Comparing Unsupervised Land Use Classification of Landsat 8 OLI Data Using K-means and LVQ Algorithms in Google Earth Engine: A Case Study of Casablanca

Ouchra, H.,[1]* Belangour, A.[1] and Erraissi, A.[2]
[1]Laboratory of Information Technology and Modeling LTIM, Hassan II University, Faculty of Sciences Ben M'sik Casablanca, Morocco, E-mail: ouchra.hafsa@gmail.com*
[2]Chouaib Doukkali University, Polydisciplinary Faculty of Sidi Bennour, El Jadida, Morocco
**Corresponding Author*

DOI: https://doi.org/10.52939/ijg.v19i12.2981

Abstract
Accurate and up-to-date land use information is essential for effective urban planning and environmental management. This paper presents a methodology for the unsupervised classification of Casablanca's land cover using Google Earth Engine (GEE). The study exploits multispectral satellite imagery, in particular Landsat data, to extract meaningful land-use classes without the need for manual labeling. The workflow includes data collection, pre-processing, unsupervised clustering, and visualization of results. By applying the k-means and Learning Vector Quantization (LVQ) clustering algorithms, the city's land area is divided into distinct clusters, each representing a specific land-use class. The resulting land-use map provides valuable information on Casablanca's urban landscape, highlighting forest areas, crops, built-up infrastructure, water bodies, and barren areas. This automated approach demonstrates the potential of GEE as a powerful tool for land use analysis, enabling effective, data-driven decision-making for urban development and environmental monitoring. The methodology presented can serve as a basis for similar studies in other regions, contributing to the advancement of remote sensing and geospatial analysis techniques for urban and environmental studies. This study evaluates the effectiveness of these two algorithms in terms of overall accuracy and kappa coefficient. The K-means algorithm recorded moderate accuracy. The LVQ algorithm, on the other hand, performed the least well.

Keywords: Google Earth Engine, Land Cover, Machine Learning, Satellite Image, Unsupervised Classification

1. Introduction

The accurate characterization and mapping of land cover in urban areas are essential for effective urban planning [1] and [2], environmental management, and sustainable development. As cities continue to grow and evolve, the demand for up-to-date and reliable land cover information becomes increasingly crucial. This paper addresses this demand by presenting a comprehensive study on land cover mapping in the vibrant urban context of Casablanca, utilizing the power of unsupervised learning techniques [3], Google Earth Engine, and Landsat satellite data. In appendix, we present URL_GEE.

Urban areas, like Casablanca, are intricate landscapes characterized by a diverse mix of land cover classes, ranging from natural features such as vegetation and water bodies to anthropogenic structures including buildings, roads, and industrial areas. Traditional manual methods of land cover classification [4] and [5] are time-consuming and labor-intensive, often struggling to capture the complexity and dynamics of these rapidly changing environments. Consequently, the application of advanced remote sensing technologies [6] and [7] and machine learning algorithms has emerged as a promising solution to address the challenges of urban land cover analysis.

In this context, the synergy between Google Earth Engine and Landsat satellite data provides an unprecedented opportunity for efficient and accurate land cover mapping.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) |

Google Earth Engine [8] and [9], a cloud-based geospatial processing platform, offers access to a vast repository of multispectral imagery and a suite of analytical tools. Combined with Landsat satellite data, which provides a rich source of information across multiple spectral bands, this platform enables the exploration of unsupervised learning methods for land cover classification. The primary objective of this paper is to present a detailed methodology for utilizing unsupervised learning techniques within the Google Earth Engine framework to perform land cover mapping in Casablanca. This approach involves data acquisition, preprocessing, unsupervised clustering, and result visualization. The inherent capability of unsupervised learning algorithms, particularly the widely used k-means [10], and LVQ clustering [11], to identify patterns and groupings in complex datasets allows for the automatic extraction of distinct land cover classes without the need for manual labeling.

The outcomes of this study extend beyond the generation of a land cover map; they contribute to a deeper understanding of the spatial arrangement and distribution of various land cover classes in Casablanca. Furthermore, the findings demonstrate the potential of Google Earth Engine and Landsat data in streamlining the land cover mapping process, enabling informed decision-making, and supporting sustainable urban development [12].

In the following sections, we delve into the methodology employed for data processing, unsupervised learning implementation, and result interpretation. We also discuss the implications of the study's outcomes for urban planning and environmental management. By advancing the field of unsupervised land cover mapping through the integration of cutting-edge technologies, this paper aims to offer insights and methodologies that can be applied not only in Casablanca but also in other urban areas facing similar challenges.

2. Related Work

The literature in the field of machine learning [13][14][15][16] and [17] for land cover mapping has grown significantly with the advent of advanced geospatial technologies and the availability of high-resolution satellite data. Previous work [10][17][18] and [19] highlights the ability of unsupervised learning to autonomously identify and categorize distinct land cover classes, offering a data-driven approach to land cover analysis. Urban land cover mapping has received particular attention due to the challenges associated with dynamic and heterogeneous urban environments [20]. Previous studies have highlighted the complexity of urban land cover analysis, including spectral heterogeneity, mixed pixels, and temporal changes [21].

The integration of Google Earth Engine into geospatial analysis has revolutionized the field, enabling efficient processing and analysis of data at scale [8][22] and [23]. Landsat satellite data has served as a cornerstone for land cover mapping due to its multispectral capabilities and historical continuity. Previous research has demonstrated the usefulness of Landsat data for urban land use analysis, leading to a better understanding of urban growth and dynamics [24].

Building on these foundations, this paper presents a comprehensive methodology for unsupervised learning-based land-use mapping in the urban context of Casablanca. By leveraging the capabilities of Google Earth Engine and the multispectral richness of Landsat satellite data, this study aims to contribute to the advancement of automated land use analysis techniques in urban areas. This review highlights the importance of our study in the context of existing literature, emphasizing the unique contribution of integrating unsupervised learning techniques, Google Earth Engine, and Landsat satellite data for land use mapping in Casablanca.

3. Materials and Methods

3.1 Study area and datasets

Casablanca is a major city in Morocco, located on the country's western coast along the North Atlantic Ocean. Conducting a study in the Casablanca region involves analyzing land cover and other geographical features in and around the city. It is the largest city in the country and serves as its economic and business hub. Urban, suburban, and coastal environments characterize the city. Casablanca is situated at around 33.5731° N latitude and 7.5898° W longitude. The geographical location of Casablanca is depicted in Figure 1. The Landsat 8 Collection 1 Tier 1 dataset [25] provides calibrated Digital Number (DN) values of satellite imagery captured by the Landsat 8 Operational Land Imager (OLI) and Thermal Infrared Sensor (TIRS) instruments. This dataset is widely used for various remote sensing and geospatial analyses, including land cover classification, change detection, and environmental monitoring. It is used for a wide range of applications, including [26]:

- Land cover classification and mapping.
- Monitoring urban growth and changes in natural landscapes.
- Assessing vegetation health, crop monitoring, and forestry analysis.
- Studying water bodies, coastal dynamics, and aquatic ecosystems.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

- Analyzing surface temperature and thermal anomalies.
- Change detection and environmental impact assessment.

The Landsat 8 Operational Land Imager (OLI) captures pictures utilizing 11 spectral bands, encompassing wavelengths from visible to near-infrared and shortwave infrared. It utilized for this investigation were procured during various temporal intervals, allowing for temporal assessments to track transformations in land cover across distinct periods. The spatial precision of Landsat 8 OLI images varies, with resolution ranging from 15 meters for the panchromatic band to 30 meters for the multispectral bands.

3.2 Methodology

The unsupervised learning methodology for mapping the land cover of Casablanca using Google Earth Engine involves a sequence of steps encompassing data preparation, unsupervised clustering, interpretation, accuracy assessment, and visualization. The integration of Google Earth Engine's [9] capabilities and satellite imagery facilitate efficient and accurate land cover classification within the urban context of Casablanca. The following steps describe the methodology shown in Figure 2.

3.2.1 Data acquisition and preprocessing

The study commences by acquiring and pre-processing the Landsat satellite imagery [26] from the Google Earth Engine repository. These multispectral images, covering a temporal range of interest, are essential for capturing the dynamic land cover patterns of the urban area. The images are rectified, atmospherically corrected, and georeferenced to ensure accurate and consistent data input. We selected the period from January 1, 2021 to December 31, 2021, to capture the different land use conditions and changes over time.

3.2.2 Region of Interest (ROI) selection

Next, we defined the specific area of Casablanca that will be the focus of the classification. The ROI should encompass the land cover types of interest and account for the urban complexity of the city.

3.2.3 Unsupervised learning algorithm

Subsequently, unsupervised learning algorithms, specifically k-means clustering [27], and LVQ [28] and [29] are applied to the pre-processed Landsat imagery. K-means clustering [27] is chosen for its capability to autonomously group pixels with similar spectral characteristics into distinct clusters, which are representative of different land cover classes.

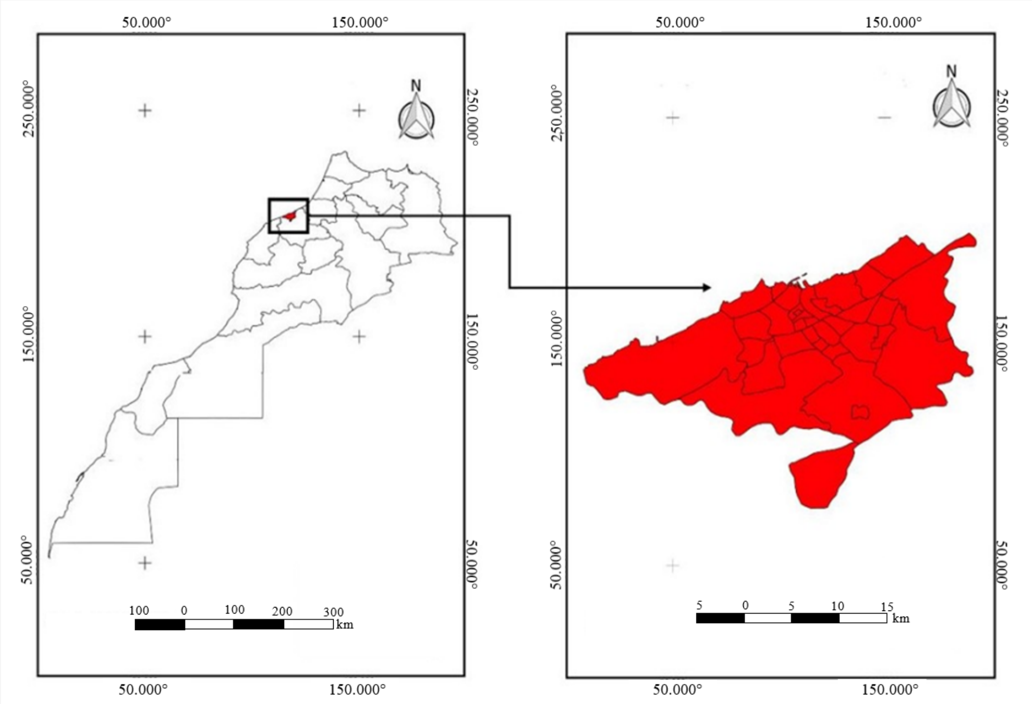

Figure 1: Study area (left) Morocco (right) Casablanca city in western Morocco

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

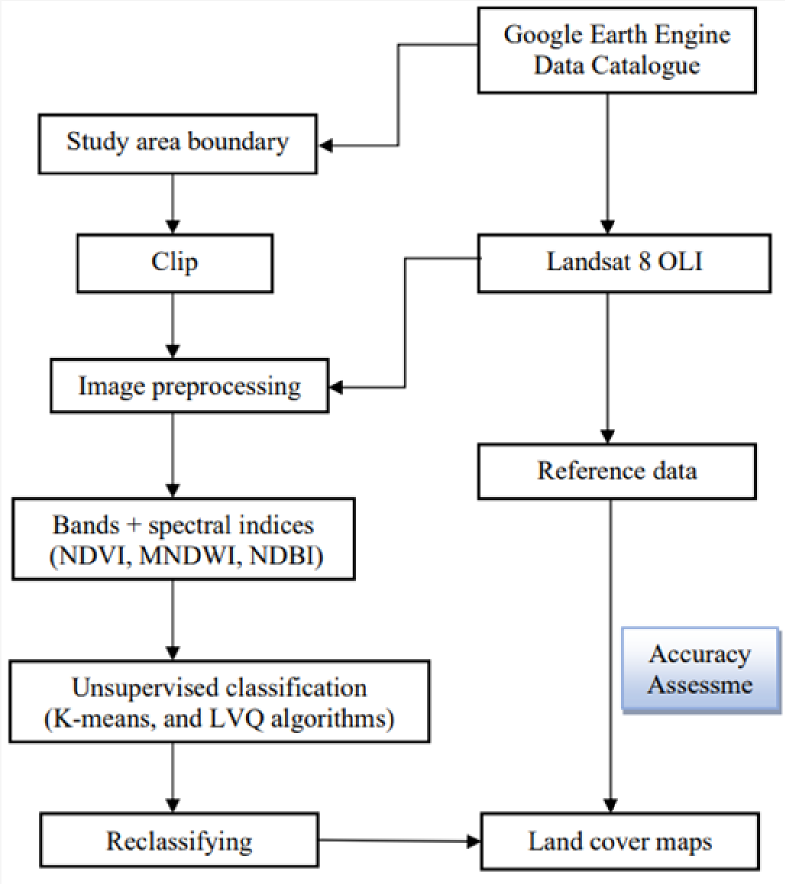

Figure 2: Workflow of methodology implemented in Google Earth Engine (GEE)

The algorithm is initialized with a user-defined number of clusters and iteratively assigns pixels to the nearest cluster centroid based on spectral similarity. Google Earth Engine's cloud-based processing capabilities play a pivotal role in the scalability of the methodology. Learning vector quantization (LVQ) [28], is a supervised learning algorithm, but it always works with unlabeled data, i.e., it can also be said to be an unsupervised learning algorithm. It has been used to compare the differences between unsupervised algorithms and supervised learning algorithms. It uses a training data set to classify new data points according to their similarity to labeled examples in the training set. LVQ [11] and [34] is a competitive learning algorithm, meaning that it trains several classifiers and chooses the one that works best on the training data. A key feature of LVQ is that it allows classifiers to adjust their bounds to better fit the data, which can improve model accuracy. Overall, LVQ is a useful tool for classification tasks in machine learning and can be applied to a wide range of applications.

In this paper, we used LVQ as an unsupervised clustering algorithm to group pixels or data points based on their similarity. The LVQ algorithm in GEE works by iteratively adjusting a set of cluster prototypes to minimize the distance between prototypes and data points. It assigns each data point to the nearest prototype, effectively clustering the data into distinct groups. The platform allows for efficient computation of clustering algorithms across the extensive spatial and temporal coverage of Landsat imagery, enabling rapid analysis of the entire study area.

3.2.4 Accuracy assessment and validation

Validation and accuracy assessment are integral components of the methodology. A stratified random sampling approach is employed to generate reference data points across the study area. These reference points are used to validate the accuracy of the unsupervised classification results. Table 2 shows the reference data used. Various accuracy metrics, including overall accuracy and kappa coefficient, are calculated to quantify the agreement between the classified land cover map and the reference data.

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Table 1: Index spectacles description

Metrics	Equation	Description
Overall Accuracy (OA)	$OA = \frac{Number\ of\ correctly\ classified\ sample}{Number\ of\ samples}$	It is used to determine the proportion of correctly mapped reference sites among all reference sites, expressed as a percentage. It is calculated by dividing the number of correctly classified samples by the total number of samples [31] and [32].
Kappa Coefficient (*K*)	$K = \frac{P_{Accord} - P_{Hazard}}{1 - P_{Hazard}}$ where: P_{Accord} is the observation of inter-rater agreement P_{Hazard} is the overall probability that graders agree	It provides an overall evaluation of the classification performance compared to random assignment. It ranged from -1 to 1 and is derived from a statistical test to evaluate the classification accuracy, determining if the classification performed better than random [32].

Table 2: Land cover classes/ Reference data

Land cover Classes	Numbers of points
Water body	115
Forest	174
Barren	150
Built-up	130
Cropland	119
Total	**688**

3.2.5 Visualization and mapping

The final step of the methodology involves visual interpretation and refinement. The unsupervised classification results are visually compared with high-resolution imagery and existing land cover maps to identify and rectify any misclassifications or inconsistencies.

In summary, throughout the methodology, Google Earth Engine's computational capabilities are leveraged for efficient processing, enabling the analysis of large-scale Landsat datasets. The cloud-based nature of the platform facilitates scalability and accelerates data processing to achieve accurate land cover mapping in the urban environment of Casablanca.

4. Result and Discussion

This section presents the results and subsequent analysis of the Weka K-means and LVQ algorithm applied in Google Earth Engine (GEE) for land cover mapping of Casablanca using Landsat 8 OLI satellite data for the specified period 1-1-2021 to 31-12-2021. The image was segmented into five distinct land-use classes: water, forest, cultivated areas, barren land, and built-up areas. Performance metrics, including accuracy and Kappa coefficient, provide insight into the effectiveness of these algorithms. Table 1 shows the equations for these evaluation metrics. In the Google Earth Engine platform, clusters are used in the same way as classifiers. The training data are Feature Collection properties which are fed into the clusterer. Unlike classifiers, there is no input class value for a clusterer file. The general clustering workflow we followed in this paper is presented in figure 2: first, we collected features with numerical properties in which we searched for clusters. then we instantiated our cluster and trained the clusterer using the training data, then we applied the clusterer to the Landsat satellite image of our study area (Casablanca) and finally we labeled the clusters.

In this study, several underlying factors contribute to the accuracy achieved by the K-means [27] and LVQ algorithm:

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

- Spectral discrimination: The ability of the algorithm to discriminate between land-use classes is highly dependent on the selection of spectral bands and derived indices. It is essential to make an informed choice of features with high discriminating power.
- Number of clusters (*K*): Determining the optimum number of clusters (*K*) is crucial. An inappropriate choice can lead to classes being merged or split, affecting classification accuracy.
- Spectral similarity: Spectral similarity between certain land use classes can cause confusion and reduce classification accuracy.

Based on the results obtained by the algorithms used, it can be concluded that unsupervised methods do not achieve optimum performance in the classification and clustering of satellite images for land cover mapping. Supervised approaches, on the other hand, achieve significantly higher accuracy rates [15][35] and [36]. Nevertheless, in the context of this study, considerable efforts have been made to improve the performance of the K-means and LVQ algorithms. Prior to the application of techniques to improve data quality, the results of K-means and LVQ had demonstrated particularly low accuracy, amounting to 31.47% and 19.96%, respectively, with a Kappa coefficient of 0.14 and 0.0009, respectively, as shown in Table 3. The land-use maps obtained for each algorithm are shown in Figure 4. To improve these results, a series of improvements were made to the pre-processing stages. Atmospheric correction and cloud removal were applied to improve source data quality. Next, the integration of spectral indices [35] such as NDBI, NDVI, and WNDBI was carried out in conjunction with the image bands. The incorporation of these indices was aimed at increasing separability between the different classes. Table 4 shows the equations for these spectral indices. This approach proved successful, generating a clear improvement in the accuracy of the K-means algorithm, reaching an accuracy of 38.66% with 0.06 Kappa coefficients, and for the LVQ algorithm, we have an accuracy of 26.01% with 0.06 Kappa coefficients. Figure 3 illustrates these results. This improvement reflects moderate accuracy but represents a significant advance in the initial results.

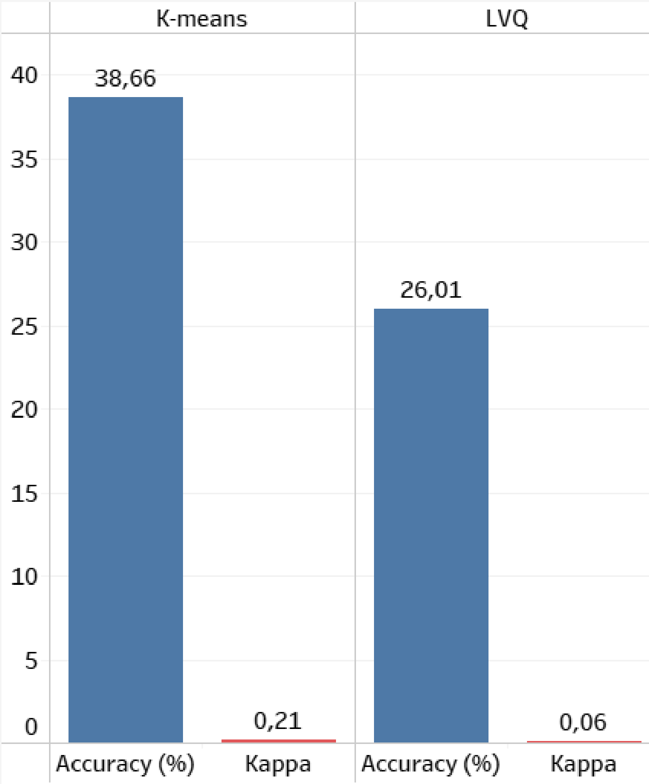

Figure 3: The overall accuracy and Kappa coefficient values of each adopted algorithm

ISSN: 1686-6576 (Printed) | ISSN 2673-0014 (Online) | © Geoinformatics International

Table 3: Accuracy and Kappa coefficient of each algorithm

Unsupervised algorithms	**After the image is rectified, atmospherically corrected, and georeferenced**		**Before the image is rectified, atmospherically corrected, and georeferenced**	
	Accuracy (%)	**Kappa**	**Accuracy (%)**	**Kappa**
K-means	38.66	0.21	31.47	0.14
LVQ	26.01	0.06	19.96	0.00092

Table 4: spectacle indices used

Index	**Equation**	**References**
Normalized Difference Vegetation Index (NDVI)	$NDVI = \frac{NIR - RED}{NIR + RED}$	[36] and [37]
Modified Normalized Difference Water Index (MNDWI)	$MNDWI = \frac{Green - SWIR1}{Green + SWIR1}$	[38]
Normalized Difference Built-Up Index (NDBI)	$NDBI = \frac{SWIR - NIR}{SWIR + NIR}$	[39]

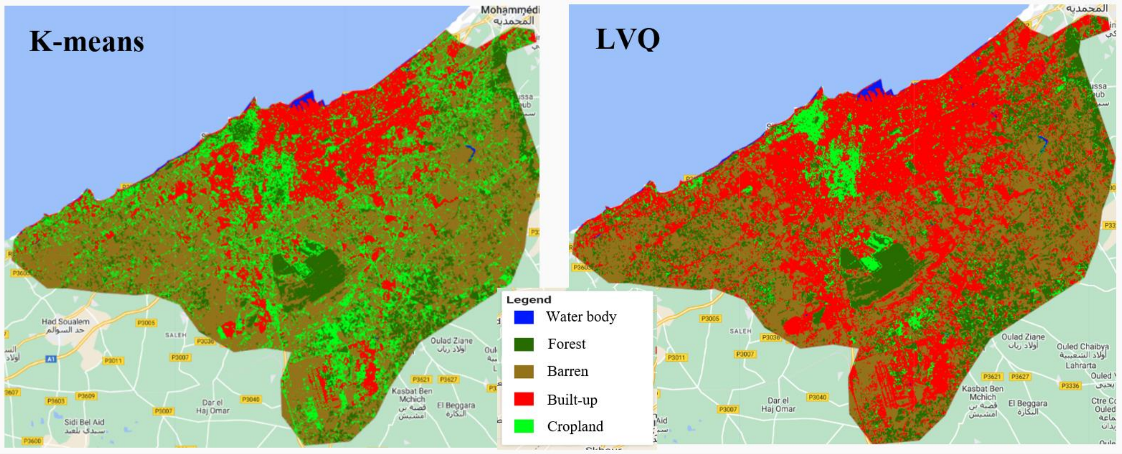

Figure 4: Land cover classification of Casablanca using K-means (left) and LVQ (right)

From these results, we can say that the K-means algorithm performs better than the LVQ algorithm, but we concluded that unsupervised learning algorithms do not perform well for land cover mapping. On the other hand, in these previous studies and in our previous work [15] and [35], we demonstrated that supervised learning algorithms perform well for satellite image classification and land cover mapping. In our future work, we will seek to improve these algorithms. We will try to integrate other techniques to improve the algorithm's performance. Strategies under consideration include:

- Data fusion: Integrating auxiliary data, such as high-resolution images or complementary spectral datasets, could improve class separability.
- Advanced clustering approaches: Exploring more advanced clustering methods, such as spectral clustering or Gaussian mixture models, could yield superior results in complex landscapes.
- Ensemble techniques: The use of ensemble methods, which combine the results of several algorithms, could leverage their respective strengths to achieve higher classification accuracy.
- In conclusion, the K-means algorithm applied to Landsat 8 OLI imagery for land cover classification in Casablanca showed moderate accuracy. While these results are valid, future research should refine the pre-processing steps, optimize feature selection, and further investigate sophisticated clustering techniques to achieve more precise and accurate results.

5. Conclusion

In conclusion, this paper presents the successful application of unsupervised learning techniques and Google Earth Engine (GEE) for land use mapping in the urban context of Casablanca. The use of Landsat multispectral imagery, combined with the k-means and LVQ clustering algorithm, enabled the extraction of meaningful land cover classes without the need for manual labeling, demonstrating the effectiveness of unsupervised methods for large-scale urban land cover analysis. The results of the study provide valuable information on the composition of Casablanca's land cover, highlighting the distribution and extent of the different land cover categories, including urban, forest, agriculture, water bodies and barren areas. The resulting land use map contributes to a comprehensive understanding of the city's spatial organization, enabling informed urban planning, sustainable development, and environmental management.

The efficiency and automation offered by GEE streamlines the data acquisition, pre-processing, and analysis process, making it a powerful tool for researchers and practitioners working on similar studies. GEE accessibility and potential for customization and scalability further enhance its applicability to complex urban challenges.

Although this article demonstrates the performance of unsupervised learning algorithms, in particular K-means and LVQ, and demonstrates the capability of GEE for land use analysis and mapping in Casablanca, there is scope for further refinement and improvement. Future research could explore the integration of additional data sources, such as higher-resolution images or ancillary data, to improve classification accuracy and capture land-use details at a finer scale. In addition, a comparison of unsupervised results with ground truth data or other classification methods could provide a more comprehensive validation of the approach. In summary, this study contributes to the growing body of knowledge in remote sensing and geospatial analysis, highlighting the importance of automated techniques for characterizing urban land use. By leveraging GEE capabilities, researchers and stakeholders can make informed decisions that contribute to the sustainable development and environmental well-being of rapidly evolving urban areas such as Casablanca.

References

[1] Abualghanam, O., Adwan, O. Y., Al Shariah, M. A. and Qatawneh, M., (2022). Enhancing the Speed of the Learning Vector Quantization (LVQ) Algorithm by Adding Partial Distance Computation. *Cybernetics and Information Technologies*, Vol. 22(2), 36–49. https://doi.org/10.2478/CAIT-2022-0015.

[2] Ali, I., Rehman, A. U., Khan, D. M., Khan, Z., Shafiq, M. and Choi, J. G., (2022). Model Selection Using K-Means Clustering Algorithm for the Symmetrical Segmentation of Remote Sensing Datasets. *Symmetry*, Vol. 14(6). https://doi.org/10.3390/sym14061149.

[3] Artelt, A. and Hammer, B. (2020). Efficient Computation of Counterfactual Explanations of LVQ Models. *ESANN 2020 - Proceedings, 28th European Symposium on Artificial Neural Networks, Computational Intelligence and Machine Learning*. 19–24.

[4] Celik, T., (2009). Unsupervised change Detection in Satellite Images Using Principal Component Analysis and κ-means Clustering. *IEEE Geoscience and Remote Sensing Letters*, Vol. 6(4), 772–776. https://doi.org/10.1109/LGRS.2009.2025059.

[5] Coefficient Kappa de Cohen (K), (2023). IRDP: Institut de Recherche Et De Documentation Pédagogique. Available at: https://www.irdp.ch/institut/coefficient-kappa-cohen-2039.html.

[6] Duda, T. and Canty, M., (2002). Unsupervised Classification of Satellite Imagery: Choosing a Good Algorithm. *International Journal of Remote Sensing*, Vol. 23(11), 2193–2212. https://doi.org/10.1080/01431160110078467.

[7] Duro, D. C., Franklin, S. E. and Dubé, M. G., (2012). 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 Sensing of Environment*, Vol. 118, 259–272. https://doi.org/10.1016/j.rse.2011.11.020.

[8] Gascon, M., Cirach, M., Martinez, D., Dadvand, P., Valentin, A., Plasencia, A. and Nieuwenhuijsen. M. J., (2016). Normalized Difference Vegetation Index (NDVI) as a Marker of Surrounding Greenness in Epidemiological Studies: The Case of Barcelona City. *Urban Forestry & Urban Greening*. Vol. 19, 88–94. https://doi.org/10.1016/J.UFUG.2016.07.001.

[9] Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D. and Moore, R., (2017). Google Earth Engine: Planetary-Scale Geospatial Analysis for Everyone. *Remote Sensing of Environment*, Vol. 202, 18–27. https://doi.org/10.1016/J.RSE.2017.06.031.

[9] Heydarian, M., Doyle, T. E. and Samavi, R., (2022). MLCM: Multi-Label Confusion Matrix. *IEEE Access*, Vol. 10, 19083–19095. https://doi.org/10.1109/ACCESS.2022.3151048

[10] Jamali, A., (2019). 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 Applied Sciences*, Vol. 1(11). https://doi.org/10.1007/s42452-019-1527-8.

[11] Landsat 8 Bands: Landsat Science: Available at: https://landsat.gsfc.nasa.gov/satellites/landsat-8/landsat-8-bands/# (Accessed: 30 January 2023).

[12] Li, N., Martin, A. and Estival, R., (2018). Combination of Supervised Learning and Unsupervised Learning Based on Object Association for Land Cover Classification. *2018 Digital Image Computing: Techniques and Applications (DICTA), Canberra, ACT, Australia*, 1-8. https://doi.org/10.1109/DICTA.2018.8615871.

[13] Lynch, P., Blesius, L. and Hines, E., (2020). Classification of Urban Area Using Multispectral Indices for Urban Planning. *Remote Sensing*, Vol. 12(15), 1-28. https://doi.org/10.3390/RS12152503.

[14] Masek, J. G., Huang, C., Wolfe, R. E., Cohen, W. B., Hall, F. G., Kutler, J. and Nelson, P., (2008). North American Forest Disturbance Mapped from A Decadal Landsat Record. *Remote Sensing of Environment*, Vol. 112(6), 2914–2926. https://doi.org/10.1016/j.rse.2008.02.010.

[15] Masoumi, F., Eslamkish, T., Honarmand, M. and Abkar, A. A. (2017). A Comparative Study of Landsat-7 and Landsat-8 Data Using Image Processing Methods for Hydrothermal Alteration Mapping, *Resource Geology*, Vol. 67(1), 72–88. https://doi.org/10.1111/rge.12117.

[16] NDBI—ArcGIS Pro Documentation, (2023). https://pro.arcgis.com/en/pro-app/latest/arcpy/spatial-analyst/ndbi.htm. (Accessed: 19 May 2023).

[17] Mfondoum, A. H., Etouna, J., Nongsi, B. K., Moto, F. A. and Deussieu, F. G., (2016). Assessment of Land Degradation Status and its Impact in Arid and Semi-Arid Areas by Correlating Spectral and Principal Component Analysis Neo-Bands. International Journal of Advanced Remote Sensing and GIS, Vol. 5(1), 1539–1560. https://doi.org/10.23953/CLOUD.IJARSG.77.

[18] Ouchra, H. and Belangour, A., (2021). Object Detection Approaches in Images: A Survey. *Proceedings Thirteenth International Conference on Digital Image Processing (ICDIP 2021)*. Vol. 11878, 132–141. https://doi.org/10.1117/12.2601452.

[19] Ouchra, H. and Belangour, A. (2021). Object Detection Approaches in Images: A Weighted Scoring Model based Comparative Study. *International Journal of Advanced Computer Science and Applications,* Vol. 12(8), 268–275. https://doi.org/10.14569/IJACSA.2021.0120831.

[20] Ouchra, H. and Belangour, A., (2021). Satellite Image Classification Methods and Techniques: A Survey. *2021 IEEE International Conference on Imaging Systems and Techniques (IST), Kaohsiung, Taiwan*. 1-6, https://doi.org/10.1109/IST50367.2021.9651454.

[21] Ouchra, H., Belangour, A. and Erraissi, A., (2022). A Comparative Study on Pixel-based Classification and Object-Oriented Classification of Satellite Image. *International Journal of Engineering Trends and Technology*, Vol. 70, 206–215. https://doi.org/10.14445/22315381/IJETT-V70I8P221.

[22] Ouchra, H., Belangour, A. and Erraissi, A., (2022). A Comprehensive Study of Using Remote Sensing and Geographical Information Systems for Urban Planning. *Internetworking Indonesia Journal*, Vol. 14(1), 15–20.

[23] Ouchra, H., Belangour, A. and Erraissi, A. (2022) 'Machine learning for satellite image classification: A comprehensive review', *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, pp. 1–5. Available at: https://doi.org/10.1109/ICDABI-56818.2022.10041606.

[24] Ouchra, H., Belangour, A. and Erraissi, A., (2022). Satellite Data Analysis and Geographic Information System for Urban Planning: A Systematic Review. *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 558–564. https://doi.org/10.1109/ICDABI56818.2022.10041487.

[25] Ouchra, H., Belangour, A. and Erraissi, A. (2022). Spatial Data Mining Technology for GIS: A Review. *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 655–659. https://doi.org/10.1109/ICDABI56818.2022.10041574.

[26] Ouchra, H., Belangour, A. and Erraissi, A., (2022). Spatial Data Mining Technology for GIS: A Review. *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 655–659. https://doi.org/10.1109/ICDABI56818.2022.10041574.

[27] Ouchra, H., Belangour, A. and Erraissi, A. (2022). Machine Learning for Satellite Image Classification: A Comprehensive Review. *2022 International Conference on Data Analytics for Business and Industry (ICDABI)*, 1–5. https://doi.org/10.1109/ICDABI56818.2022.10041606.

[28] Vaibhav, J., (2020). Performance Metrics: Confusion Matrix, Precision, Recall, and F1 Score. Towards Data Science. https://towardsdatascience.com/performance-metrics-confusion-matrix-precision-recall-and-f1-score-a8fe076a2262 (Accessed: 10 August 2023).

[29] Rashed, T. and Jürgens, C., (2010). Remote Sensing of Urban and Suburban Areas. Dordrecht: Springer Netherlands (Remote Sensing and Digital Image Processing). https://doi.org/10.1007/978-1-4020-4385-7.

[30] Reddy, G.P.O. (2018) 'Satellite Remote Sensing Sensors: Principles and Applications', pp. 21–43. Available at: https://doi.org/10.1007/978-3-319-78711-4_2.

[31] Reddy, G. P. O., (2018). Satellite Remote Sensing Sensors: Principles and Applications. Geospatial Technologies in Land Resources

[32] Mapping, Monitoring and Management. Springer, 21–43. https://doi.org/10.1007/978-3-319-78711-4_2.

[33] Semadi, P. N. A. and Pulungan, R., (2019). Improving Learning Vector Quantization Using Data Reduction. *International Journal of Advances in Intelligent Informatics*, Vol. 5(3), 218–229. https://doi.org/10.26555/ijain.v5i3.330.

[34] Spectral Indices with Multispectral Satellite Data. (2023). https://www.geo.university/pages/blog?p=spectral-indices-with-multispectral-satellite-data (Accessed: 22 August 2023).

[35] Abbas, A. W., Minallh, N., Ahmad, N., Abid, S. A. R. and Kha, M. A. A. (2016). K-Means and ISODATA Clustering Algorithms for Landcover Classification Using Remote Sensing. Sindh Univ. Res. Jour. (Sci. Ser.), Vol. 48(2). 315-318. https://www.researchgate.net/publication/303971825.

[36] Yang, L., Driscol, J., Sarigai, S., Wu, Q., Chen, H. and Lippitt, C. D., (2022). Google Earth Engine and Artificial Intelligence (AI): A Comprehensive Review. *Remote Sensing*, Vol. 14(14), https://doi.org/10.3390/RS14143253.

[37] Yengoh, T., Dent, D. L., Olsson, L., Tengberg, A. E., and Tucker, C. J., (2016). Use of the Normalized Difference Vegetation Index (NDVI) to Assess Land Degradation at Multiple Scales. Springer

[38] Zurqani, H. A., Post, C. J., Mikhailova, E. A. and Allen, J. S., (2019). Mapping Urbanization Trends in a Forested Landscape Using Google Earth Engine. *Remote Sensing in Earth Systems Sciences*, Vol. 2(4), 173–182. https://doi.org/10.1007/s41976-019-00020-y.