An In-depth Investigation of OBIA Classification with High-Resolution Imagery: Unravelling the Explanations Behind Deep Learning and Machine Learning

Elif Ozlem Yilmaz [1], Taskin Kavzoglu [1]

[1] Dept. of Geomatic Engineering, Gebze Technical University, 41400, Gebze/Kocaeli, Turkey – eoyilmaz@gtu.edu.tr; kavzoglu@gtu.edu.tr

Keywords: Object-Based Image Analysis, Convolutional Neural Network, XAI, XGBoost, Multiresolution Segmentation, SHAP.

Abstract

Object-Based Image Analysis (OBIA) is a method employed in the field of remote sensing with the objective of enhancing classification accuracy. This is achieved by focusing on image segments comprising groups of pixels, rather than evaluating individual pixels. By addressing the limitations of traditional pixel-based methods, OBIA is employed for the classification of segments based on their attributes. The present study evaluates the use of OBIA-based classification in conjunction with deep learning and machine learning classifiers. A study area, approximately 210 km² located in Ankara, was selected and SPOT-6 imagery with a spatial resolution of 1.5 meters and 4 spectral bands (red, green, blue and near infrared) was employed for this purpose. In the segmentation stage, a multiresolution segmentation approach was employed, and classification process was conducted using a Convolutional Neural Network (CNN) and Extreme Gradient Boosting (XGBoost). The CNN classifier demonstrated superior performance compared to the XGBoost algorithm, with an improvement of 2.7%. The Shapley Additive Explanations (SHAP) technique, an effective Explainable Artificial Intelligence (XAI) method, was employed to assess the explainability of the classifiers. The SHAP analysis indicated that the HSI transform was the most influential factor in the XGBoost algorithm's decision-making process whereas the average DN values of the green band were the most effective feature for the CNN model. Global SHAP analyses elucidated the overarching model decision-making process, whereas class-specific analyses furnished insights into the classification of each land use and land cover (LULC) class.

1. Introduction

The use of remote sensing technology for the purpose of obtaining land use and land cover (LULC) information is widespread due to the technology's ability to collect data on a periodic basis over vast geographical areas (Steinhausen et al., 2018; Singh et al., 2017). The dynamic nature of LULC necessitates continuous monitoring and accurate predictions (Alshari and Gawali, 2021). However, traditional pixel-based methods are inadequate for high-resolution imagery as they are unable to account for the spatial context of neighboring pixels. Object-based image analysis (OBIA) methods are a more effective approach, particularly for high-resolution satellite imagery. OBIA employs groups of pixels as the unit of analysis, rather than individual pixels, integrating spatial, textural and contextual features to provide more accurate LULC classifications (Kavzoglu and Tonbul, 2017; Panda et al., 2024). The success of this method is contingent upon the quality of the image segmentation process and the correct choice of parameters. The principal advantage of OBIA is that it markedly enhances mapping accuracy by combining spectral data with textural and contextual information (Kavzoglu et al., 2024).

Machine learning techniques, particularly Extreme Gradient Boosting (XGBoost), are extensively employed in the domain of computer vision and data science research (Chen and Guestrin, 2016). XGBoost is an efficient and powerful machine learning method based on the Gradient Boosting algorithm, which is characterized by high performance in both classification problems (Kavzoglu and Teke, 2022). In recent years, methods based on deep learning, particularly convolutional neural networks (CNNs), have been increasingly employed in the field of OBIA. They offer substantial benefits in the analysis of intricate and high-resolution images, largely due to their automated feature extraction capabilities (Kavzoglu and Yilmaz, 2022). The integration of CNNs with OBIA can demonstrate superior performance in LULC classifications, with high accuracy rates. These models are capable of effectively processing segmented image objects and accurately classifying different types of LULC. In addition, advances in image processing, in the field of Explainable Artificial Intelligence (XAI), have led to the development of methods such as SHapley Additive Explanations (SHAP), which provide invaluable insights into the decision-making processes of machine and deep learning models (Salih et al., 2024).

SHAP analysis employs Shapley scores to elucidate the way individual characteristics contribute to the generation of classification outcomes (Lundberg et al., 2020). This method enhances the transparency of the decision-making processes within the models, thereby improving the interpretability of the results and strengthening the reliability of these models. XAI techniques, particularly SHAP, have been demonstrated to markedly enhance the reliability, accountability and accuracy of classification results obtained through the analysis of remote sensing data. Consequently, the integrated utilization of innovative techniques, such as XGBoost, CNN and XAI, optimizes the efficacy of classification and analysis applications within the domain of remote sensing.

The objective of this study is to evaluate the effectiveness of OBIA in the interpretation of a very high-resolution image, while simultaneously investigating the performance of machine and deep learning algorithms, including CNN and XGBoost, in the classification process. The model is employed to examine the potential for high accuracy in deep learning approaches, with XGBoost regarded as a benchmark representing robust machine learning techniques. The primary objective of the study is to assess the capacity of each classifier to learn and accurately classify the spatial and spectral features of the image segments. Moreover, XAI methodologies, particularly SHapley Additive explanations (SHAP) analysis, are integrated to enhance the reliability and transparency of the classification outcomes. By

https://doi.org/10.5194/isprs-archives-XLVIII-M-6-2025-317-2025

offering a comprehensive assessment of the decision-making processes of each classifier, both in aggregate and at the class level, SHAPE can explain the influence of individual features on the classification results.

2. Study Area and Dataset

The study area comprised approximately 210 km^2 of Ankara, the capital of Turkey (Figure 1). The area encompasses a multitude of LULC classes, including urban, agricultural, and natural areas. To meet the objectives of the study, SPOT-6 optical imagery, provided free of charge by Airbus, with a high spatial resolution, was employed. The image comprises a panchromatic band with a spatial resolution of 1.5 m and four spectral bands (red, green, blue, and near infrared) with a spatial resolution of 6 m. The panchromatic band enables the spatial resolution of the four spectral bands with a spatial resolution of 6 m to be reduced to 1.5 m. All produced bands are combined with each other. The study area contains 6 LULC classes: shadow, bare ground, road, vegetation, white roof and red roof. Samples of these classes were also collected as 2000 sample points for each class prior to image processing.

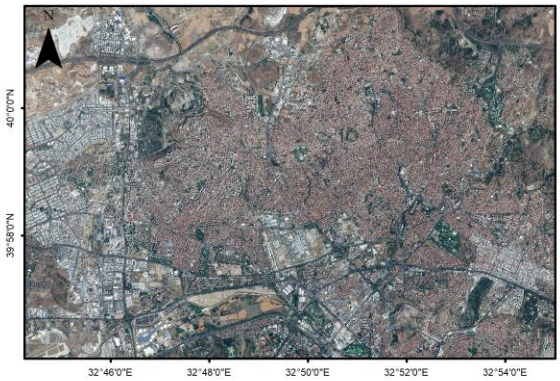

Figure 1. The study area situated in the Turkey capital, Ankara.

3. Methodology

In this study, LULC maps were created using OBIA with very high-resolution satellite imagery. OBIA consists of two main stages. In the first stage, the segmentation stage, meaningful image segments are created by grouping image pixels. In the second stage, these segments are assigned to one of the LULC classes by a classifier. In the segmentation stage, a resolution segmentation (MRS) approach was used, while machine and deep learning classifiers were performed in the classification process. Furthermore, the performance of the classifiers was evaluated by calculating recall, precision, F-score, overall accuracy and Kappa coefficient. Subsequently, SHAP was employed to enhance comprehension of the classification outcomes and to clarify the decision-making process of the model. This approach was utilized to ascertain the attributes on which artificial intelligence models base their results and to assess the influence of segment characteristics on classification.

3.1 Segmentation Algorithms

In the segmentation phase, multiresolution segmentation (MRS) algorithms were employed. The algorithm performs an analysis at varying resolution levels while generating image segments. At each resolution level, comparisons are performed between pixels regarding color, brightness and texture. Similar regions are combined, while regions with different properties are separated and converted into segments (Baatz and Schäpe, 2000). Additionally, there are three fundamental parameters in MRS, namely scale, shape and compactness. The determination of these parameters directly affects the segmentation process and enhances the quality of the segmentation (Kavzoglu et al., 2016). Scale is of paramount importance (Yilmaz and Kavzoglu, 2024). In accordance with the objective of the study, the efficacy of the algorithm is enhanced by selecting the optimal scale, thereby yielding the most favorable outcome (Kavzoglu and Tonbul, 2017).

3.2 Extreme Gradient Boosting

Through the combination of multiple machine learning models, models can have high performance and increase their generalization capacity. Several machine learning approaches are described in the literature that are based on this principle. One such approach is the Extreme Gradient Boosting (XGBoost) algorithm, developed by Chen and Guestrin (2016). This algorithm can produce rapid and highly accurate results through the utilization of many sequential decision trees, which serve to minimize errors. In other words, the outcomes of the jointly employed decision trees exert an influence on one another, thereby facilitating parallel computation. Furthermore, this algorithm is a highly flexible and versatile tool that can be deployed in both regression-based and classification-based applications (Kavzoglu and Teke, 2022).

3.3 Convolutional Neural Network

Convolutional neural networks (CNNs) are a widely utilized tool with the capacity to effectively analyze both spatial and spectral features from images due to the unique structures they contain. The convolutional layers enable the network to identify the distinctive characteristics of the data set without the need for human intervention (Kavzoğlu and Yılmaz, 2022). The general structure of CNNs is composed of three principal layers: the input layer, the hidden layer and the output layer. In a one-dimensional CNN, the input layer accepts a vector matrix containing a feature value corresponding to each data input. The convolution layers consist of filters that have been optimized by a back-propagation algorithm, which scans the dataset to generate different feature maps. This process identifies the most significant aspects of the data while simultaneously reducing its size, the number of parameters, and the computational workload. Consequently, the potential for overlearning is mitigated. The activation function such as ReLU, employed after the convolution layers, endows the model with a non-linear structure and accelerates the computational process. Optimization techniques utilized during model training aim to minimize the loss function and dynamically adjust the learning rate. These methodologies facilitate expedient and effective performance in gradient-based optimization processes (Song et al., 2019).

3.4 Explainable Artificial Intelligence

Explainable Artificial Intelligence (XAI) is a technique that aims to provide insight into the decision-making processes of machine learning and deep learning models. By unveiling the intrinsic characteristics and variables that inform the predictions and decisions of these models, XAI empowers users to comprehend the underlying logic and rationale behind the models' conclusions. These models, often characterized as 'black boxes', are challenging to interpret due to their intricate parameterization and sophisticated computational processes.

XAI methodologies elucidate internal structure and decision-making mechanisms of these models, enhancing their transparency and reliability (Kavzoglu et al., 2024).

SHapley Additive exPlanations (SHAP) analysis is one of the XAI methods and is based on game theory. It is used to interpret the predictions of artificial intelligence models (Kavzoglu and Teke, 2022). This method employs Shapley values to ensure a fair distribution of the impact of each feature. While positive and negative values reflect the contribution of each feature to a particular class, a high average absolute Shapley value over all data points indicates that a feature is generally more important. The reliability of SHAP results is superior to that of standard measurements, as it also considers the interactions between features. This method offers a comprehensive view of the overall performance of the algorithms, while simultaneously elucidating the contribution of each feature to the classification result in an objective and intelligible manner through example-based analyses. SHAP is model-agnostic and can be used to explain the outputs of both deep learning and machine learning algorithms. Furthermore, it provides various forms of graphical representation to visualize the decision-making processes of classification algorithms, thereby enhancing the reliability and interpretability of these algorithms (Chen et al., 2021).

4. Results and Discussion

4.1. Design of the Classifiers and Training Process

In this study, the MRS method was employed for the purpose of segmentation. In the segmentation process, the scale parameter was set to 30, the shape parameter to 0.1, and the compactness parameter to 0.5, based on a process of trial and error. The segmentation was conducted on a four-band image, and the OBIA features were calculated for each segment. The features encompass brightness, maximum difference, mean, standard deviation, skewness, ratio, minimum and maximum pixel values, color space characteristics (hue, saturation, intensity), geometric parameters (area, boundary length, length, pixel count, asymmetry, boundary index, compactness) and GLCM-based texture features (homogeneity, contrast, similarity, entropy, mean and standard deviation). Following the extraction of the features of the segments, the pixels collected for training purposes were assigned to the nearest segment. Subsequently, the data set was employed. The data set was randomly divided into 60% training, 20% validation and 20% testing, to be utilized in the training, validation and testing phases of the model. The training dataset was employed to ascertain the parameters of the model, whereas the validation dataset was utilized to optimize the hyperparameters. The test dataset was used to evaluate the overall performance of the model.

The OBIA features were then classified by a one-dimensional CNN model. The model structure is as follows: three convolutional 1D layers are combined with max pooling 1D operations to enable the learning of deep spatial features. To enhance the accuracy of the model, fully connected layers (dense) are employed. Additionally, dropout layers are incorporated to prevent overfitting. In the final layer of the model, multiclass classification is conducted with a Softmax activation function. Moreover, the model was trained using the Adam optimization algorithm and a categorical cross-entropy loss function, with a total of 200 epochs completed. A batch size of 32 was employed during the training of the model. The results indicated that the model exhibited a loss of 0.075 and an accuracy of 97.8% on the training dataset, while the loss on the test dataset was 0.139 and the accuracy was 96.9%. On the other hand, The XGBoost algorithm was employed for multiclass classification purposes. The model was structured as 'multi:softprob', with the max_depth parameter set to 200 and the number of trees (n_estimators) set to 5. The hyperparameters of the model were determined through testing to achieve optimal performance. The validation accuracy of the XGBoost classifier was 95.0%.

4.2. Classification Results

In this study, the classification of LULC classes was evaluated using both the XGBoost and CNN models. The performance of the models was evaluated using a range of metrics, including F1-score, recall, precision, overall accuracy and Kappa coefficient (Table 1). The results demonstrated that the CNN model exhibited superior performance compared to the XGBoost model, achieving higher accuracy rates across all classes. For the Shadow class, the CNN model demonstrated superior performance compared to the XGBoost model, achieving an F1-score of 96.34% and a recall of 96.94%. Similarly, for the Bare Soil class, the CNN model exhibited a higher F1-score (97.03%) than the XGBoost model (92.71%). For the Road class, the F1-score of the CNN model (92.93%) is higher than that of XGBoost (91.26%). Additionally, for the Vegetation class, the CNN model achieved an F1-score of 97.89%, surpassing XGBoost's 97.64% by a slight margin. For the White Roof class, the CNN model demonstrates a notable advantage, with an F1-score of 97.51% and a precision of 99.85%. Finally, for the Red Roof, the CNN model achieved an F1-score of 98.53%, which was significantly higher than that of the XGBoost model (94.15%). In terms of overall model performance, the overall accuracy of the CNN model (96.71%) is higher than that of the XGBoost model (94.75%). Moreover, the Kappa coefficient is 0.960 for the CNN model and 0.936 for the XGBoost model. These findings demonstrate that the CNN model exhibits greater consistency and success in LULC classification.

LULC Classes	XGBoost			CNN		
	F1-score (%)	Recall (%)	Precision (%)	F1-score (%)	Recall (%)	Precision (%)
Shadow	96.34	96.94	95.74	96.42	99.30	93.70
Bare Soil	92.71	95.41	90.15	97.03	97.63	96.43
Road	91.26	90.69	91.84	92.93	93.19	92.68
Vegetation	97.64	97.50	97.77	97.89	97.08	98.72
White Roof	96.38	96.25	96.51	97.51	95.27	99.85
Red Roof	94.15	91.66	96.77	98.53	97.77	99.29
Overall Acc. (%)	94.745			96.712		
Kappa Coef.	0.936			0.960		

Table 1. Assessment of the accuracy of the thematic maps produced with XGBoost and CNN

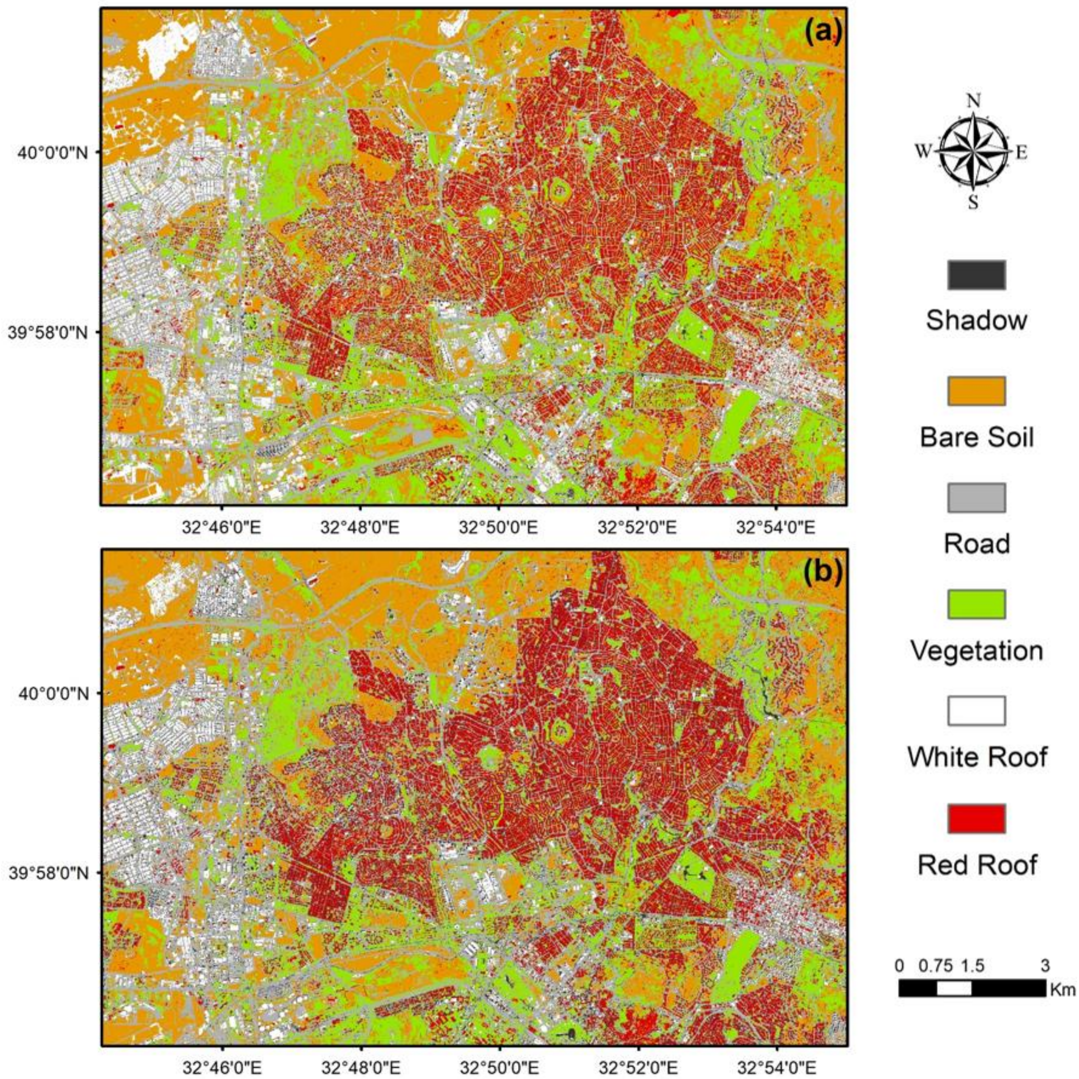

Figure 2. Thematic maps are produced using classifiers of (a) XGBoost and (b) CNN.

4.3. XAI Results

To gain insight into the decision-making processes of machine learning and deep learning algorithms, SHAP analysis, a technique for describing the output of machine learning models, was employed. In the case of the XGBoost model, a class-based analysis was initially conducted using SHAP analysis (Figure 3). For each class (Shade, Bare Soil, Road, Vegetation, White Roof, and Red Roof), the contribution of the features that affect the model's decision-making process is illustrated. A positive SHAP value indicates that the feature is the primary determinant of the model's decision in favor of that class. Conversely, a negative value suggests that the feature is a contributing factor to the model's bias towards other classes. In the Shadow class, spectral features such as "Min_pixel_Red" and "Mean_Red" were of particular significance, exerting a positive influence on the model predictions and serving as pivotal determinants in the differentiation between classes. In the Bare Soil class, the "HSL_Transf" and "Ratio_Green" features exhibited high SHAP values and were the most influential in the model's classification performance. Furthermore, statistical features, such as "Standard_d" and "Mean_Red" also exert a considerable influence. For the Road class, the "Ratio_Green" and "Ratio_Blue" features play a pivotal role in class prediction, exhibiting strong positive SHAP values. Furthermore, the application of geometric and spectral features, such as "Mean_Red" and "Length_Pxl," has also proven to be effective. In other words, road segments exhibit distinctive characteristics when compared to other LULC classes. One such difference is their elongated, narrow structure, which sets them apart from other LULC classes in terms of length. For the Vegetation class, the "Ratio_NIR" and "Mean_Blue" features were identified as the most effective variables. This highlights the particular importance of reflectance characteristics in the NIR band for classification purposes. For the White Roof class, spectral features such as "Mean_Red" and "Ratio_Blue" were identified as having a significant impact on the prediction performance. In the Red Roof class, the variables "HSI_Transf" and "Min_pixel_Green" were found to have a significant impact on the classification process, as indicated by their high SHAP values. This suggests that certain color transformations are crucial for differentiating between classes in the Red Roof class. This analysis provides insights into the features that the model relies on for each class and the extent to which these features influence the decision-making process.

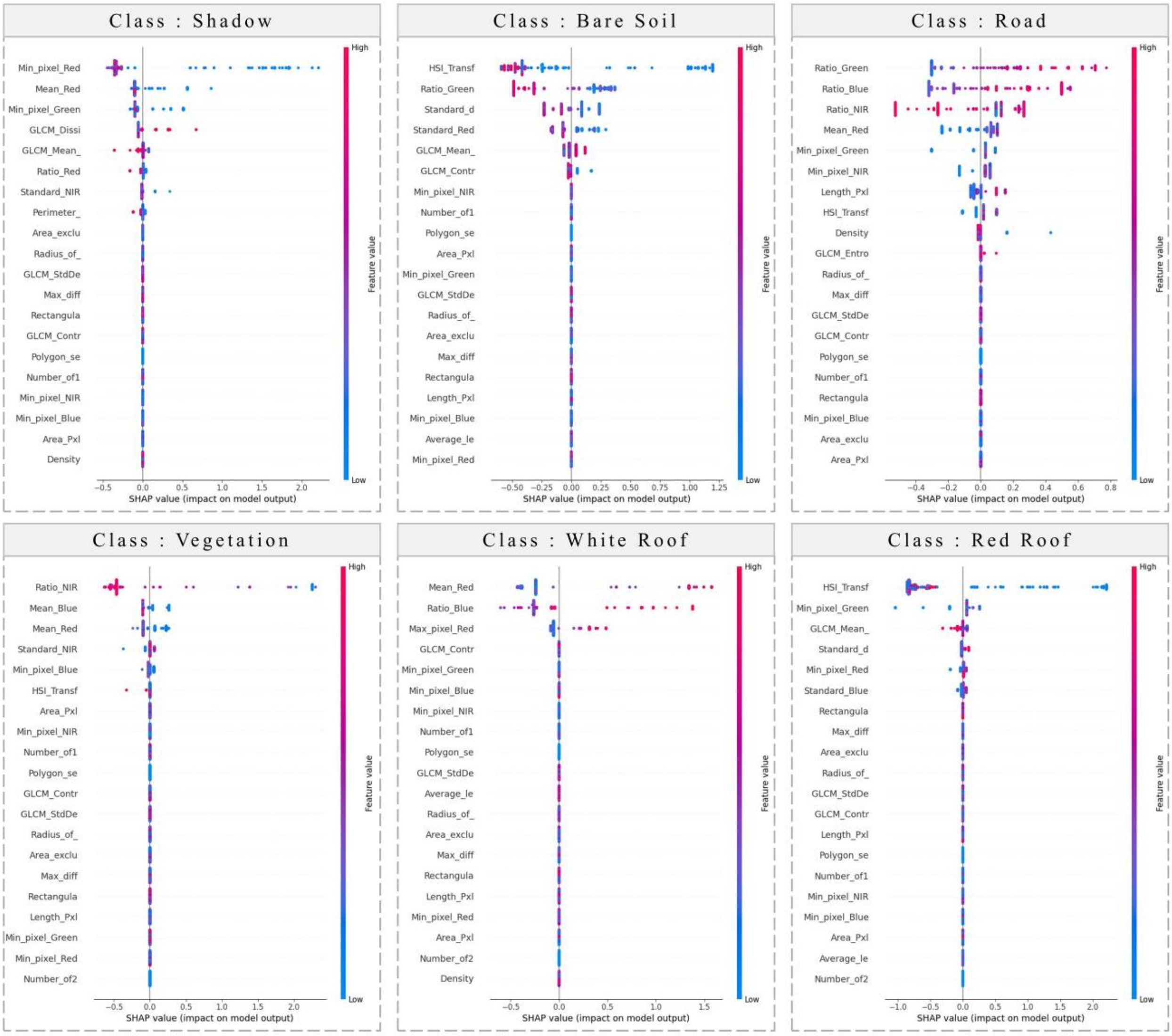

Figure 3. LULC class based SHAP graphs for XGBoost classifier.

A class-based SHAP analysis was conducted on the XGBoost model and subsequently applied to the CNN model (Figure 4). The results of the analysis demonstrate the influence of features on model predictions and their relative importance for each class. In the case of the Shadow class, the spectral features "Mean_Green" and "Min_pixel_Red" emerge as the variables with the most significant impact on the model's predictive performance. This demonstrates that the shadow classes can be differentiated based on the average reflectance values in the green and red bands. For the Bare Soil class, the "Mean_Red" and "Mean_Green" features contribute the most to the prediction accuracy, exhibiting strong SHAP values. In particular, the reflectance values of the soil surfaces in the red spectrum are of significant consequence regarding the distinction of classes. For the Road class, the mean red and standard near-infrared features have a positive effect on the model prediction. These findings highlight the significance of the mean values of the road class in spectral bands as discriminating factors. For the Vegetation class, the variables "Min_Pixel_Green" and "HSI_Transf" are of particular significance, exhibiting the highest SHAP values. For the White Roof class, the "Mean_Green" and "Standard_NIR" features contribute the most to the model prediction. These findings indicate that spectral reflectance differences are a significant factor in class distinction within the White Roof class. For the Red Roof class, the dominant variable was identified as "HSI_Transf." In general, spectral bands (especially mean and standard values) were found to be prominent in the classification using the CNN model. These results demonstrate that the learning mechanism of the model is predominantly informed by spectral features and that the decision-making process of the model can be elucidated through SHAP analysis. The study makes significant contributions to the accuracy and reliability analysis of deep learning models in the context of XAI.

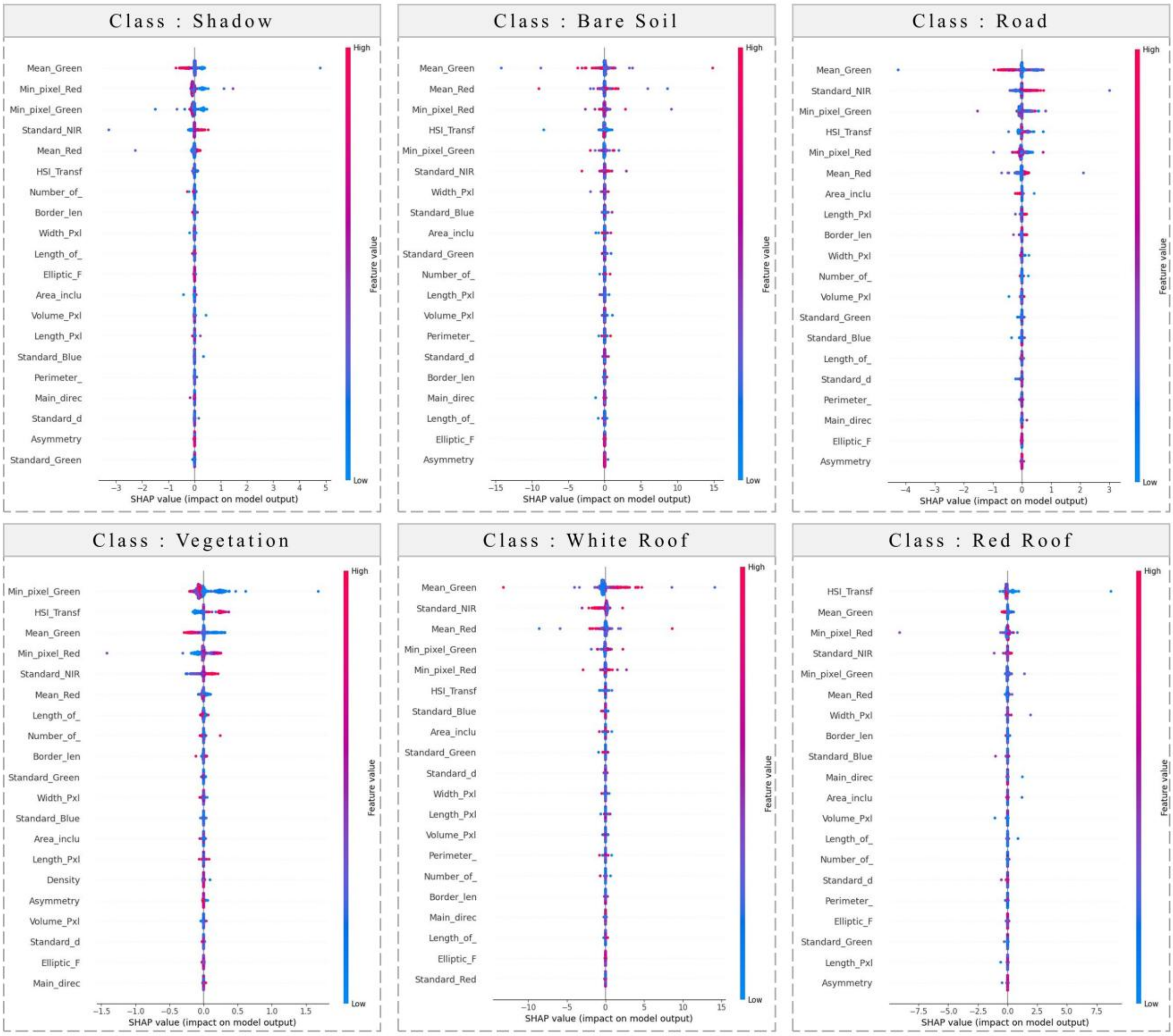

Figure 4. LULC class based SHAP graphs for CNN classifier.

The global SHAP analysis of the XGBoost and CNN models provides a clear indication of the relative effectiveness of features in the decision-making processes of both models, as demonstrated by the contribution of these features (Figure 5). In the XGBoost model, spectral features were of primary importance, with variables such as “HSI_Transf”, “Ratio_NIR“, “Mean_Red” and “Ratio_Blue” contributing the most to the classification performance. Furthermore, it is evident that spectral features are considered in this model, facilitating more precise differentiation between classes. In other words, it can be said that XGBoost is capable of effectively modelling spectral variations, with these variations contributing significantly to the decision processes through SHAP values. In the CNN model, the decision-making process was significantly influenced by spectral features. Features such as “Mean_Green”, “Standard_NIR’ and “Min_pixel_Green” were found to be the most effective in the classification decisions of the model. Additionally, variables such as “HSI_Transf” and “Mean_Red” also made significant contributions to the performance of the model. The CNN model demonstrated a greater focus on spectral information than on textural and geometric features, a pattern that is clear in the density distribution of SHAP values.

5. Conclusion

The objective of this study was to evaluate the performance of XGBoost and CNN models in classifying LULC using features obtained through the OBIA approach. The findings of the study demonstrated that both models exhibited high accuracy rates; however, the CNN model demonstrated superior performance compared to the XGBoost model. In particular, the CNN model demonstrated superior performance in terms of F1-score, recall, and precision across all classes, with an overall accuracy of 96.71%. In contrast, the XGBoost model achieved an accuracy rate of 94.75%. With regard to the Kappa coefficient, the CNN model demonstrated greater consistency (0.960) than the XGBoost model (0.936). Furthermore, the decision-making processes of the models were elucidated through the utilization of SHAP analysis, a prominent XAI method. The SHAP analysis demonstrated that both models place significant reliance on spectral features. The XGBoost model employed spectral features, including “HSI_Transf” “Mean_Red,” and “Ratio_NIR” whereas the CNN model emphasized features such as “Mean_Green” “Standard_NIR”, and “Min_pixel_Green.” These analyses have made a substantial contribution to the elucidation of the decision-making processes and the enhancement of the reliability of the models. In conclusion, the findings of this study demonstrate that CNN models exhibit superior performance in LULC classifications, due to their capacity to learn deep spectral features. Concurrently, the application of SHAP analysis facilitated the explication of the classification processes, enabling a comprehensive interpretation of the model outputs and enhancing the reliability of the models. In light of the findings presented in this study, it is recommended that future research consider combining XGBoost and CNN models in a hybrid structure. This approach has the potential to enhance classification performance by leveraging the strengths of both models. Furthermore, the application of XAI techniques could facilitate a more comprehensive understanding of the decision-making processes involved in hybrid models, thereby improving the explainability of this hybrid approach. This, in turn, could lead to more reliable LULC classification results.

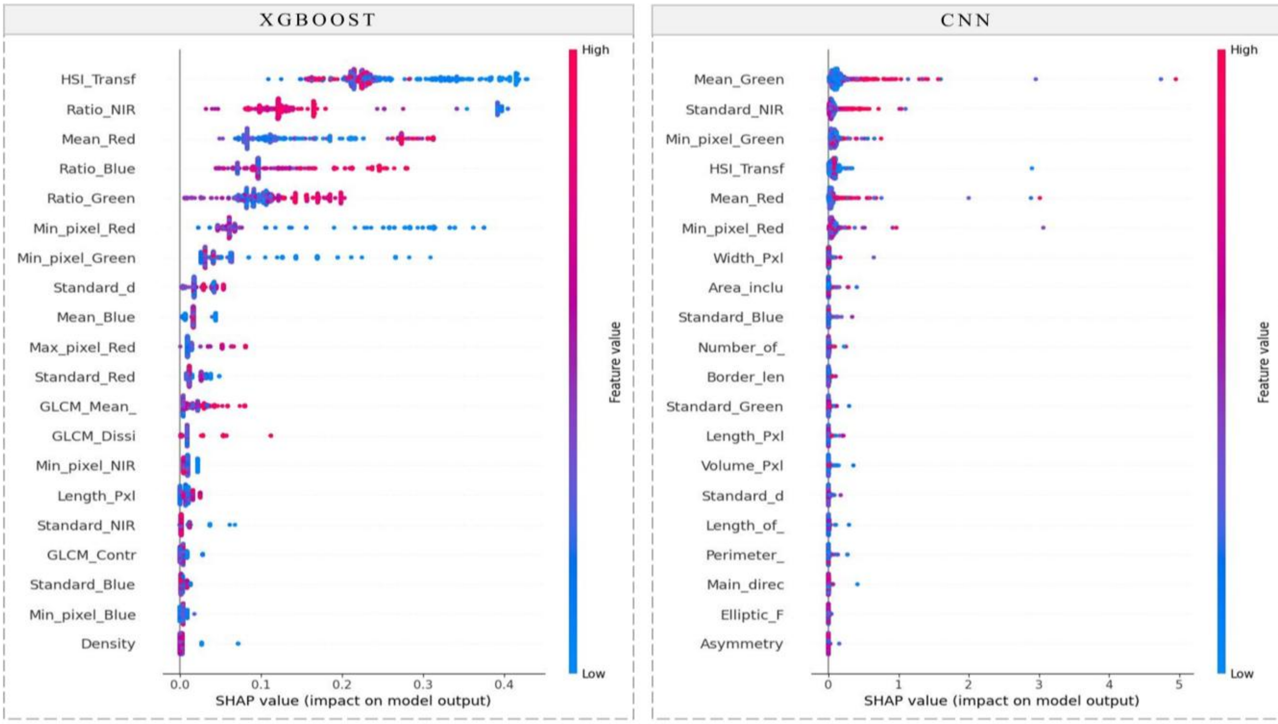

Figure 5. Global SHAP graphs for (a) XGBoost and (b) CNN.

Acknowledgements

The authors would like to express their sincerest gratitude to AIRBUS for providing the SPOT 6 imagery utilized in this study (https://space-solutions.airbus.com/imagery/sample-imagery/).

References

Alshari, E.A., Gawali, B.W., 2021. Development of classification system for LULC using remote sensing and GIS. *Glob. Transit. Proc.*, 2(1), 8–17. doi.org/10.1016/j.gltp.2021.01.002.

Baatz, M., Schäpe, A., 2000. Multiresolution segmentation: an optimization approach for high quality multi-scale image segmentation. In: Strobl, J., Blaschke, T., Griesebner, G. (Eds.), *Angewandte Geographische Informationsverarbeitung XII*, Wichmann, Heidelberg, Germany, 12–23.

Chen, T., Guestrin, C., 2016. Xgboost: A scalable tree boosting system. *Proc. 22nd ACM SIGKDD Int. Conf. Knowl. Discov. Data Min.*, 785–794.

Kavzoglu, T., Tonbul, H., 2017. A comparative study of segmentation quality for multi-resolution segmentation and watershed transform. *Proceedings of the 8th Int. Conf. Recent Adv. Space Technol. (RAST)*, Istanbul, Turkey, 19–22 June, 113–117. doi.org/10.1109/RAST.2017.8002984.

Kavzoğlu, T., Yilmaz, E.Ö., 2022. Analysis of patch and sample size effects for 2D-3D CNN models using multiplatform dataset: hyperspectral image classification of ROSIS and Jilin-1 GP01 imagery. *Turk. J. Electr. Eng. Comput. Sci.*, 30(6), 2124–2144. doi.org/10.55730/1300-0632.3929.

Kavzoglu, T., Teke, A., 2022. Predictive performances of ensemble machine learning algorithms in landslide susceptibility mapping using random forest, extreme gradient boosting (XGBoost) and natural gradient boosting (NGBoost). *Arab. J. Sci. Eng.*, 47(6), 7367–7385.

Kavzoglu, T., Tso, B., Mather, P.M., 2024. *Classification Methods for Remotely Sensed Data*. Third Edition, CRC Press, Boca Raton. doi.org/10.1201/9781003439172.

Kavzoglu, T., Yildiz Erdemir, M., Tonbul, H., 2016. A region-based multi-scale approach for object-based image analysis. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.*, XLI-B7, 241–247. doi.org/10.5194/isprs-archives-XLI-B7-241-2016.

Lundberg, S.M., Erion, G., Chen, H., DeGrave, A., Prutkin, J.M., Nair, B., Katz, R., Himmelfarb, J., Bansal, N., Lee, S.I., 2020. From local explanations to global understanding with explainable AI for trees. *Nat. Mach. Intell.*, 2(1), 56–67. doi.org/10.1038/s42256-019-0138-9.

Panda, K.C., Singh, R.M., Singh, S.K., 2024. Advanced CMD predictor screening approach coupled with cellular automata-artificial neural network algorithm for efficient land use-land cover change prediction. *J. Clean. Prod*., 449, 141822. doi.org/10.1016/j.jclepro.2024.141822.

Salih, A.M., Raisi-Estabragh, Z., Galazzo, I.B., Radeva, P., Petersen, S.E., Lekadir, K., Menegaz, G., 2024. A perspective on explainable artificial intelligence methods: SHAP and LIME. *Adv. Intell. Syst.*, 2400304.

Singh, S.K., Srivastava, P.K., Szabó, S., et al., 2017. Landscape transform and spatial metrics for mapping spatiotemporal land cover dynamics using earth observation datasets. *Geocarto Int.,* 32(2), 113–127. doi.org/10.1080/10106049.2015.1130084.

Song, J., Gao, S., Zhu, Y., et al., 2019. A survey of remote sensing image classification based on CNNs. *Big Earth Data,* 3(3), 232–254. doi.org/10.1080/20964471.2019.1657720.

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.org/10.1016/j.jag.2018.08.011.

Yilmaz, E.O., Kavzoglu, T., 2024. Quality Assessment for Multi-Resolution Segmentation and Segment-Anything Model Using WORLDVIEW-3 Imagery. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci.,* XLVIII-4/W9, 383-390. doi.org/10.5194/isprs-archives-XLVIII-4-W9-2024-383-2024.