify complex boundaries between the different features, which effectively classified general patterns within different features in the image resulting in more accurate classification across all categories compared to other techniques. However, maximum likelihood was the best among the algorithms that used Landsat 8 but with only 68% (Table 4).

Meanwhile, focusing on the specific class accuracy, Support Vector Machine using ArcGIS Pro successfully recognized the built-up areas with more than 92% based on user's accuracy (UA) and 100% for producer's accuracy (PA). In the process of classification systems evaluation, UA is the probability of a classified data point that truly belonging to the class, which is significant for user confidence. However, PA represents the ability of a system to accurately identify actual data of a specific class. By using both UA and PA, researchers provide a comprehensive evaluation for performance of the classification algorithms [24].

Table 4: Land use and land cover scheme categories for Sentinel-2.

Sentinel-2		SVM	ML	ISO cluster	RT	KNN	CART
Kappa accuracy %		0.74	0.81	0.77	0.83	0.68	0.67
PA%	Built-up	1	0.96	0.92	1	0.77	0.88
	Vegetation	0.59	0.68	0.97	0.75	0.64	0.6
	Tree	0.96	0.96	0.88	0.85	0.78	1
	Bare land	0.87	0.95	0.94	0.95	1	0.79
UA%	Built-up	0.92	0.92	0.88	0.84	0.68	0.84
	Vegetation	0.92	0.92	1	0.96	1	1
	Tree	0.88	0.88	0.84	0.88	0.72	0.4
	Bare land	0.52	0.72	0.6	0.8	0.64	0.76

While SVM did not achievebest overall accuracy, it surpassed other algorithms in classifying urban regions accurately using Sentinel-2, which makes it useful for applications such as mapping urban features, monitoring urban sprawl and building footprint extraction. Moreover, the highest accurate accuracy for detecting vegetation was the ISO Cluster unsupervised algorithm using Sentinel-2 with PA 97% (Table 4). The ISO Cluster from ArcGIS Pro effectively organized pixels with similar characteristics based on the spectral properties to separate vegetation features precisely with 97% PA and 100% UA. The outcome suggests that ISO Cluster is the most suitable for monitoring agriculture, crop type identification, and crop health assessment using Sentinel-2 imagery based on the accuracy of the results (Table 4, 5). Regarding the tree classification, due to the ability to handle the complex relationship with the satellite data, both SVM and ML algorithms showed great performance with 96% in PA and 88% UA in classifying tree cover over the study area for Sentinel-2 data. This makes both of them suitable for application related to generating thematic maps, deforestation monitoring and classifying data with distinct spectral characteristics. The study showed that random tree, which had the highest overall accuracy, is best among the other algorithms for classifying bare lands and with 95% PA and 80% UA, suggesting that it's most suitable for change detection applications such as rapid deforestation urbanization and drought monitoring using Sentinel-2 (Table 4 and 5).

Table 5: Land use and land cover scheme categories for Landsat 8.

Landsat 8		SVM	ML	ISO cluster	RT	KNN	CART
Kappa accuracy %		0.64	0.68	0.57	0.59	0.55	0.66
PA%	Built-up	0.92	0.91	0.82	0.83	0.71	0.59
	Vegetation	0.6	0.75	0.51	0.5	0.63	0.88
	Tree	1	0.77	1	1	0.89	0.77
	Bare land	0.65	0.67	0.77	0.79	0.56	0.82
UA%	Built-up	0.92	0.76	0.92	1	0.88	0.93
	Vegetation	0.84	0.96	0.88	0.84	0.88	0.88
	Tree	0.36	0.52	0.12	0.2	0.32	0.4
	Bare land	0.8	0.8	0.8	0.72	0.56	0.75

4. Discussion

After accuracy assessment for all algorithms, our research shows that choosing Sentinel-2 data (Table 3.) is more suitable compared to Landsat 8 (Table 4.). The main reason for algorithms achieving higher accuracy is Sentinel-2 having higher spatial resolution, which offers 10 meter resolution for several bands, compared to 30 meter resolution of Landsat 8 for most bands. Having higher spatial resolution allows the satellite to capture more detail resulting to improved accuracy in classifying different features in the image. Moreover, Sentinel-2 has more spectral bands, providing a wider range of information about the surface features. These distinct bands are significant for differentiating small features and enhancing the performance of the classifiers. However, broader research and using different locations with different features, also utilizing different software and algorithms, is recommended to choose the best data and algorithms for classifying LULC maps. Moreover, despite the lower accuracy compared to Sentinel-2 for recent imagery, we can only depend on Landsat series for time series analysis due to its longer data archives,as shown in figure 3 and 4.

The varying accuracies of different type of algorithmsused for the same image are due to their unique designs and operational mechanisms. Random tree builds multiple decision trees, having high performance in handling high-dimensional data and complex boundaries, resulting in high overall accuracy. Conversely, SVM distinguishes closely related classes effectively, leading to higher accuracy in specific areas like urban regions. The higher spatial and spectral resolutions of Sentinel-2 satellite imagery also improve the performance of the used algorithm. Understanding these characteristics is significant for selecting the optimal algorithm for mapping and creating accurate and reliable LULC [25 - 27].

We applied SVM, ML, ISO cluster, RT, and KNN classifiers using ArcGIS Pro and CART classifier using GEE for classifying Landsat 8 and Sentinel-2 imagery. Both ArcGIS Pro and GEE have their benefits and limitations. The study showed that some algorithms used in ArcGIS Pro have better performance compared to CART algorithm in GEE. For easy and user friendly interface with efficient accuracy, choosing ArcGIS Pro is preferable. Conversely, GEE's huge data library and cloud-based processing is more suitable for users with coding experiences.

The study showed that choosing the best algorithms should be based on the user's specific objectives. If the overall accuracy is the priority, RT is the most suitable one. However, for specific class or application, considering individual class accuracy and utilizing the best algorithms is essential, such as SVM for built-up areas and ML for bare land classification.

5. Conclusions

LULC mapping is essential for monitoring environmental changes. In this study, we used different datasets, platforms and classifiers to assess the performance of different classification algorithms in overall and class-based accuracy. ArcGIS Pro and GEE with Landsat 8 and Sentinel-2 datasets using SVM, ML, RT, CART, KNN and ISO cluster algorithms were utilized to classify the satellite imagery of the Ranya city area. The research identified RT algorithm in ArcGIS Pro as the best classifier with overall accuracy of 83% using Sentinel-2, also the ML algorithm is giving the best overall accuracy with 68% for Landast 8 imagery and second best for Sentinel 2 with 81%. Furthermore, regarding the accuracy of specific classes, the SVM algorithm successfully identified built-up areas with 92% UA and 100% PA, which was better than all other methods and ISO Cluster showed best accuracy for identifying vegetation areas with 97% PA and 100% UA. SVM and ML had minimum misclassification regarding classifying trees with 96% PA and 88% UA accuracy, also the best algorithmthat could be used for deforestation and drought monitoring was the RT classifier as it accurately recognized the bare land areas from Sentinel-2 imagery with PA of 95% and UA of 80%. Our research recommends the use of Sentinel-2 for more accurate LULC mapping over the Landsat 8 based on the results of all algorithms in which they have better accuracy in Sentinel-2 using ArcGIS Pro.

Authors contributions: Kaifi Chomani: Writing – original draft, Conceptualization, Investigation, Methodology, Project administration. **Shaki Pshdari:** Writing – review & editing.

Data availability: Data will be available upon reasonable request.

Conflicts of interest: The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Funding: The authors did not receive support from any organization for the submitted work.

References

[1] S. Basheer et al., “Comparison of Land Use Land Cover Classifiers Using Different Satellite Imagery and Machine Learning Techniques,” Remote Sensing, vol. 14, no. 19, p. 4978, 2022, doi: 10.3390/rs14194978.

[2] J. Mäyrä et al., “Utilizing historical maps in identification of long-term land use and land cover changes,” Ambio, vol. 52, no. 11, p. 1777-1792, 2023, doi: 10.1007/s13280-023-01838-z.

[3] F. Ahmad, L. Goparaju, and A. Qayum, “LULC analysis of urban spaces using Markov chain predictive model at Ranchi in India,” Spatial Information Research, vol. 25, no. 3, p. 351-359, 2017, doi: 10.1007/s41324-017-0102-x.

[4] R. Pacheco Quevedo et al., “Land use and land cover as a conditioning factor in landslide susceptibility: a literature review,” Landslides, vol. 20, no. 5, p. 967-982, 2023, doi: 10.1007/s10346-022-02020-4.

[5] K. Thiagarajan et al., “Satellite Image Classification Using a Hierarchical Ensemble Learning and Correlation Coefficient-Based Gravitational Search Algorithm,” Remote Sensing, vol. 13, no. 21, p. 4351, 2021, doi: 10.3390/rs13214351.

[6] H. Ferdous et al., “Machine Learning Approach Towards Satellite Image Classification,” in Proceedings of International Conference on Trends in Computational and Cognitive Engineering, Singapore, 2021, doi: 0.1007/978-981-33-4673-4_51.

[7] I. Atef, W. Ahmed, and R.H. Abdel-Maguid, “Modelling of land use land cover changes using machine learning and GIS techniques: a case study in El-Fayoum Governorate, Egypt,” Environmental Monitoring and Assessment, vol. 195, no. 6, p. 637, 2023, doi: 10.1007/s10661-023-11224-7.

[8] G. Abebe, D. Getachew, and A. Ewunetu, “Analysing land use/land cover changes and its dynamics using remote sensing and GIS in Gubalafito district, Northeastern Ethiopia,” SN Applied Sciences, vol. 4, no. 1, p. 30, 2021, doi: 10.1007/s42452-021-04915-8.

[9] A. Tassi and M. Vizzari, “Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms,” Remote Sensing, vol. 12, no. 22, p. 3776, 2020, doi: 10.3390/rs12223776.

[10] S. Talukdar et al., “Land-Use Land-Cover Classification by Machine Learning Classifiers for Satellite Observations—A Review,” Remote Sensing, vol. 12, no. 7, p. 1135, 2020, doi: 10.3390/rs12071135.

[11] S. Dhingra and D. Kumar, “A review of remotely sensed satellite image classification,” International Journal of Electrical and Computer Engineering, vol. 9, no. 3, p. 1720, 2019, doi: 10.11591/ijece.v9i3.pp.1720-1731.

[12] Z. Abbas and H.S. Jaber, “Accuracy assessment of supervised classification methods for extraction land use maps using remote sensing and GIS techniques,” in IOP Conference Series: Materials Science and Engineering, 2020, doi: 10.1088/1757-899X/745/1/012166.

[13] E.A. Alshari and B.W. Gawali, “Development of classification system for LULC using remote sensing and GIS,” Global Transitions Proceedings, vol. 2, no. 1, p. 8-17, 2021, doi: 10.1016/j.gltp.2021.01.002.

[14] K. Nivedita Priyadarshini et al., “A COMPARATIVE STUDY OF ADVANCED LAND USE/LAND COVER CLASSIFICATION ALGORITHMS USING SENTINEL-2 DATA,” Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., vol. XLII-5, p. 665-670, 2018, doi: 10.5194/isprs-archives-XLII-5-665-2018.

[15] Lubis, A. R., & Lubis, M. ‘Optimization of distance formula in K-Nearest Neighbor method’, Bulletin of Electrical Engineering and Informatics, 9(1), 326-338, 2020, doi: 10.11591/eei.v9i1.1464.

[16] A. Tassi and M. Vizzari, "Object-Oriented LULC Classification in Google Earth Engine Combining SNIC, GLCM, and Machine Learning Algorithms," Remote Sensing, vol. 12, no. 22, p. 3776, 2020. doi: 10.3390/rs12223776.

[17] B. Norovsuren et al., “Land cover classification using maximum likelihood method (2000 and 2019) at Khandgait valley in Mongolia,” IOP Conference Series: Earth and Environmental Science, vol. 381, no. 1, p. 012054, 2019, doi: 10.1088/1755-1315/381/1/012054.

[18] Y. Hu et al., “Land Use/Land Cover Change Detection and NDVI Estimation in Pakistan’s Southern Punjab Province,” Sustainability, vol. 15, no. 4, p. 3572, 2023, doi: 10.3390/su15043572.

[19] C.C. Fonte et al., “Assessing The Accuracy Of Land Use Land Cover (Lulc) Maps Using Class Proportions In The Reference Data,” ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., vol. V-3-2020, p. 669-674, 2020, doi: 10.5194/isprs-annals-V-3-2020-669-2020.

[20] K.-S. Cheng et al., “Quantifying Uncertainty in Land-Use/Land-Cover Classification Accuracy: A Stochastic Simulation Approach,” Frontiers in Environmental Science, vol. 9, 2021, doi: 10.3389/fenvs.2021.628214.

[21] M. Jamil, H.U. Rehman, SaleemUllah, I. Ashraf, and S. Ubaid "Smart Techniques for LULC Micro Class Classification Using Landsat8 Imagery," Comput. Mater. Contin., vol. 74, no. 3, pp. 5545-5557. 2023,doi: 10.32604/cmc.2023.033449.

[22] Landis, J. R., & Koch, G. G “The measurement of observer agreement for categorical data,” Biometrics, vol. 33(1), pp. 159-174, 1977, doi: 10.2307/2529310.

[23] S. G. Setegn, R. Srinivasan, B. Dargahi, and A. M. Melesse, "Spatial delineation of soil erosion vulnerability in the Lake Tana Basin, Ethiopia," Hydrological Processes: An International Journal, vol. 23, no. 26, pp. 3738-3750, Dec. 2009, doi: 10.1002/hyp.7476.

[24] Nicolau, A. P., Dyson, K., Saah, D., & Clinton, N. “Accuracy Assessment: Quantifying Classification Quality. In Cloud-Based Remote Sensing with Google Earth Engine: Fundamentals and Applications” Springer International Publishing, pp. 135-145, 2024, doi: 10.1007/978-3-031-26588-4_7.

[25] R. C. Chen, C. Dewi, S. W. Huang, and others, "Selecting critical features for data classification based on machine learning methods," Journal of Big Data, vol. 7, p. 52, 2020. doi: 10.1186/s40537-020-00327-4.

[26] R. G. McClarren, "Decision Trees and Random Forests for Regression and Classification," in Machine Learning for Engineers. Cham, Switzerland: Springer, 2021. doi: 10.1007/978-3-030-70388-2_3.

[27] M. Awad and R. Khanna, "Support Vector Machines for Classification," in Efficient Learning Machines. Berkeley, CA: Apress, 2015. doi: 10.1007/978-1-4302-5990-9_3.