ADVANCES IN GEODESY AND GEOINFORMATION
(formerly GEODESY AND CARTOGRAPHY)
Vol. 74, no. 1, article no. e61, 2025

https://doi.org/10.24425/agg.2025.150691

Original paper

Part of Topical collection:
"Advancements in Applied Geoinformatics"

Integration of LIME with segmentation techniques for SVM classification in Sentinel-2 imagery

Elif Ozlem Yilmaz*, Yakup Kaan Uzun, Emre Berkan, Taskin Kavzoglu

Gebze Technical University, Kocaeli, Turkey
e-mail: eoyilmaz@gtu.edu.tr; ORCID: http://orcid.org/0000-0002-6853-2148
e-mail: y.uzun2020@gtu.edu.tr; ORCID: http://orcid.org/0009-0007-2556-8363
e-mail: e.berkan2019@gtu.edu.tr; ORCID: http://orcid.org/0009-0009-9861-6113
e-mail: kavzoglu@gtu.edu.tr; ORCID: http://orcid.org//0000-0002-9779-3443

*Corresponding author: Elif Ozlem Yilmaz, e-mail: eoyilmaz@gtu.edu.tr

Received: 2024-10-15 / Accepted: 2025-02-10

Abstract: Recent advancements in remote sensing technology have facilitated the acquisition of images with higher spatial resolution. In response to this rapid technological evolution, the paradigm of OBIA has emerged as a key approach. An essential component of OBIA is image segmentation, where the careful selection of an appropriate segmentation algorithm and its parameters significantly influences the quality of the segmentation output. This study aims to conduct LULC analysis on Sentinel-2 imagery and compare the accuracy of the SVM classifier across different segmentation methods produced by MRS, SLIC, Mean Shift, and Quick Shift algorithms. The selected study area is located in the Marmara region of Turkey and characterized by seven major LULC classes. The segmentation was conducted though four algorithms, with 60 segment features being extracted for each output, considering spectral, textural, and geometric attributes separately. Following the classification process with SVM, overall accuracies of 96.14% for the MRS, 91.00% for the SLIC, 89.95% for the Mean Shift and 87.95% for the Quick Shift approach were estimated. These results underscore the superior performance of the MRS algorithm with significant level of improvement. This high level of accuracy holds significant potential for delivering more dependable and precise outcomes in planning and decision-making processes. Moreover, integrating XAI, specifically the LIME algorithm, enhances the transparency and comprehensibility of classification analysis within the OBIA framework. Features associated with the NIR and SWIR bands were found to have predominantly positive effects. This integration contributes to improved transparency, enabling more informed and reliable decision-making processes.

Keywords: LIME, multiresolution segmentation, simple linear iterative clustering, Mean Shift, Quick Shift

1. Introduction

In recent decades, the exponential growth of the human population has led to significant changes on Earth's surface. Most of the world's population now prefers to live in urban settlements, which is referred to urbanization, due to the social, economic, and environmental reasons. This is a major change from the past, when people predominantly lived in rural areas. It is reported that by the year of 2050, urbanization is projected to reach 64% in developing countries and 86% in industrialized nations (Humbal et al., 2023). Cities not only change the natural environment in which they are located, but also affect the well-being of the people they live in. Therefore, the rapid population growth, particularly in urban areas, combined with increasing industrialization, has made the efficient management of natural resources and the monitoring of environmental changes crucial. Hence, remote sensing techniques have become a powerful tool for observing and studying changes on Earth's surface (Steinhausen et al., 2018). Technological improvements in this field have greatly enhanced the ability to acquire multispectral images at higher spatial, temporal, and spectral resolutions (Tassi et al., 2021; Kavzoglu et al., 2024). These advances provide a unique opportunity to systematically detect and monitor up-to-date changes in land use and land cover (LULC) on a large scale (Alshari and Gawali, 2021). LULC maps are an important tool for understanding the impact of human activities on the ecosystem and highlight the relationship between human socio-economic activities and natural ecological processes (Zhu et al., 2022).

There are several freely available EO satellites, including Sentinel and Landsat, which are frequently employed for LULC monitoring (Gani et al., 2023). The LULC classification methodology can be complex because of the different resolutions of the images obtained from satellites, classification techniques, and numerous image processing software used in the selection of sample sets for training (Mohan et al., 2021). Furthermore, various types of classification methodologies, such as contextual-based, supervised, and unsupervised methods, require comprehensive automation and have specific constraints (Quader, 2019). However, the rapid development of this technology has led to a significant increase in the amount of data, requiring the development of creative methods for image analysis. Object-Based Image Analysis (OBIA) is an important strategy that provides effective solutions for data interpretation and classification. It is a widely used remote sensing approach that classifies pixels in images according to their common characteristics and merges them into objects or regions (Kavzoglu and Tonbul, 2017). The first essential step in OBIA is image segmentation because the quality of segmentation mainly affects the subsequent classification results (Pal and Pal, 1993; Kavzoglu et al., 2017; Su and Zhang, 2017). Image segmentation involves the division of images into homogeneous regions, which are referred to as "image objects or segments". There are LULC classes that include buildings, agricultural parcels, wetlands, trees, roads, and grasslands. The process involves dividing images into regions that have comparable characteristics and assigning these regions to different object classes in the classification stage (Kavzoglu and Tonbul, 2018). By integrating image segmentation techniques with OBIA, it is possible to achieve greater classification accuracy (Kavzoglu et al., 2016).

The existing literature contains segmentation algorithms based on various fundamental principles for image processing applications. Multi-resolution segmentation (MRS), simple linear iterative clustering (SLIC), Mean Shift, and Quick Shift are frequently employed techniques (Kavzoglu et al., 2024; Yilmaz and Kavzoglu, 2024). The MRS approach is commonly used, and it relies on three parameters: shape, compactness, and scale. On the other hand, SLIC, which is another widely used segmentation algorithm, is inspired by the k-means algorithm and aims to group image pixels into segments or clusters, considering both color similarity and spatial proximity within the image plane (Tonbul and Kavzoglu, 2020a). Furthermore, the Mean Shift algorithm clusters pixels according to their similarities by shifting data points towards high-density regions, whereas the Quick Shift algorithm follows a local structure and performs hierarchical segmentation based on color and spatial similarities between pixels, providing faster but generally coarser segmentation (Comaniciu and Meer, 2002; Vedaldi and Soatto, 2008; Pipaud and Lehmkuhi, 2017). All segmentation methodologies have advantages and disadvantages. Therefore, selecting the segmentation method based on the objective of the study is crucial. Classification is the second stage in OBIA process, for which various machine learning techniques have been reported in the literature, including support vector machine (SVM), random forest (RF), and boosting algorithms.

SVM is a robust classification method that can distinguish linear and nonlinear distinctions on complex datasets. It is widely used in various fields, including LULC analysis (Kavzoglu and Colkesen, 2009; Aryal et al., 2023). Machine learning methods like SVM used in the LULC classification stage can be considered black boxes due to their many parameters and computational complexity. Making the decision-making process of these algorithms understandable is essential for the reliability and reproducibility of the research. For this purpose, explainable artificial intelligence (XAI) has become a popular topic, focusing on interpreting machine learning models to test the reliability of algorithms' working principle. There are two primary approaches for interpreting machine learning models. The first one involves the use of the entire data set to perform a global interpretation, while the second one involves the selection of samples from the data set to carry out the interpretation (Ishikawa et al., 2023; Teke and Kavzoglu, 2024).

The aim of this study is to present a comparative performance analysis of LULC maps produced with different segmentation algorithms using advanced remote sensing techniques, shedding light on the importance of segmentation algorithms. Thus, the potential of XAI to increase the transparency and reliability of classification processes within the OBIA framework is investigated. The present study offers an original contribution to the existing literature on this subject, both in terms of its methodological perspective and the specific LULC designated for the study area. The novelty of this study lies in the integration of segmentation-based LULC classification with the XAI technique. In this context, this study focuses on LULC analysis using Sentinel-2 imagery and evaluated the performance of the SVM classifier for different segmentation methods (e.g. MRS, SLIC, Mean Shift and Quick Shift). In addition, XAI approach was incorporated into OBIA to improve the clarity and understandability of the classification analysis, specifically using the local interpretable model-agnostic explanations (LIME) algorithm. In other

words, the effect of segmentation parameters on the classification results is evaluated using XAI methods such as LIME. This approach enables the determination of which segment feature contributes most to the accuracy of the model. Furthermore, this study can provide useful information for parameter optimizations, allowing better segmentation and classification results specific to the study areas.

2. Study area and dataset

The study area located in the Gebze district of Kocaeli province covering a total area of approximately 500 km^2 land in the Marmara region of Turkey, which is densely populated and of great economic and social importance in Turkey, was selected for a comprehensive study of its LULC (Fig. 1). Seven land classes, namely, water, road, forest, grassland, soil, white-roof, and red-roof, covering the bulk of the study area were identified after

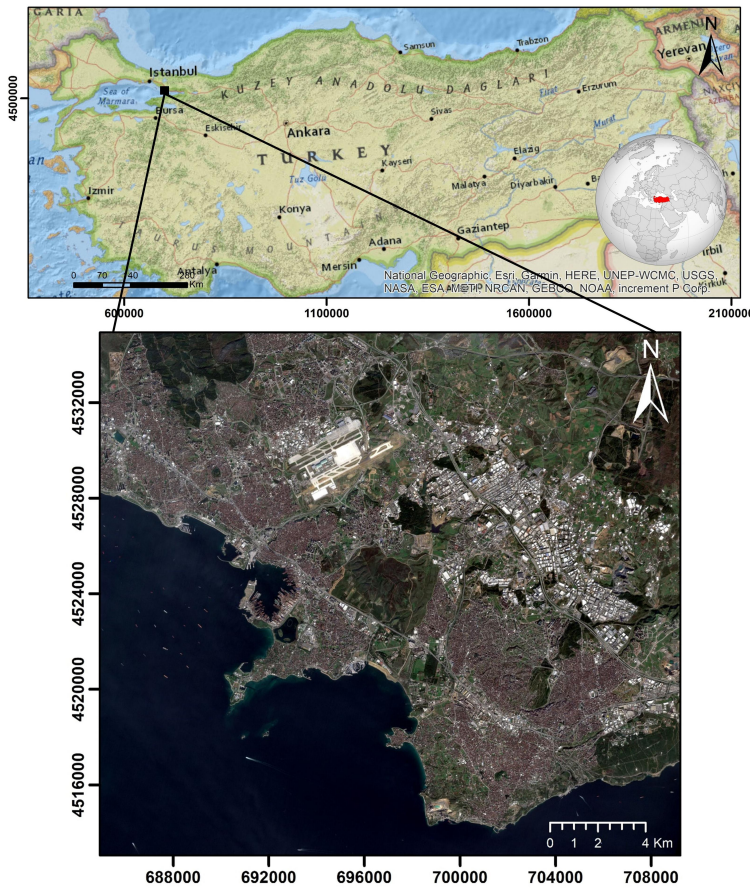

Fig. 1. The area of the study: Gebze district of Kocaeli provience in Turkey. The coordinate system: WGS 1984 UTM Zone 35N

field visits and visual interpretation of aerial images. The region contains red/white roof buildings used for both residential and industrial purposes due to its proximity to a high density city. The LULC class termed 'red roof' was typically employed for residential or public buildings, whilst the 'white roof' class was generally utilized for industrial buildings in this study. Moreover, the proximity to the industrial zone has resulted in a high level of road infrastructure in the study area. Despite the industrial activities, there are dense areas of forest and other vegetation in the region. There are also significant areas of bare ground. Finally, as the study area is located on the coast of the Sea of Marmara, the water class was used to represent the pixel on the sea surface. Samples were collected from homogeneous regions in order to represent the LULC class. 750 training samples and 250 test samples were collected separately for each LULC class to be used in the classification stage. These samples were selected before the segmentation stage and assigned to the segments.

Sentinel-2 Level 2A imagery acquired by the European Space Agency on December 13, 2023, was used in local coordinate system of the World Geodetic System 1984 / Universal Transverse Mercator (WGS 84/UTM). The image was subjected to atmospheric and radiometric corrections, with a radiometric resolution of 12 bits and 13 bands of varying spatial resolution (i.e., 4 bands at 10 m, 6 bands at 20 m, and 3 bands at 60 m). In particular, the 60 m resolution bands were excluded from the pre-processing, resulting in the use of 10 m and 20 m spectral bands for subsequent analysis (Table 1). The bands with 60 m resolution have been excluded from the dataset on the grounds of their low spatial resolution and their irrelevancy for LULC map classification. Also, bands with 10 m and 20 m resolutions are more suitable for data-based analysis in determining the LULC classes of the study area. The 20 m bands were resampled to 10 m resolution using the Gram-Schmidt pan sharpening technique (Tonbul et al., 2020). All the selected spectral bands were then combined using the layer stack function. Additionally, samples were collected separately for the seven LULC classes using the QGIS program. 70% of the collected samples were used for training, and 30% were used to check the accuracy of the produced thematic maps.

Table 1. Properties of Sentinel 2A spectral bands used in the study

Band number	Spectral region	Wavelength (nm)	Resolution (m)
B2	blue	458–523	10
B3	green peak	543–578	10
B4	red	650–680	10
B5	red edge	698–713	20
B6	red edge	733–748	20
B7	red edge	773–793	20
B8	NIR	785–899	10
B8a	NIR narrow	855–875	20
B11	SWIR	1565–1655	20
B12	SWIR	2100–2280	20

3. Methodology

In this study, OBIA was used to produce LULC maps from Sentinel-2A imagery. Instead of analyzing image pixels individually, OBIA creates more meaningful and homogeneous data structures by grouping neighboring pixels into object segments (Kavzoglu et al., 2024). The proposed methodology can be used to produce more accurate LULC maps, particularly for complex areas. The satellite image was segmented using different segmentation algorithms, and LULC maps were generated using machine learning algorithms. In addition, classification performance was compared by calculating the class-based F-score values, overall accuracy and Kappa coefficient of each segmentation method. Finally, LIME, one of the Explainable Artificial Intelligence (XAI) techniques, was used to better understand the classification results and make the model's decision-making process transparent. It was used to explain the features on which classification models based their decisions and to analyze the impact of segment features on classification. In this way, not only the accuracy but also the explainability of the model's decisions were evaluated. A flowchart of the above-metioned methodology adopted in this study is shown in Figure 2.

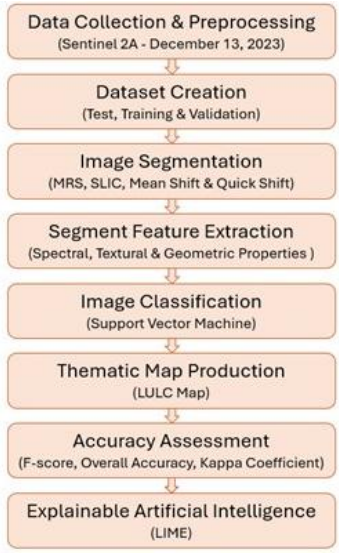

Fig. 2. The workflow of the study

3.1. Segmentation Algorithms

Segmentation is a method of dividing images into a spatially coherent collection of objects, each consisting of an adjacent group of pixels that exhibit uniformity or semantic relevance (Tonbul and Kavzoglu, 2020a; Lourenço et al., 2021; Kavzoglu et al., 2024). OBIA classifications primarily include image segmentation and classification in that image segmentation directly affecting the classification accuracy (Feng et al., 2023). Therefore, the segmentation stage is the most important stage in OBIA applications.

In this study, four algorithms were used in the segmentation stage. In order to observe the different working principles of the segmentation algorithms used in the study and how the results affect the LULC classification, MRS, SLIC, Mean Shift and Quick Shift algorithms were applied. While MRS plays an important role in the analysis of large areas by creating more homogeneous segments, the SLIC algorithm creates small and detailed segments in regions with complex LULC classes such as urban areas. On the other hand, the Mean Shift and Quick Shift algorithms operate in a similar way, although Mean Shift extracts regional boundaries more accurately, while Quick Shift performs this process more quickly and in greater detail.

3.1.1. Multi-Resolution Segmentation

Multi-resolution segmentation (MRS) provides a comprehensive perspective by examining objects or areas at multiple resolution levels (Baatz and Schäpe, 2000). The first step in MRS is to downsample the dataset or image to different resolution levels. The reduction process is often performed within a pyramidal framework. Each level represents an image or data of larger size and reduced resolution. At each resolution level, the image is segmented into sections with analogous properties. This segmentation is often related to the similarity of pixels in terms of color, brightness, texture, and other attributes. Similar areas derived from the segmentation process are transformed into segments by merging or dividing. The segments specify the characteristics at each resolution level. This creates a hierarchy that illustrates the relationships between segments at different resolution levels (Tonbul and Kavzoglu, 2020b).

3.1.2. Simple Linear Iterative Clustering

Simple linear iterative clustering (SLIC) algorithm is an exceptionally rapid and effective image segmentation technique that analyzes low-level attributes. The SLIC algorithm segments images by merging pixels that exhibit similar color and spatial characteristics (Achanta et al., 2010). In the SLIC method, the initial centers are selected to be uniformly distributed across a specified distance range within the image. Each of these starting points represents an area. Each pixel is aligned with the closest centroid based on its color and spatial coordinates via updating. This update is determined as the weighted average of the pixels inside the target area. Therefore, regional centers are updated iteratively. Pixels exhibiting similar color attributes and spatial proximity are likely to be part of the same area. This integration facilitates the formation of homogenous and uniform areas. The merging and updating operations proceed iteratively until a specified convergence threshold is achieved (Kavzoglu and Tonbul, 2018).

3.1.3. Mean Shift

The Mean Shift is a non-parametric method designed to study complicated spaces. The framework relies on three primary parameters: the spatial radius to define the neighborhood, the interval radius to define the color space interval, and the minimum size of the areas to be preserved after grouping (Cheng, 1995). The settings are modified to construct segments based on a pixel, forming a collection of adjacent pixels within a specified spatial

radius and color range. In the Mean Shift algorithm, the initial positions of the objects are used as reference points. The kernel window is then positioned near the selected start location. A weighted average is calculated to evaluate the pixel intensity within the window in feature space. The weighted average is determined by evaluating the location of each pixel in the feature space and its corresponding intensity at that location. This procedure results in a shift of the pixels within the window towards the intensity centers of the objects. This iterative process moves objects closer to their density centers. As the algorithm continues to be applied, it often reaches a convergence criterion (Comaniciu and Meer, 2002; Su et al., 2015).

3.1.4. Quick Shift

The Quick Shift algorithm uses a region-based method that merges pixels with similar attributes. The radian value is calculated to assess the similarity between the color and spatial location of each pixel. The radian values quantify the color similarity and spatial proximity of pixels. The algorithm aggregates pixels within a specified distance radius with analogous color and spatial characteristics. A zone is defined by aggregating similar pixels in the color space. The merging process continues until a specified threshold is reached, and this iterative process continues during the construction of regions. This iteration improves segmentation by clustering similar pixels. The merging process terminates when a certain threshold is reached. The threshold determines the similarity level required to continue the merging process (Vedaldi and Soatto, 2008).

3.1.5. Configuration of Segmentation Algorithms

In this study, the results of classification and segmentation techniques applied to Sentinel-2 satellite imagery were analyzed in detail. In the segmentation step, two methods were first applied independently. The Automated Estimation of Scale Parameter (ESP2) tool has been developed for the purpose of determining the scale parameter for the MRS algorithm. The MRS algorithm proceeded by determining a scale parameter of 40 using the ESP2 tool proposed by Drăgut et al. (2014) (Fig. 3). The shape and compactness characteristics were

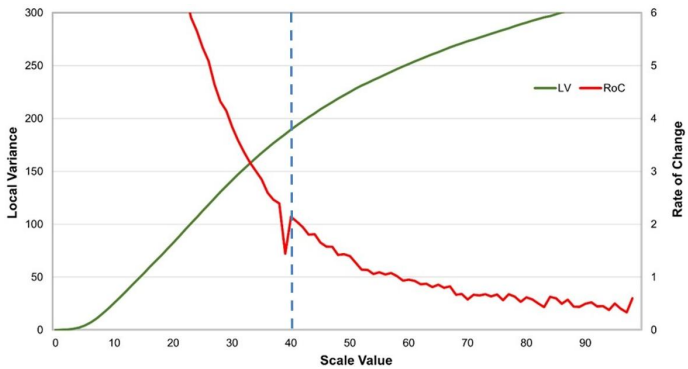

Fig. 3. Determination of optimal scale parameter for the MRS algorithm using ESP2 tool

set to 0.1 and 0.5, respectively. However, in the SLIC method, the number of segments was set to 150,000, and the compactness parameter was set to 1 by applying trial and error strategy. For the Mean Shift algorithm, the spatial radius was set to 5, while for the Quick Shift algorithm the kernel size was set to 1 and the maximum distance was set to 10. After the segmentation process, a total of 113,045, 143,138, 21,682, 87,913 segments were produced using the MRS, SLIC, Mean Shift, and Quick Shift algorithms, respectively.

3.2. Image Classification Algorithm

A variety of machine learning algorithms can be used for OBIA-based image classification, which is the next step after segmentation processing (Kavzoglu et al., 2018). SVM is selected due to its ability to operate with minimal data requirements and low computational demands, which are crucial in the context of resource-constrained environments for this study. Moreover, the utilization of SVM in this study was motivated by its demonstrated capacity for high accuracy, a feature that is particularly advantageous in analytical endeavours. It uses high-dimensional feature spaces and nonlinear decision boundaries. The fundamental premise of SVM is to identify the optimum hyperplane in the feature space that most effectively distinguishes data points from disparate classes. The hyperplane serves as the decision border. Its position is established by maximizing the margin, which is defined as the distance between the closest data points in each class (Cortes et al., 1995). SVM relies on the principle of linear discriminability, where the classification of data points as either above or below a line enables the prediction of the label for an unknown input. The approach can be mathematically extended to higher dimensions, resulting in the creation of a hyperplane. This hyperplane acts as a demarcation that differentiates various classes (Ben-Hur et al., 2008). In addition, the kernel trick is employed to enhance the algorithm performance. This approach facilitates the acquisition of non-linear decision boundaries by translating data points into higher-dimensional regions. To mitigate computing expenses in high-dimensional spaces, it executes this action via kernel functions rather than directly within such areas. The method enhances efficiency by minimizing computational expenses. The frequently used kernels are linear, polynomial, radial basis functions (RBF), and sigmoid, and each is appropriate for distinct data distributions (Schölkopf, 2000; Kavzoglu and Colkesen, 2009). The classifier attempts to optimize the margin in establishing the decision border while simultaneously ensuring accurate classification of the data points. It optimizes the distinction between classes and categorizes data points with the highest accuracy. Margin maximization enhances the generalization capacity of the classifier and mitigates the risk of overfitting. The regularization parameter is employed to regulate the model's complexity (Evgeniou et al., 2000). The SVM classifer used in this study was applied using the Python Programming Language and the Sckit-learn libraries. Linear kernel was chosen as it was sufficient to discriminate LULC classes after segmentation, thus reducing the computational cost. 'c=1' was used for the regularization, which prevented overfitting and increased the generalisation capacity of the model.

3.3. Image Explainable Artificial Intelligence (XAI)

XAI can be used to make the decision-making processes of the deep learning models used in OBIA understandable. By revealing the features and parameters that underlie the model's predictions and decisions, XAI allows users to understand how the model works and why it reaches certain conclusions. Machine learning (ML) models are often referred to as "black boxes" because their inner workings and decision-making mechanisms are often complex and difficult to understand. These models contain many parameters and complex calculations, making it difficult to understand why and how decisions are made. The use of XAI helps to explain the inner workings and decision-making mechanisms of these "black box" models. This allows users to better understand these models and make more informed decisions. ML algorithms make predictions by learning from data sets. For example, algorithms such as Random Forest and SVM are often used to process complex and multidimensional data sets. These algorithms can be used in conjunction with XAI techniques to identify features and parameters that support model's predictions and decisions (Ghnemat et al., 2023).

The application of XAI in image analysis aims to explain the inner workings and decision-making mechanisms of the model. This approach increases the reliability of models and helps to explain the logic of the decisions made by the model, as well as the reasons and ways of making the decisions. The use of XAI increases the transparency and reliability of the model while allowing users to better understand the data and make more informed decisions. XAI can be divided into global and local analysis methods. Global methods are used to understand the overall performance of the model and calculate the contribution of each feature to the prediction. Local methods are used to comprehend the prediction of specific data points. In the field of XAI, Local Interpretable Model Explanations (LIME) is a notable technique used to elucidate model decisions, and it is particularly useful for understanding the rationale behind black-box models such as SVM. By highlighting the significant features that influence classification decisions, LIME improves the interpretability of the models, thereby providing more reliable and descriptive results. XAI approaches, exemplified by LIME, play a central role in elucidating and understanding the decisions made by sophisticated classification models (Temenos et al., 2023). Such methodologies provide insights into how models categorize or predict specific samples, thereby facilitating deeper understanding and analysis (Teke and Kavzoglu, 2024).

4. Results and discussion

The segmentation process was completed with the MRS, SLIC, Mean Shift, and Quick Shift algorithms. For the purpose of visual evaluation, a sample area containing different LULC classes in Bayramoglu district was selected from the study area (Fig. 4). An examination of the segmentation results for this area reveals that the segments produced by MRS provide a suitable foundation for LULC classification. In contrast, the SLIC algorithm generated smaller, grid-like segments compared to MRS. The Mean Shift and Quick Shift algorithms, however, resulted in fewer segments and were less effective in representing

LULC. In conclusion, based on the findings of this analysis, it can be suggested that the MRS and SLIC algorithms are likely to demonstrate superior performance in LULC classification compared to alternative algorithms.

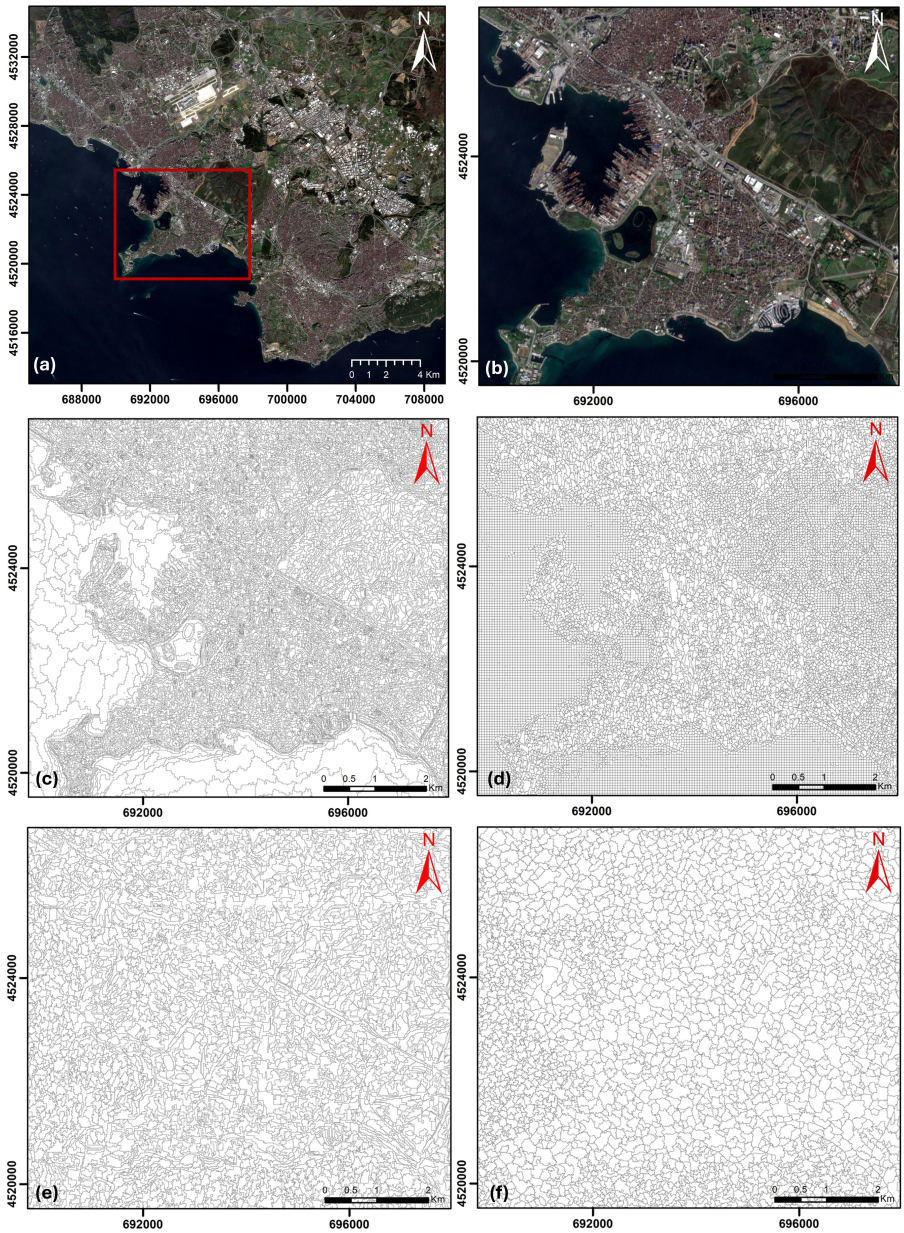

Fig. 4. Segmentation results of the study area and segmentation performed with different algorithms: (a) the study area, (b) a selected region from the study area for visual segmentation analysis, (c) MRS, (d) SLIC, (e) Mean Shift and (f) Quick Shift

After the segmentation stage, 60 features were extracted for each segment to apply classification. These features include area, border length, brightness, GLCM contrast, GLCM dissimilarity, GLCM entropy, GLCM homogeneity, HSI transformation, length, maximum difference, maximum pixel value (for each layer), mean (for each layer), minimum pixel value (for each layer), ratio (for each layer), and standard deviation (for each layer). The classification performance of the segmentation techniques using SVM was evaluated using metrics, including F-score, overall accuracy, and Kappa coefficient (Table 2). Results on the test dataset revealed that an overall accuracy of 96.14% for the MRS, 91.00% for the SLIC, 89.95% for the Mean Shift and 87.95% for the Quick Shift approach were estimated. The MRS method demonstrated higher overall accuracy compared to the other methods. Furthermore, the MRS approach produced a Kappa coefficient of 0.96, whereas the SLIC method produced a Kappa coefficient of 0.90. Also, the Kappa coefficient of the thematic maps produced by the other segmentation methods were calculated as 0.88 for the Mean Shift and 0.84 for the Quick Shift. It is clear from the OBIA classification results that the MRS methods were more consistent than the other methods.

Table 2. Class-based performance analysis for MRS, SLIC, Mean Shift and Quick Shift

F-score (%)	MRS	SLIC	Mean Shift	Quick Shift
Water	99.83	99.67	99.67	98.31
Road	91.08	80.82	79.59	74.35
Forest	98.99	98.50	97.35	95.13
Grassland	98.36	95.67	92.10	92.13
Soil	95.85	92.86	91.75	91.85
Red-roof	96.48	80.08	79.61	76.08
White-roof	92.42	89.44	89.68	88.23
Overall Acc. (%)	96.14	91.00	89.95	87.95
Kappa Coef.	0.96	0.90	0.88	0.84

Upon evaluation of the segmentation algorithms in accordance with the F-score (%) values presented in Table 2, notable disparities emerged with respect to each LULC class. For the water class, all algorithms demonstrated relatively high accuracies. While the MRS and SLIC algorithms exhibited the highest success, with 99.83% and 99.67%, respectively, the Quick Shift algorithm demonstrated an acceptable performance, although it exhibited lower accuracy (98.31%). This is likely due to the distinct spectral reflectance of the water class compared to the other land cover classes. A notable decline in performance was observed in the road class. The MRS algorithm yielded the most favorable results with a success rate of 91.08%. However, the SLIC (80.82%), Mean Shift (79.59%), and Quick Shift (74.35%) demonstrated comparatively inferior performance. The forest class also demonstrated high F-score values, with the MRS algorithm achieving the best result (98.99%). In contrast, the Quick Shift algorithm exhibited the lowest success rate (95.13%). Similarly, in the grassland class, the MRS algorithm demonstrated the greatest degree of success, with a score of 98.36%. On the contrary, the success of the other algorithms exhibited a more variable range, with scores between 95.67% and 92.10%. The performance of the algorithms in the soil class exhibits less discrepancy than the

other classes. The MRS algorithm yielded the highest F-score (95.85%), whereas the other algorithms obtained results between 91.75% and 92.86%. In the red-roof class, the MRS algorithm demonstrated the highest degree of success, with an F-Score of 96.48%. In comparison, the other algorithms exhibited lower F-scores, namely 80.08% for SLIC, 79.61% for Mean Shift, and 76.08% for Quick Shift. For the white-roof class, the SLIC, Mean Shift, and Quick Shift algorithms yielded relatively similar results. However, the MRS algorithm exhibited the highest degree of success, with an F-score of 92.42%. In conclusion, the results demonstrate that the MRS technique exhibits superior performance in accurately distinguishing between all LULC classes.

Figure 5 illustrates thematic maps produced by the MRS, SLIC, Mean Shift, and Quick Shift segmentation algorithms using the SVM classifier. The MRS method enables the identification of distinct classes, including water, forest, and red-roof, with clear borders. It was determined that the SLIC algorithm produced larger and more homogeneous segments when compared with other algorithms. Conversely, the Mean Shift algorithm divides segmentation into smaller segments, which leads to the merger of certain classes (e.g., grassland and bare soil). The Quick Shift method is less comprehensive and introduces difficulties, especially in the classification of road, white-roof, and bare soil classes. The MRS algorithm delivers superior detail and precision, whereas the Quick Shift algorithm exhibits worse performance; SLIC and Mean Shift algorithms yield moderate outcomes.

The contribution of each feature to the classification decision is shown alongside the feature importance ranking generated by LIME. The explainability of the LULC classification problem, in which four segmentation algorithms were used, was investigated with LIME (Fig. 6). The probability that a particular feature will either increase or decrease the value of the classification decision is expressed by the value provided for each feature. In contrast, negative values indicate that a particular feature has a negative effect on the categorization decision, and positive values indicate that the feature has a positive effect. For the MRS algorithm, the features ‘Mean9’, ‘Mean7’ and ‘Min3’ corresponding to the mean DN value of Band 9 (SWIR), mean DN value of Bant 7 (NIR) and minimum DN value of Band 3 (red) had the most significant positive influence on the system. The probability that the MRS algorithm reaches the correct classification conclusion increases proportionally to the values of these features reaching higher levels. On the other hand, features of ‘Max6’ (maximum DN values of red edge), ‘Max’ (maximum DN values of blue), and ‘Max5’ (maximum DN values of red edge) had the most negative impact. As the total values of these features continue to increase. The likelihood that the MRS model reaching the correct classification conclusion continues to decrease. ‘Mean6’ (mean DN values of red edge), ‘Mean9’ (mean DN values of SWIR), and ‘Mean7’ (mean DN values of red edge) are examples of properties that had the highest values for SLIC. There is a correlation between the values of these features and the probability that the SLIC algorithm will arrive at the correct classification choice. On the other hand, there are also features that have a negative influence, such as ‘Max7’ (maximum DN values of red edge), ‘Min8’(minimum DN values of NIR narrow) and ‘Max6’(maximum DN values of red edge). As the total values of these features continue to increase, the likelihood of SLIC estimating the correct categorization decreases.

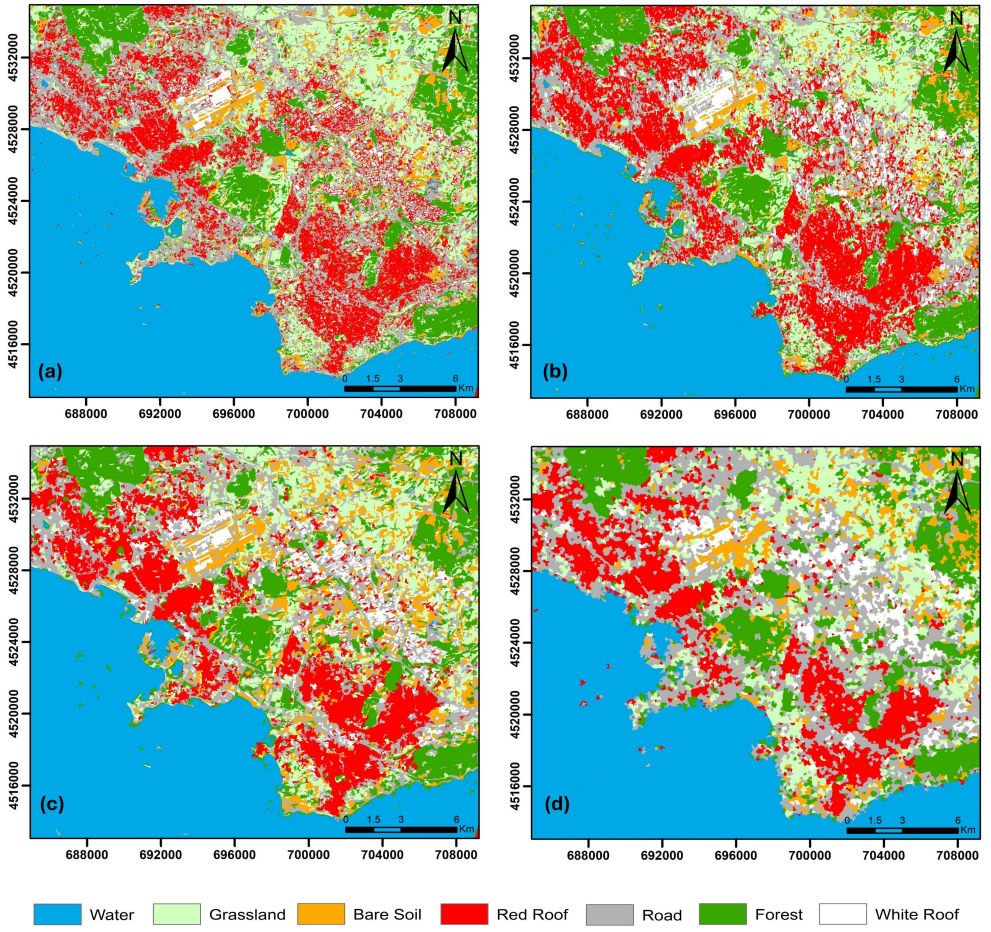

Fig. 5. Thematic maps produced with SVM classifier using (a) MRS, (b) SLIC, (c)Mean Shift, and (d) Quick Shift segmentation algorithms. The coordinate system: WGS 1984 UTM Zone 35N

The Mean Shift algorithm tended to highlight the 'Mean' (mean DN values of blue) and 'Max' (maximum DN values of blue) values among the features that have the most influence on the model's decision. Features such as 'Mean7' (mean DN values of NIR) and 'Max7' (maximum DN values of NIR narrow) have a positive impact on the model's prediction, whereas low 'Min' (minimum DN values of blue) values lead to predictions closer to the negative class. Compared to other algorithms, the Mean Shift performs segmentation based on pixel density; thus, classification decisions are based on pixel density and summary statistics of that density. This demonstrates that density-based features are more important in the algorithm's decision-making process than texture or geometric features. Furthermore, the Quick Shift algorithm follows a classification strategy based on both texture-based and statistical features. In addition to statistical summaries such as 'Min8' (minimum DN values of NIR narrow), 'Max8' (maximum DN values of NIR narrow) and 'Mean9'(mean DN values of SWIR), GLCM-based texture

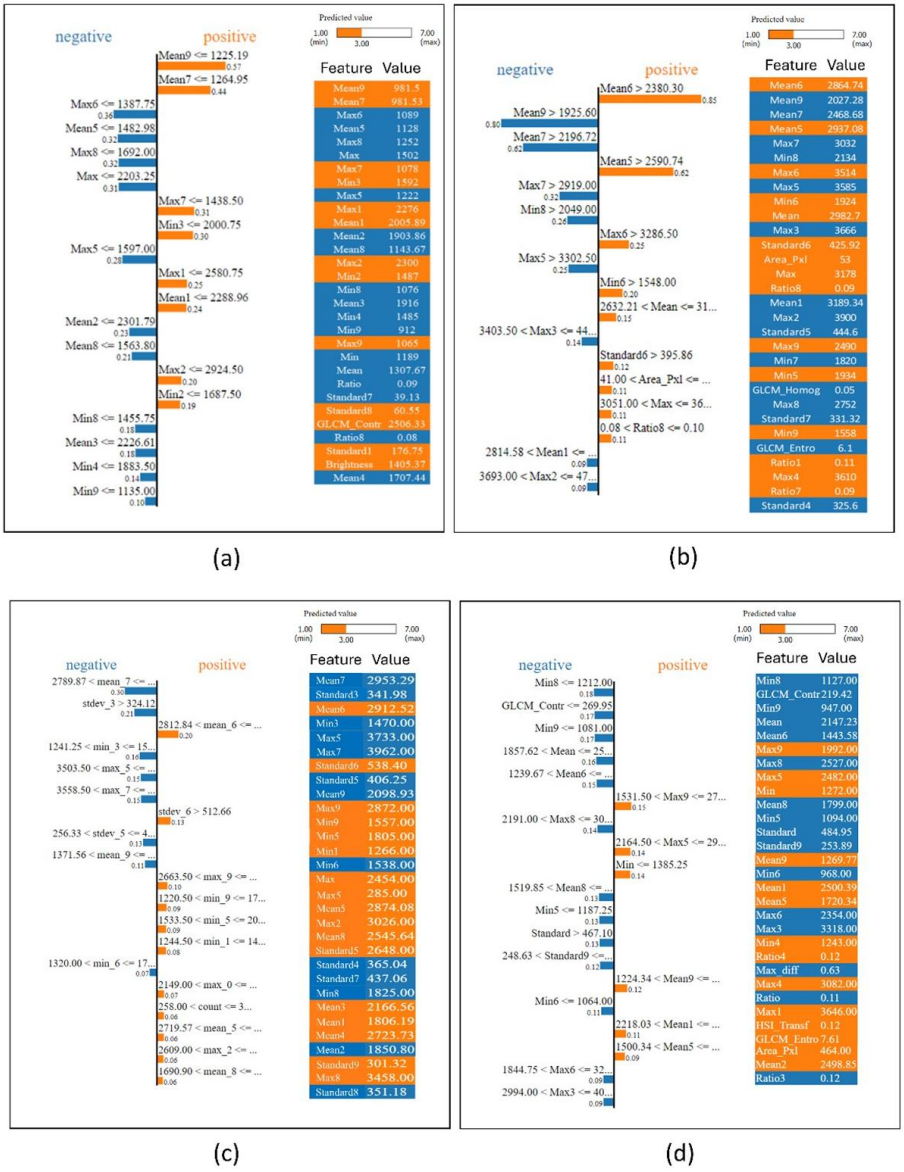

Fig. 6. LIME graphs for classifications using (a) MRS, (b) SLIC, (c) Mean shift and (d) Quick Shift segmentations. Note that the numbers next to the features in OBIA represent the bands used. For example, features without numbers (such as max, ratio, min) were calculated using the blue band. Features marked '2' are used for the green band, '3' for the red band, '4-5-6' for the red-edge bands, '7' for the NIR band, '8' for the narrow NIR band and '9-10' for the SWIR bands

features (i.e., 'GLCM Contrast') also have a significant impact on classification. This shows that the Quick Shift algorithm uses a segmentation strategy based on both the intensity of pixels and texture differences between pixels. In addition, geometric features ('Area') also contribute to the model decision. This indicates that Quick Shift considers

texture and geometric features in addition to intensity-based segmentation. As a result, each segmentation technique offers a distinct representation of the data and emphasizes specific factors that influence classification success. It can be concluded that the MRS and Mean Shift algorithms mostly depend on statistical summaries. On the other hand, SLIC and Quick Shift demonstrate that textural and geometric indications can be more efficient in classification decisions in this study. This indicates that each segmentation technique focuses on distinct dynamics in data representation and classification, offering significant insights into the explainability of outcomes.

5. Conclusion

To perform LULC mapping in the Marmara region of Turkey, this study considered the use of different segmentation techniques employing the SVM methods, a machine learning classifier. After identifying homogeneous zones using the MRS, SLIC, Mean Shift, and Quick Shift algorithms during the segmentation stage, the SVM classifier was used to complete the classification process. When comparing the MRS segmentation method with other segmentation algorithms, the results showed that the MRS algorithm produced superior performance. It was noted that MRS had a better prediction with approximately 5% to 8% improvement in terms of overall accuracy. According to the LIME results, the features associated with the NIR and SWIR bands mostly produce positive effects, whereas some features such as minimum and maximum values produce negative effects. A detailed evaluation of the results produced by each algorithm revealed that the SWIR and NIR bands contributed the most to the classification accuracy. Although elevated values of these features enhanced the likelihood of accurate classification, the maximum DN values on the red edge and blue bands diminish the probability of correct classification. In the SLIC algorithm, it was determined that the average DN values of the red edge, SWIR and NIR bands contributed positively to the classification. However, the maximum and minimum DN values in the red edge and NIR narrow bands negatively affected the performance. In the Mean Shift algorithm, it was observed that the average and maximum DN values of the blue band contributed positively to the classification decisions, while low minimum DN values caused the model to make predictions that were close to the negative class. The Quick Shift algorithm employs both texture-based features and statistical summaries to inform classification decisions. This result demonstrates that the algorithm assesses intensity and texture in conjunction with geometric features (area), which enhances classification accuracy. These findings demonstrate that the selection of the segmentation algorithm influences classification decisions based on specific data characteristics, which highlights the importance of considering these differences when determining the optimal classification strategy. These explainability methods facilitate a deeper understanding of the features that contribute to the enhanced performance of the classification models, thereby enhancing the transparency and interpretability of the model. In conclusion, the results of this research underscore the crucial role of employing advanced segmentation and classification techniques in remote sensing and image processing for the accurate mapping of LULC.

Author contributions

Conceptualization and methodology: T.K., E.O.Y.; formal analysis and investigation: E.O.Y., Y.K.U., E.B.; writing original draft preparation: T.K., E.O.Y., Y.K.U., E.B.; writing – review and editing: T.K., E.O.Y., Y.K.U., E.B.; supervision: T.K.

Data availability statement

The datasets used in this study are available from the corresponding author.

Acknowledgements

This research was not funded by external sources.

References

Achanta, R., Shaji, A., Smith, K. et al. (2010). SLIC Superpixels, Technical Report no. 149300, EPFL.

Alshari, E.A. and Gawali, B.W. (2021). Development of Classification System for LULC Using Remote Sensing and GIS. *Global Trans. Proc.*, 2(1), 8–17. DOI: 10.1016/j.gltp.2021.01.002.

Aryal, J., Sitaula, C., and Frery, A.C. (2023). Land Use and Land Cover (LULC) Performance Modeling Using Machine Learning Algorithms: A Case Study of the City of Melbourne, Australia. *Sci. Rep.,* 13(13510). DOI: 10.1038/s41598-023-40564-0.

Baatz, M., and Schäpe, A. (2000). Multiresolution Segmentation: An Optimization Approach for High Quality Multi-Scale Image Segmentation. In: J. Strobl, T. Blaschke and G. Griesebner (Eds.): Angewandte Geographische Informationsverarbeitung XII, pp. 12–23, Wichmann, Heidelberg, Germany.

Ben-Hur, A., Ong, C.S., Sonnenburg, S. et al. (2008). Support vector machines and kernels for computational biology. *PLoS Comput. Biol.*, 4(10), e1000173. DOI: 10.1371/journal.pcbi.1000173.

Cheng, Y. (1995). Mean Shift, Mode Seeking, and Clustering. *IEEE Trans. Pattern Anal. Mach.*, 17(8), 790–799. DOI: 10.1109/34.400568.

Comaniciu, D., and Meer, P. (2002). Mean shift: A robust approach toward feature space analysis. *IEEE Trans. Pattern Anal. Mach. Intell.*, 24(5), 603–619. DOI: 10.1109/34.1000236.

Cortes, C., and Vapnik, V. (1995). Support-Vector Networks. *Mach. Learn.,* 20. 273–297.

Drăgut, L., Csillik, O., Eisank, C. et al. (2014). Automated Parameterisation for Multi-Scale Image Segmentation on Multiple Layers. *ISPRS J. Photogramm. Remote Sens.,* 88, 119–127. DOI: 10.1016/j.isprsjprs.2013.11.018.

Evgeniou, T., Pontil, M., and Poggio, T. (2000). Regularization networks and support vector machines. *Adv. Comput. Math.*, *13*, 1-50. DOI: 10.1023/A:1018946025316.

Feng, C., Zhang, W., Deng, H. et al. (2023). A Combination of OBIA And Random Forest Based on Visible UAV Remote Sensing for Accurately Extracted Information About Weeds in Areas with Different Weed Densities in Farmland. *Remote Sens.,* 15(19), 4696. DOI: 10.3390/rs15194696.

Gani, M.A., Sajib, A.M., Siddik, M.A. et al. (2023). Assessing The Impact of Land Use and Land Cover on River Water Quality Using Water Quality Index and Remote Sensing Techniques. *Environ. Monit. Assess.*, 195(4), 449. DOI: 10.1007/s10661-023-10989-1.

Ghnemat, R., Alodibat, S., and Abu Al-Haija, Q. (2023). Explainable Artificial Intelligence (XAI) for deep learning based medical imaging classification. *J. Imaging*, 9(9), 177. DOI: 10.3390/jimaging9090177.

Humbal, A., Chaudhary, N., Pathak, B. (2023). Urbanization Trends, Climate Change, and Environmental Sustainability. In: Pathak, B., Dubey, R.S. (Eds.) Climate Change and Urban Environment Sustainability. Disaster Resilience and Green Growth. Springer, Singapore. DOI: 10.1007/978-981-19-7618-6_9.

Ishikawa, S.N., Todo, M., Taki, M. et. al. (2023). Example-Based Explainable AI and Its Application for Remote Sensing Image Classification. *Int. J. Appl. Earth Obs. Geoinf.*, 118, 103215. DOI: 10.1016/j.jag.2023.103215.

Kavzoglu, T., and Colkesen, I. (2009). A kernel Functions Analysis for Support Vector Machines for Land Cover Classification. *Int. J. Appl. Earth Obs. Geoinf.*, 11(5), 352–359. DOI: 10.1016/j.jag.2009.06.002.

Kavzoglu, T., Yildiz Erdemir, M. and Tonbul, H. (2016). A Region-Based Multi-Scale Approach for Object-Based Image Analysis. *ISPRS Archives*, 41(B7), 241–247. DOI: 10.5194/isprs-archives-XLI-B7-241-2016.

Kavzoglu, T., and Tonbul, H. (2017). A Comparative Study of Segmentation Quality for Multi-Resolution Segmentation and Watershed Transform, Istanbul, Turkey. 8th International Conference on Recent Advances in Space Technologies (RAST), 19–22 June (pp. 113–117), Istanbul, Turkey. DOI: 10.1109/RAST.2017.8002984.

Kavzoglu, T., Erdemir, M.Y., and Tonbul, H. (2017). Classification of Semiurban Landscapes from Very High-Resolution Satellite Images Using a Regionalized Multiscale Segmentation Approach. *J. Appl. Remote Sens.*, 11(3), 035016–035016. DOI: 10.1117/1.JRS.11.035016.

Kavzoglu, T., and Tonbul, H. (2018). An Experimental Comparison of Multi-Resolution Segmentation. SLIC And K-Means Clustering for Object-Based Classification of VHR Imagery. *Int. J. Remote Sens.*, 39(18), 6020–6036. DOI: 10.1080/01431161.2018.1506592.

Kavzoglu, T., Tonbul, H., Yildiz Erdemir, M. et al. (2018). Dimensionality Reduction and Classification of Hyperspectral Images Using Object-Based Image Analysis. *J. Indian Soc. Remote Sens.*, 46, 1297–1306. DOI: 10.1007/s12524-018-0803-1.

Kavzoglu, T., Tso, B., and Mather, P.M. (2024). Classification Methods for Remotely Sensed Data. Third Edition, Boca Raton: CRC Press. DOI: 10.1201/9781003439172.

Lourenço, P., Teodoro, A.C., Gonçalves, J.A. et al. (2021). Assessing The Performance of Different OBIA Software Approaches for Mapping Invasive Alien Plants Along Roads with Remote Sensing Data. *Int. J. Appl. Earth Obs. Geoinf.*, 95, 102263. DOI: 10.1016/j.jag.2020.102263.

Mohan, A., Singh, A.K., Kumar, B. et al. (2021). Review on Remote Sensing Methods for Landslide Detection Using Machine and Deep Learning. *T. Emerg. Telecommun.. T.*, 32(7), e3998. DOI: 10.1002/ett.3998.

Pal, N.R., and Pal, S.K. (1993). A Review on Image Segmentation Techniques. *Pattern Recognit.*, 26(9), 1277–1294.

Pipaud, I., and Lehmkuhl, F. (2017). Object-based Delineation and Classification of Alluvial Fans by Application of Mean-Shift Segmentation and Support Vector Machines. *Geomorphology*, 293, 178–200. DOI: 10.1016/j.geomorph.2017.05.013.

Quader, M.A. (2019). Impact Of Rohingya Settlement on The Landcovers at Ukhiya Upazila in Cox's Bazar, Bangladesh. *Jagannath University Journal of Life and Earth Sciences*, 5(1), 13–28.

Schölkopf, B. (2000). The kernel trick for distances. *Adv. Neural Inf. Process. Syst.*, *13*.

Steinhausen, M.J., Wagner, P.D., Narasimhan, B. et al. (2018). Combining Sentinel-1 and Sentinel-2 Data for Improved Land Use and Land Cover Mapping of Monsoon Regions. *Int. J. Appl. Earth Obs. Geoinf.*, 73, 595–604. DOI: 10.1016/j.jag.2018.08.011.

Su, T., and Zhang, S. (2017). Local and Global Evaluation for Remote Sensing Image Segmentation. *ISPRS J. Photogramm. Remote Sens.*, 130, 256–276. DOI: 10.1016/j.isprsjprs.2017.06.003.

Su, T., Li, H., Zhang, S. et al. (2015). Image Segmentation Using Mean Shift for Extracting Croplands from High-Resolution Remote Sensing Imagery. *Remote Sens. Lett.*, 6(12), 952–961. DOI: 10.1080/2150704X.2015.1093188.

Tassi, A., Daniela, G., Giuseppe, M. et al. (2021). Pixel- vs. Object-Based Landsat 8 Data Classification in Google Earth Engine Using Random Forest: The Case Study of Maiella National Park. *Remote Sens.*, 13(12), 2299. DOI: 10.3390/rs13122299.

Teke, A., and Kavzoglu, T. (2024). Exploring the Decision-Making Process of Ensemble Learning Algorithms in Landslide Susceptibility Mapping: Insights from Local and Global eXplainable AI analyses. *Adv. Space Res*., 74(8), 3765–3785. DOI: 10.1016/j.asr.2024.06.082.

Temenos, A., Temenos, N., Kaselimi, M. et al. (2023). Interpretable Deep Learning Framework for Land Use and Land Cover Classification in Remote Sensing Using SHAP. *IEEE Geosci. Remote S.*, 20, 8500105. DOI: 10.1109/LGRS.2023.3251652.

Tonbul, H., and Kavzoglu, T. (2020a). Semi-Automatic Building Extraction from Worldview-2 Imagery Using Taguchi Optimization. *Photogramm. Eng. Remote Sensing*, 86(9), 547–555. DOI: 10.14358/PERS.86.9.547.

Tonbul, H., and Kavzoglu, T. (2020b). A Spectral Band Based Comparison of Unsupervised Segmentation Evaluation Methods for Image Segmentation Parameter Optimization. *Int. J. Environ. Geoinformatics*, 7(2), 132–139. DOI: 10.30897/ijegeo.641216.

Tonbul, H., Colkesen, I., and Kavzoglu, T. (2020). Classification of poplar trees with object-based ensemble learning algorithms using Sentinel-2A imagery. *J. Geod. Sci.*, 10(1), 14–22. DOI: 10.1515/jogs-2020-0003.

Vedaldi, A., and Soatto, S. (2008). Quick Shift and Kernel Methods for Mode Seeking. In Computer Vision–ECCV 2008: 10th European Conference on Computer Vision, Marseille, France, 12-18 October 2008, Proceedings, Part IV 10 (pp. 705–718). DOI: 10.1007/978-3-540-88693-8_52.

Yilmaz, E.O., and Kavzoglu, T. (2024). Quality Assessment for Multi-Resolution Segmentation and Segment-Anything Model Using WORLDVIEW-3 Imagery. *ISPRS Archives*, 48, 383–390. DOI: 10.5194/isprs-archives-XLVIII-4-W9-2024-383-2024.

Zhu, L., Song, R., Sun, S. et al. (2022). Land Use/Land Cover Change and Its Impact on Ecosystem Carbon Storage in Coastal Areas of China from 1980 to 2050. *Ecol. Indic.*, 142, 109178. DOI: 10.1016/j.ecolind.2022.109178.