6th Intercontinental Geoinformation Days

igd.mersin.edu.tr

An evaluation of the effectiveness of spectral bands and indices on semantic segmentation with Attention U-Net using Sentinel-2A imagery and ESA WorldCover products

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

[1]Gebze Technical University, Faculty of Engineering, Department of Geomatics Engineering, Kocaeli, Türkiye

Keywords
Attention U-Net
Deep Learning
Semantic Segmentation
Sentinel-2A
Spectral Index

Abstract

In remote sensing applications, semantic segmentation is applied to assign a semantic label (e.g., settlement, forest, and meadow) to the pixels of an image. Due to the diversity of natural and unnatural landscapes, semantic segmentation remains a difficult task for remotely sensed imagery. In this study, the Attention U-Net model was trained as a deep learning network for semantic segmentation of a Sentinel-2A image together with the WorldCover reference dataset to generate the land cover map of the study site covering Bolu, Duzce and Zonguldak provinces of Turkey. The Attention U-Net was applied to two datasets (original Sentinel-2A dataset with 10 spectral bands and Sentinel-2A dataset with additional spectral indices including Normalized Difference Vegetation Index-NDVI, Soil Adjusted Vegetation Index-SAVI, Normalized Difference Build-Up Index-NDBI and Normalized Difference Water Index-NDWI) and performance comparison was performed. The performances of deep learning models were investigated using Intersection Over Union (IoU) for segmentation model training and mean IoU for model results. The results show that the use of spectral indices as auxiliary data in semantic segmentation increases the mean IoU by approximately 10%. The results indicate that spectral indices have a higher degree of effectiveness in accuracy assessment than the original spectral bands.

1. Introduction

The capability of remote sensing (RS) platforms to gather diverse data has significantly increased with the recent technological developments in sensors. Remotely sensed data that are currently accessible have emerged as crucial components for environmental surveillance as they facilitate the identification of developments, alterations, and modifications on the Earth's surface (Shafique et al., 2022). They have been widely utilized for land cover and land use mapping, urban planning, change detection, object detection, and disaster management (Kavzoglu and Colkesen, 2013; Kavzoglu and Mather, 2003). The increasing accessibility of remotely sensed imagery, as a result of the accelerated development of remote sensing technology, provides broader opportunities for image databases (van der Meer, 2012). There exists great potential for data-driven applications, given the coverage capacity, high spatial resolution, high frequency of acquisition, and rich spectral resolution of remotely sensed imagery. In addition, numerous artificial intelligence (i.e., machine learning and deep learning) based techniques have been presented to analyze and process these data (Tsagkatakis et al., 2019). The deep learning approach has important advantages including the minimum level of human intervention, the ability to solve complex problems, and effectiveness in processing high-resolution satellite images (Yılmaz and Kavzoğlu, 2021). To date, various deep learning models have been suggested for applications (Kavzoglu et al., 2021). In particular, semantic segmentation of satellite imagery based on deep learning has been one of the substantial research topics providing essential data for geoscience applications. Semantic segmentation is the technique of assigning a class categorization to each pixel of an image (Kushwah and Markam, 2021). Unlike image classification, it is a high-level process that facilitates scene comprehension. The segmentation of image objects has a substantial impact on the precision of the classification process because segmentation is a processing stage before land use/land cover (LULC) map classification. In the last two decades, image segmentation has been conducted through region-based and edge-based algorithms. One of the most popular

* **Corresponding Author**

*(eoyilmaz@gtu.edu.tr) ORCID ID 0000-0002-6853-2148
(kavzoglu@gtu.edu.tr) ORCID ID 0000-0002-9779-3443

Cite this study

Yilmaz, E. O., & Kavzoglu, T. (2023). Evaluating the impact of spectral indices on land cover semantic segmentation with Attention U-Net architecture utilizing Sentinel-2A WorldCover reference dataset. Intercontinental Geoinformation Days (IGD), 6, 360-363, Baku, Azerbaijan

methods has been multi-resolution segmentation having scale, shape and compactness parameters (Kavzoglu et al., 2016; Tonbul et al., 2020; Yildiz et al., 2012). Evaluation of segmentation is critical to evaluate the segmentation quality qualitatively and quantitatively (Kavzoglu and Tonbul, 2017; Kavzoglu and Tonbul, 2018; Tonbul and Kavzoglu, 2017; Maxwell et al., 2021).

Existing research has highlighted problems related to the low spatial resolution of satellite imagery, the inability to fully exploit spectral information, the use of traditional methods as classifiers, and the difficulty in selecting the ideal patch size for image classification (Kavzoglu and Yilmaz, 2022). In addition, semantic segmentation applications using deep learning methods suffer from missing or insufficiently accurate reference data. To address these issues, a semantic segmentation model based on deep learning (i.e., the Attention U-Net architecture) was applied to WorldCover reference maps provided by ESA, and the influence of spectral indices, namely Normalized Difference Vegetation Index-NDVI, Soil Corrected Vegetation Index-SAVI, Normalized Difference Buildup Index-NDBI and Normalized Difference Water Index-NDWI was also investigated.

2. Study Area and Dataset

The western side (Duzce, Bolu and Zonguldak provinces) of the Black Sea region was selected as the study area, covering approximately 100 km^2 (Figure 1). In addition, the study area has a wide coastline and is known to contain mainly forest and grassland areas.

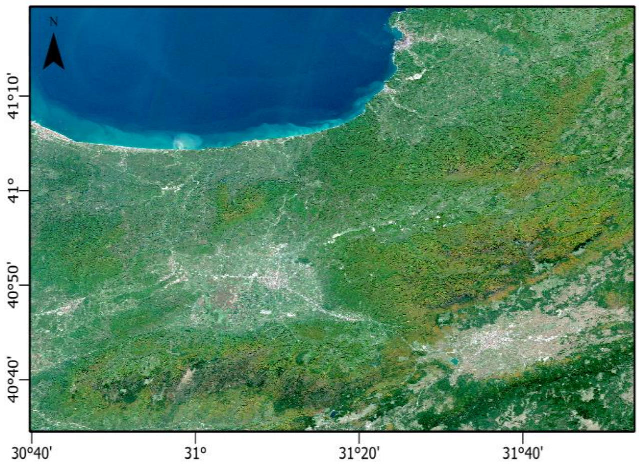

Figure 1. The location of the study area

The high-resolution, cloud-free Sentinel-2A satellite image acquired in October 2021 and the Earth Cover reference datasets produced and provided by (European Space Agency) ESA for 2021 were used for semantic segmentation and LULC labels, respectively. Prior to the training phase, the dataset underwent a process in which three bands of the Sentinel-2 image with a spatial resolution of 60 m were excluded. Furthermore, the Normalized Difference Vegetation Index (NDVI), the Soil Adjusted Vegetation Index (SAVI), the Normalized Difference Build-Up Index (NDBI) and the Normalized Difference Water Index (NDWI) were created and then used to investigate the contribution of the spectral indices to the semantic segmentation based on U-Net (Table 1).

Table 1. Formulations of Spectral Indices

Abbr.	Formula
NDVI	(NIR – Red) / (NIR + Red)
SAVI	((NIR – R) / (NIR + R + L)) * (1 + L)
NDBI	(SWIR – NIR) / (SWIR + NIR)
NDWI	(NIR – SWIR) / (NIR + SWIR)

The study area contains eight major land cover classes according to the Earth Cover reference dataset with a spatial resolution of 10 meters, produced using Sentinel-2A imagery and an overall accuracy of approximately 74% (https://esa-worldcover.org). However, as it is assumed that the limited number of pixels belonging to some land classes may affect the accuracy of the segmentation models, some land classes have been combined. As a result, the dataset included five land cover classes: tree cover, meadows, cultivated land, impervious surface including man-made structures such as roads and buildings, water bodies consisting of herbaceous vegetation, and sea (Figure 2).

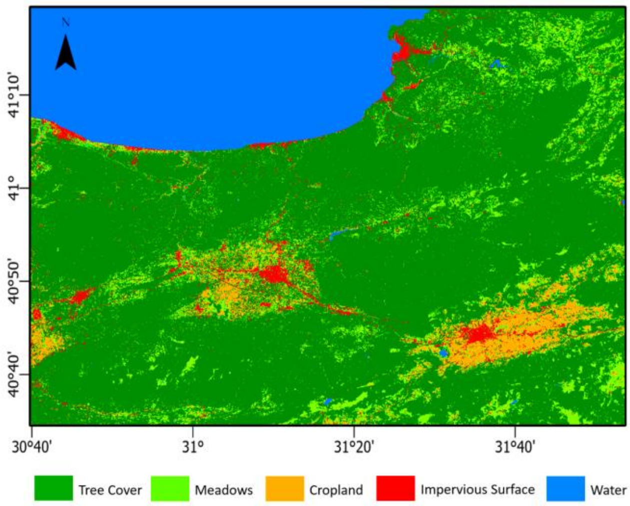

Figure 2. The study area is based on WorldCover reference data with five land cover classes

3. Methodology

In this study, the Attention U-Net was employed for semantic segmentation (Chen et al., 2023). The attention module is added to the U-Net architecture to draw attention to prominent features passing through the skip links. To distinguish between irrelevant and noisy details in skip connections, information derived from a rough scale is employed in the attention function. This occurs just prior to the concatenation process to ensure that only pertinent activations are combined. Moreover, the U-net architecture includes 2D convolutional and up-sampling layers of encoder and decoder blocks. Attention gates are located between the encoding and decoding layers to emphasize the important features of the image. Furthermore, the optimizer function of Adamax was chosen for the model, and categorical cross entropy was used as the loss function to solve the multiclass segmentation problem. After arranging the architecture of the model, it was created with 10,000,000 trainable parameters.

The Attention U-Net architecture inherently requires inputs in the form of patches. Therefore, the satellite image and reference data were separately divided into 256x256 patches. In total, 811 patches were created for training and 542 patches were produced for both testing (271 patches) and validation (271 patches) analyses. After the datasets were normalized, the training of Attention U-Net architectures was completed. In the section on performance analysis of the deep learning models, Intersection Over Union (IoU) and meanIoU metrics were employed in the training stage while the meanIoU metric was calculated for the model findings. IoU, is characterized as the region of overlap between the anticipated and the ground truth region, which is then divided by the region of combination between both regions. Moreover, the evaluation of segmentation algorithms for multiclass problems often involves the utilization of meanIoU metric to assess their efficacy throughout all classes (Yeghiazaryan & Voiculescu, 2018).

4. Results

Semantic segmentation models trained on two datasets were completed without overfitting in the training phase. In both applications, the training and validation accuracies of the models exceeded 0.92 while the training and validation losses of the models converged to 0.1 (Figure 3). Furthermore, 270 patches were used as test data to measure the accuracy of the trained models. Some of the predicted images are shown in Figure 4. In addition, both meanIoU values and class based IoU values were calculated separately for the two semantic segmentation outputs for accuracy evaluation (Table 2).

When examining the class-based IoU values in both datasets, it was observed that the dataset containing spectral indices related to water yielded the highest IoU value of 0.991, while the dataset containing only spectral bands yielded an IoU value of 0.899 (Table 2). Conversely, it was observed that the datasets consisting of spectral indices and spectral bands for the meadow class yielded the minimum IoU values of 0.597 and 0.570 in both datasets. Furthermore, one of the reasons for the low IoU of the meadow class is that it is mixed with cropland and impervious surface classes. The findings prove that using spectral indices obtained from satellite images for land cover classification improves the accuracy of semantic segmentation applications.

Table 2. Class-based IoU values for both datasets

LULC Class	Spectral Bands	Spectral Indices
Tree cover	0.832	0.956
Meadows	0.570	0.597
Cropland	0.721	0.820
Impervious Surface	0.593	0.733
Water	0.899	0.991
mIoU	0.723	0.819

5. Conclusion

In recent years, significant progress has been made in image segmentation applications. Recent research based on deep learning approaches has improved segmentation accuracy. In general, the main challenge in deep learning image segmentation is the lack of reference data or problems in generating such data. This study proves that the reference data, namely WorldCover provided by ESA, is useful for image segmentation applications. Furthermore, the effect of spectral indices to improve the accuracy of semantic segmentation was investigated and an increase in accuracy of about 10% was achieved. In other words, the use of spectral indices in addition to spectral bands significantly improves the accuracy of semantic segmentation applications. It can be deduced that semantic segmentation and deep learning models have the capacity to make significant contributions in a wide range of areas, such as visual understanding, object recognition and image classification/segmentation. Future work includes the development of a deep learning model that uses semantic segmentation techniques for multispectral remote sensing images to produce results with superior accuracy compared to machine learning with OBIA.

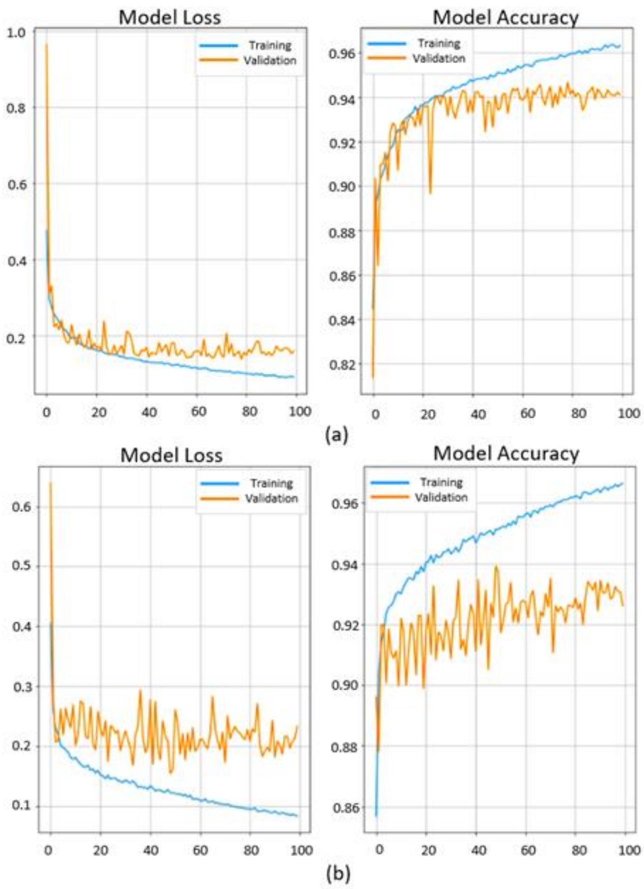

Figure 3. The learning curves for (a) Sentinel-2A with spectral indices and (b) Sentinel-2A spectral bands

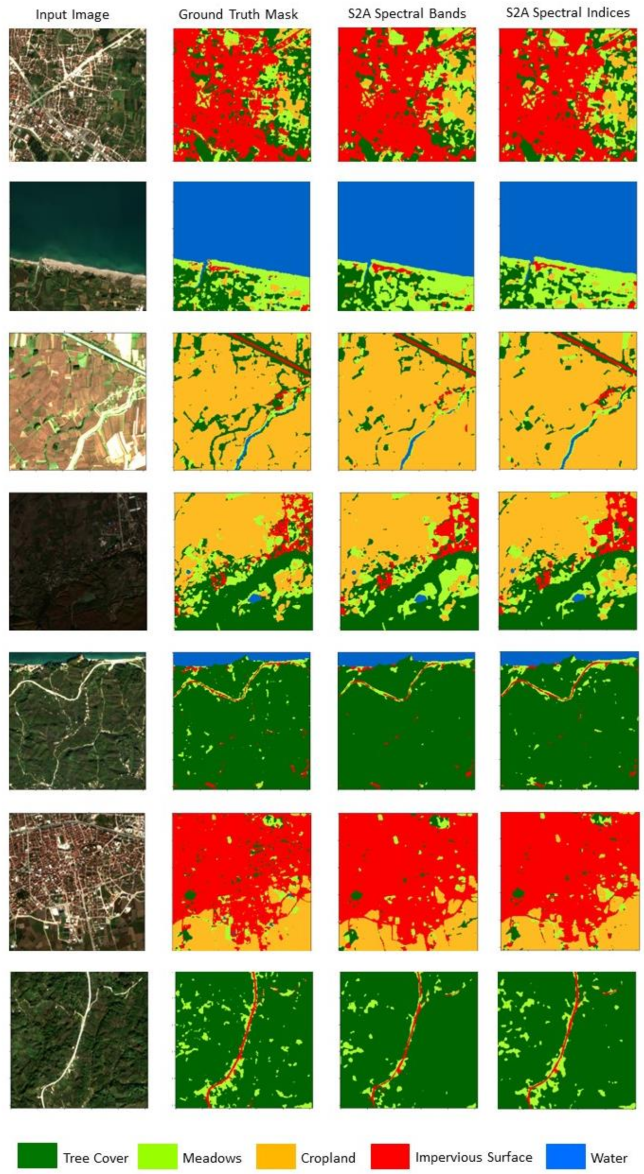

Figure 4. Predictions of semantic segmentation models

References

Chen, W., Hua, W., Ge, M., Su, F., Liu, N., Liu, Y., & Xiong, A. (2023). Severe Precipitation Recognition Using Attention-UNet of Multichannel Doppler Radar. Remote Sensing, 15(4), 1111.

Kavzoglu T. & Colkesen I. (2013). An assessment of the effectiveness of a rotation forest ensemble for land-use and land-cover mapping, International Journal of Remote Sensing, 34(12), 4224–4241.

Kavzoglu T., & Mather P. M., (2003), The use of backpropagating artificial neural networks in land cover classification, International Journal of Remote Sensing, 24(23), 4907–4938.

Kavzoglu, T., Erdemir, M. Y. & Tonbul, H. (2016). A region-based multi-scale approach for object-based image analysis. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, Volume XLI-B7, 241-247.

Kavzoglu, T., Teke, A. & Yilmaz, E.O. (2021). Shared blocks-based ensemble deep learning for shallow landslide susceptibility mapping, Remote Sensing, 2021, 13, 4776. https://doi.org/10.3390/rs13234776

Kavzoglu, T. & Tonbul, H. (2017). A comparative study of segmentation quality for multi-resolution segmentation and watershed transform. 8th International Conference on Recent Advances in Space Technologies (RAST'2017), 113-117.

Kavzoglu, T. & Tonbul, H. (2018). An experimental comparison of multi-resolution segmentation, SLIC and K-means clustering for object-based classification of VHR imagery. International Journal of Remote Sensing, 39(18), 6020-6036.

Kavzoglu, T. & Yilmaz, E.O. (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. Turkish Journal of Electrical Engineering and Computer Sciences, 30(6), 2124-2144.

Kushwah C. P. & Markam K., (2021), Semantic segmentation of satellite images using deep learning, International Journal of Innovative Technology and Exploring Engineering, 10(8), 33–37.

Maxwell A. E., Warner T. A., Guillén L. A., (2021a), Accuracy Assessment in Convolutional Neural Network-Based Deep Learning Remote Sensing Studies—Part 1: Literature Review, Remote Sensing, 13(13), 2450.

van der Meer F., (2012), Remote-sensing image analysis and geostatistics, International Journal of Remote Sensing, 33(18), 5644–5676.

Shafique, A., Cao, G., Khan, Z., Asad, M., & Aslam, M. (2022). Deep learning-based change detection in remote sensing images: A review. Remote Sensing, 14(4), 871.

Teke, A., Yilmaz, E. O. & Kavzoglu, T. (2021). Comparative assessment of deep learning and machine learning learning models in shallow landslide susceptibility. International Symposium on Applied Geoinformatics (ISAG2021) 2-3 December 2021, Riga, Latvia.

Tonbul, H., Colkesen, I. & Kavzoglu, T. (2020). Classification of poplar trees with object-based ensemble learning algorithms using Sentinel-2A imagery. Journal of Geodetic Science, 10(1), 14-22.

Tonbul, H., & Kavzoglu, T. (2017). Investigation of the effect of segmentation quality on object-based classification accuracy, Afyon Kocatepe University Journal of Science and Engineering, 17, 118-125.

Tsagkatakis G., Aidini A., Fotiadou K., Giannopoulos M., Pentari A., Tsakalides P., (2019), Survey of deep-learning approaches for remote sensing observation enhancement, Sensors (Switzerland), 19(18), 3929.

Yeghiazaryan, V., & Voiculescu, I. (2018). Family of boundary overlap metrics for the evaluation of medical image segmentation. Journal of Medical Imaging, 5(1), 015006-015006.

Yılmaz, E. Ö. & Kavzoğlu, T. (2021). Basic principles of deep learning and its applications in remote sensing. Map Journal (Harita Dergisi), 87(166), 25-43.

Yildiz, M., Kavzoglu, T., Colkesen, I. & Sahin, E. K. (2012). An assessment of the effectiveness of segmentation methods on classification performance. In 10th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 133-138.