

# AUTOMATIC CLASSIFICATION OF SELECTED CORINE CLASSES USING DEEP LEARNING BASED SEMANTIC SEGMENTATION

D. B. Demir <sup>a\*</sup>, N. Musaoglu <sup>a</sup>

<sup>a</sup> ITU, Civil Engineering Faculty, 80626 Maslak Istanbul, Türkiye - (kilincoglu, musaoglune)@itu.edu.tr

**KEY WORDS:** CORINE, LULC, U-NET, ResNet, Deep Learning.

## ABSTRACT:

In this study, deep learning-based semantic segmentation is used to automatically generate CORINE land cover (CLC) Level 2 classes for a test region in Türkiye. This is accomplished by utilizing new datasets and models created from a pilot region in Italy, which exhibits similar land use/land cover (LU/LC) characteristics to the test region in Canakkale/Türkiye. The training and validation datasets for Italy were generated by employing Sentinel-2 images from various months and different band combinations, along with CLC 2018 vector data for labelling. Different datasets were created to investigate the impact of patch sizes (128 and 256 pixels) and seasonal changes in LU/LC. For the semantic segmentation task, the U-Net architecture was selected as the primary deep learning model. Furthermore, the U-Net architecture was used in conjunction with ResNet50 and ResNet101 for transfer learning, enabling the replacement of the encoder section of the U-Net. These models were tested in the Italy region, and the best-performing ones were subsequently applied to the Canakkale test region to automatically generate CLC 2018. The results were compared with published CLC 2018 Level 2 data for the same region, and the accuracy was assessed using the Intersection over Union (IoU) metric. The findings were presented both visually and statistically

## 1. INTRODUCTION

Countries have been working for years to create, manage and develop environmental awareness and policies. The most important of these studies is the accurate, reliable and rapid establishment of land use/land cover (LU/LC) classification systems. To achieve this, countries use remote sensing technologies to produce LU/LC maps with different methods in various scales. One of these is the CORINE project carried out by the European Environment Agency, which is produced with the visual interpretation technique of satellite images at 1/100.000 scale in 6-year time intervals. It consists of an inventory of LU/LC in 44 classes including three different levels. CORINE Land Cover (CLC) uses a Minimum Mapping Unit of 25 ha for areal phenomena and a minimum width of 100 m for linear phenomena (Feranec et al., 2016; Büttner, 2016). However, the CORINE data set is produced over a long time period, typically 6 years, and its reliance on visual interpretation poses a challenge as it results in a time-consuming production process and limited ability for timely updates. Consequently, automatic determining land cover classes as accurately as possible is an extremely significant challenge faced in the field of remote sensing. Nowadays, advances in remote sensing and image processing techniques offer new approaches to the automatic determination of land cover (Ma et al., 2019)

With the development of technologies, the frequent acquisition of satellite images, along with increasing spatial and spectral resolutions has brought difficulties in combating big data while ensuring data up-to-dateness. Thus, deep learning methods have become popular in remote sensing research (Yuan et al., 2020). The deep learning models developed by the researchers and the weights produced from the data sets with transfer learning feed each other for the new models.

The U-net algorithm based on Convolutional Neural Networks (CNN), is one of the most common in LU/LC studies for semantic segmentation.

Capable of summarizing patterns in both spectral and spatial domains, this model consists of two parts: encoder and decoder. In order to improve segmentation model performance, pre-trained networks with different backbones are used instead of encoder in various U-net approaches. In this study, Residual Networks (ResNet) backbones, specifically ResNet50 and ResNet101, were utilized. These backbones incorporate convolutional residual blocks, which are capable of effectively processing a large number of layers (Herlawati, 2022). The U-net algorithm, originally designed for biomedical image segmentation, has been widely employed in numerous studies. It finds applications not only on very high spatial resolution (VHR) images but also on medium resolution images characterized by wide coverage and high temporal resolution, such as Landsat or Sentinel, particularly within the realm of land cover classification studies (Ronneberger et al., 2015; Pollatos et al., 2020; Solórzano et al., 2021; Xu et al., 2023; Tzepkenlis et al., 2023).

In the context of this study, novel datasets were constructed to facilitate the automatic generation of CLC classes through deep learning-based semantic segmentation. To accomplish these Sentinel-2 satellite images were used. The U-net model was trained using Sentinel-2 Red, Green, Blue bands as well as Sentinel-2 Near-Infrared, Red, Green bands to generate more accurate and more diverse CLC classes. The datasets were created to evaluate the potential of the U-net in combination with different seasonal images and the models trained with these datasets were tested in different regions.

## 2. METHODOLOGY

The approach adopted for automatic CLC classification consists of four main parts as shown in Figure 1. These are dataset pre-processing, model training with different parameters, testing in selected areas, and derivation of final results. Each stage of the process plays a critical role in achieving accurate and reliable CLC classification.

\* Corresponding author

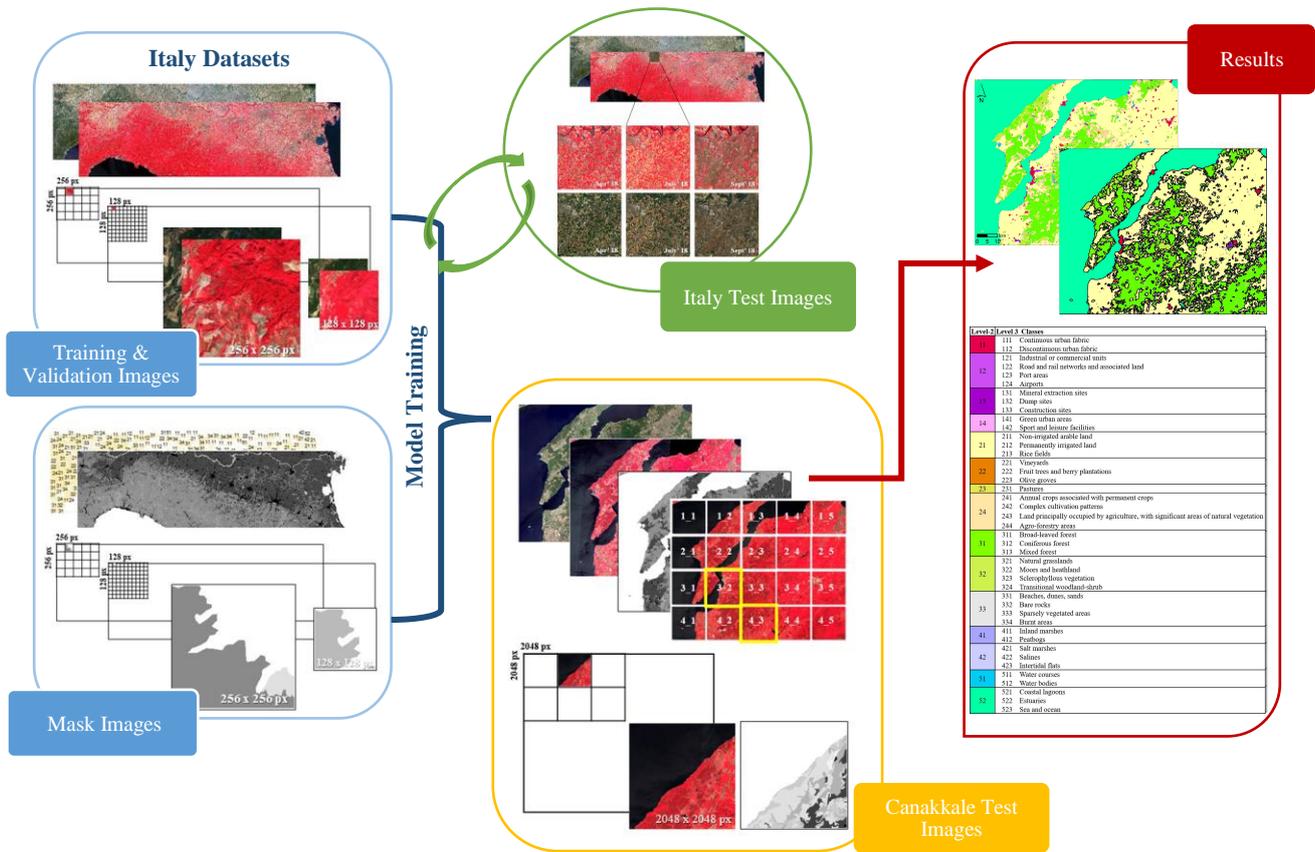


Figure 1. Flowchart of the automatic CLC classification approach.

## 2.1 Datasets

The study areas are located in two regions, one is the Italy region, which is selected for the training and validation dataset, and the other is the Canakkale region, which is located in the western part of Turkey, chosen for the testing (Figure 2). These regions were chosen to provide a diverse and representative set of conditions for evaluating the proposed methodology.

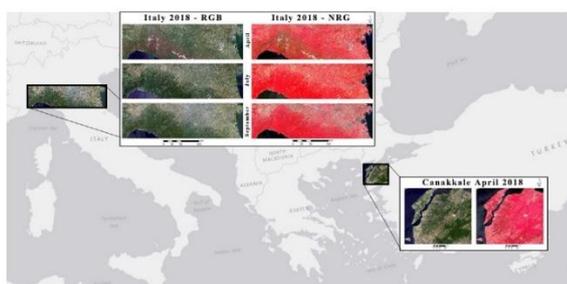


Figure 2. Italy & Canakkale Datasets.

The distribution of the CLC classes in CLC 2018 data was thoroughly analyzed, leading to the selection of the aforementioned regions. The Italy region was chosen for its diverse range of CLC classes, ensuring a comprehensive representation of land cover types. On the other hand, the Canakkale test area was selected due to its similarities to Italy region in terms of LU/LC characteristics, facilitating a meaningful evaluation of the proposed methodology in a comparable context.

New datasets were created using Sentinel-2 images and vector data downloaded from Copernicus in the same coordinate system. The cloudless Sentinel-2 images in different seasons were mosaicked for each season (April, July, September) by selecting true color composite images (RGB - Red, Green, Blue) and false color composite images (NRG - Near-infrared, Red, Green) at 10 m resolution. CLC 2018 vector data was modified according to Level 2 CLC classes and raster data with 10 m resolution was obtained for mask data.

For Italy region, April and September 2018 Sentinel-2 mosaic images were chosen as training and validation datasets and July 2018 Sentinel-2 mosaic were used for testing. For Canakkale region, April and September 2018 Sentinel-2 mosaic images were chosen in order to test model performance. All images were cropped to fit within the same boundaries in both regions. The mosaic images of Italy were automatically cut into patches of size 1,28 x 1,28 km (128 x 128 pixels) and 2,56 x 2,56 km (256 x 256 pixels) for training, validation and mask dataset (Figure 3).

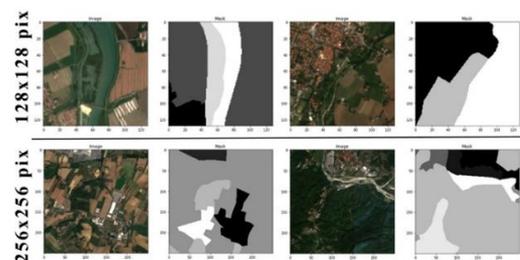


Figure 3. Patch sizes of the Italy dataset.

In order to generate the training and validation datasets, two different sizes (256 pixels, 128 pixels) of patches were created to see how patch size influence the results. In addition, effects of using multi- temporal data were examined by generating patches from both April and September images and from only September image. Total number of patches is given in Table 1-

Furthermore, to enable simultaneous testing of the models, the test images were divided into 20 tiles (1\_1, 1\_2, .. ,4\_5), each covering approximately 400 km<sup>2</sup> (2048 x 2048 pixels). This partitioning was based on the maximum size supported by the Graphics Processing Unit (GPU).

ITALY	256 pixel		128 pixel
	September	September & April	September & April
Train + Mask Images	3904	7219	15699
Validation + Mask Images	1302	2407	5233
Total	5206	9626	20932

Table 1. Italy dataset patch counts.

## 2.2 Models

**2.2.1 U-Net:** The U-Net architecture takes its name from its structure that resembles a “U” shape in the way it narrows and then expands symmetrically (Ronneberger et al., 2015). This model generates contracting (encoder) path and expansive (decoder) path. Encoder path is used to extract features from raw data, and the decoder path, consists of the up-convolutions and feature merges from the encoder. It is the decoder path that allows the network to learn spatial semantic segmentation information (Sarra et al., 2022).

**2.2.2 ResNet:** Based on the VGG-19 architecture, the ResNet network uses a 34-layer flat network architecture with fewer filters and lower complexity than VGG networks. This plain network is then transformed into a residual network by adding shortcut or residual blocks. In this way, it is aimed that faster learning will be achieved. ResNet50 is obtained by replacing every 2-layer block in the 34-layer network with a 3-layer bottleneck block, while ResNet 101–152 is generated from 101-layer and 152-layer ResNets using more 3-layer blocks. ResNet 50-101-152 layers are significantly more accurate than ResNet 34-layer ones (He et al., 2016).

## 2.3 Evaluation Metrics

To evaluate the performance of our different datasets on semantic segmentation models, we used the Intersection Over Union (IoU) metric, which is calculated as the intersection of predicted and ground truth values divided by their union (Equation (1)).

$$IoU = \frac{\text{Area of Overlap}}{\text{Area of Union}} \quad (1)$$

For an individual class, the IoU metric is defined as follows (Equation (2)).

$$IoU = \frac{TP}{TP+FP+FN} \quad (2)$$

where TP is the number of true positives, FP is the number of false positives, and FN is the number of false negatives. IoU value is bounded between 0 (when there is no overlap) and 1 (when predicted and ground truth match perfectly).

## 3. EXPERIMENTS

### 3.1 Model Training

Training and validation datasets in Italy with different band combinations, different seasons and different patch sizes were used as input to the U-Net model with different backbone and 5 different final models were generated. These model parameters shown in Table 2 were applied to the Italy and Canakkale test images.

Dataset	Bands	Methods	Transfer Learning	Optimizer	Epoch	Patch Size
Sept	RGB	U-Net	ResNet50	Adam	100	256
Sept	RGB	U-Net	ResNet101	Adam	100	256
Sept	NRG	U-Net	ResNet101	Adam	100	256
Sept & Apr	NRG	U-Net	ResNet101	Adam	100	256
Sept & Apr	NRG	U-Net	ResNet101	Adam	100	128

Table 2. Parameters of models.

### 3.2 Model Testing

Firstly, April, September, July 2018 images were divided into 2048 x 2048 pixel tiles for a test area in Italy. Subsequently, they were evaluated using U-Net models which were trained with only September RGB and NRG images utilizing 256 x 256 pixel patches. These models were produced with different backbones (ResNet50 and ResNet101) with 100 epochs (Figure 4).

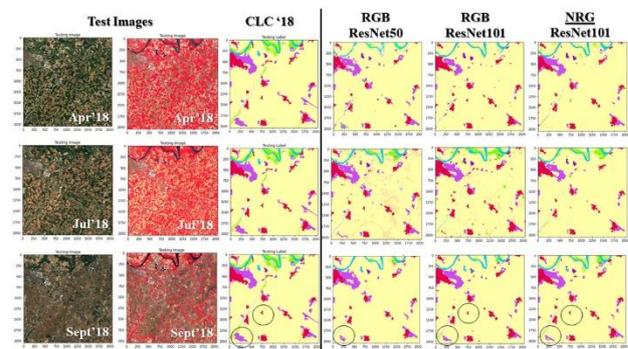


Figure 4. Model results of Italy test images.

The results show that, between April – July – September test images, September gave the best results, therefore model performances were compared using September test image.

IoU results for CLC classes of models are given in Table 3. The mean IoU values for ResNet50 and ResNet101 models trained with RGB images are 0.47 and 0.48, respectively. Although these two models produced similar results, ResNet101 achieved better results visually for CLC classes and this model was chosen for tests in Canakkale region. In addition, the results of the models trained with different band combinations on the test image are also given in Table 3. The comparison between the NRG and RGB band combinations show that NRG performed better than RGB, therefore it was chosen for Canakkale region.

Training Models	IoU for Class								
	11	12	13	21	31	32	33	41	51
Sept 256 px	RGB, ResNet50	0.64	0.57	0.72	0.92	0.17		0.49	0.82
	RGB, ResNet101	0.72	0.60	0.44	0.92	0.11	0.10	0.48	0.11
	NRG, ResNet101	0.74	0.59	0.58	0.94	0.23	0.29	0.55	0.37

Table 3. Comparison of ResNet50 & ResNet101.

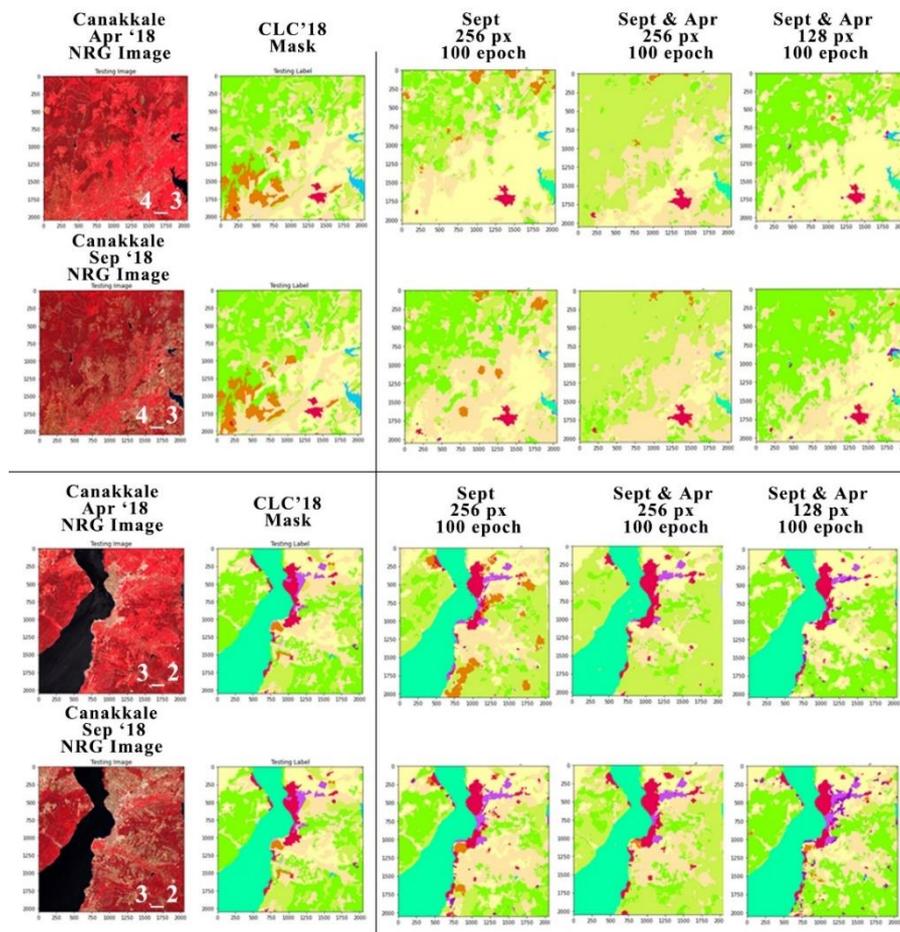


Figure 5. Model results of Canakkale 4-3 tile and Canakkale 3-2 tile.

Secondly, in order to test the applicability of the models trained in a different region to another region, which is the main objective of the study, test images taken in Canakkale in April and September were prepared.

Three U-Net models were prepared using the ResNet101 backbone trained on NRG images, and these models applied to the April and September test images from Canakkale. The first model was trained with September image of Italy with 256 patch size, second model was trained with both September and April images of Italy with 256 patch size and third model was trained with both September and April images of Italy with 128 patch size. Visual results can be seen in Figure 5, while the IoU results are presented in Table 4 and Table 5 for Canakkale 4-3 tile and Canakkale 3-2 tile.

		Canakkale 4_3 – April NRG			Canakkale 4_3 – September NRG		
IoU for class	Training Models	Sept 256px	Sept + Apr 256px	Sept + Apr 128px	Sept 256px	Sept + Apr 256px	Sept + Apr 128px
		11	0.73	0.71	0.69	0.67	0.68
	21	0.39	0.23	0.39	0.36	0.2	0.48
	24	0.14	0.28	0.2	0.32	0.32	0.31
	31	0.43	0.06	0.69	0.65	0.06	0.64
	32	0.26	0.24	0.17	0.29	0.25	0.22
	51	0.33	0.18	0.38	0.27	0.15	0.15

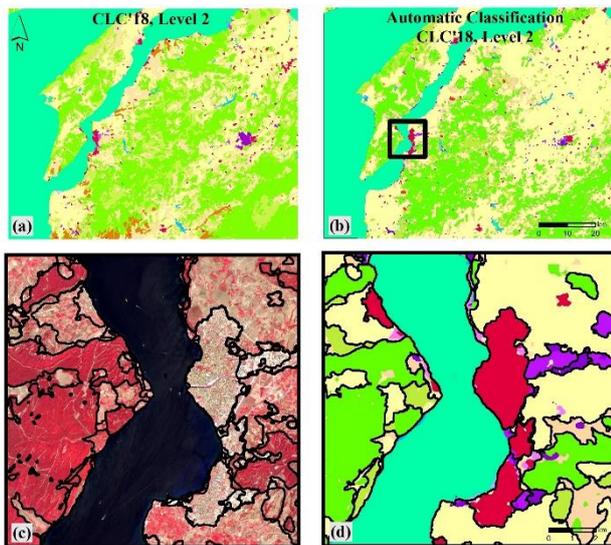
Table 4. IoU results of Canakkale 4\_3.

		Canakkale 3_2 – April NRG			Canakkale 3_2 – September NRG		
IoU for class	Training Models	Sept 256px	Sept + Apr 256px	Sept + Apr 128px	Sept 256px	Sept + Apr 256px	Sept + Apr 128px
		11	0.61	0.63	0.67	0.61	0.58
	12	0.26	0.26	0.27	0.28	0.27	0.30
	21	0.35	0.43	0.43	0.51	0.40	0.47
	22	0.01	0.02		0.24	0.17	0.23
	24	0.27	0.27	0.26	0.40	0.40	0.40
	31	0.25	0.18	0.69	0.63	0.30	0.72
	32	0.23	0.22	0.22	0.37	0.26	0.21
	52	0.98	0.98	0.98	0.99	0.98	0.99

Table 5. IoU results of Canakkale 3\_2.

According to results on tile 4\_3 and 3\_2, third model (Sept & Apr – 128 px) applied on April image was chosen to generate automatic semantic segmentation of the entire test area (20 tiles) of Canakkale.

Models generated by training & validation datasets with different parameters such as spectral band combination, image dates, patch size and backbones on Italy region, were analysed. The best performing models were then applied to sample tiles in the Canakkale test region. Afterwards, the entire Canakkale test region was automatically classified according to CLC Level 2 classes using the best performing U-Net model. This model trained with both NRG September and NRG April images of Italy, using a patch size 128 and the ResNet101 backbone. The results can be seen in Figure 6.



**Figure 6.** (a) CLC 2018 Level 2 Map by Copernicus. (b) Automatic CLC 2018 Level 2 Classification with U-Net model. (c) Automatic CLC 2018 Level 2 segmentation result. (d) Generalization of automatic CLC 2018 Level 2 segmentation to CLC standards.

The visual results demonstrate the effectiveness of our selected model to predict the CLC Level 2 classes 52, 31, 11, 21, 24, 12 (Sea, forest, urban fabric, agricultural areas, complex cultivation pattern and Industrial or commercial units) with satisfactory outcomes. However, certain rare classes such as 41, 42 (marshes) are quite difficult to predict.

#### 4. CONCLUSIONS

In this study, new datasets were created for automatically generating CLC classes by deep learning based semantic segmentation. The similarity between the seasonal characteristics of the training and test images led to improved classification results for certain CLC classes. Additionally, using small patch size in models increases number of CLC classes in results. Utilizing, different parameters can further improve model performance and expanding the dataset achieve higher accuracy. The accuracy of the original CLC dataset is very important, it contains imperfections due to the fact that some classes are not classified correctly and are generalized, thus affecting the IoU evaluation on the predicted image. Therefore, evaluating automatic classification should incorporate various accuracy assessment methods, such as thematic accuracy. Finally, the results show that deep learning methods can be used for automatic CLC classification.

#### REFERENCES

Büttner, G., 2016. CORINE land cover products. In European Landscape Dynamics: CORINE Land Cover Data, pp. 55–58.

Feranec, J., Hazeu, G., Kosztra, B., & Arnold, S., 2016. CORINE land cover nomenclature. In European Landscape Dynamics: CORINE Land Cover Data, pp. 17–26.

He, K., Zhang, X., Ren, S., & Sun, J., 2016. Deep residual learning for image recognition. Proceedings of the IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2016-December, pp. 770–778.

Herlawati, H., Handayanto, R. T., Atika, P. D., Sugiyatno, S., Rasim, R., Mugiarto, M., Hendharsetiawan, A. A., Jaja, J., & Purwanti, S., 2022. Semantic Segmentation of Landsat Satellite Imagery. 2022 7th International Conference on Informatics and Computing, ICIC 2022, pp. 1–6.

Ma, L., Liu, Y., Zhang, X., Ye, Y., Yin, G., & Johnson, B. A., 2019. Deep learning in remote sensing applications: A meta-analysis and review. *ISPRS Journal of Photogrammetry and Remote Sensing*, 152(November 2018), pp. 166–177.

Pollatos, V., Kouvaras, L., & Charou, E., 2020. Land cover semantic segmentation using ResUNet. CEUR Workshop Proceedings, 2844, pp. 55–65.

Ronneberger, O., Fischer, P., Brox, T., 2015. U-net: Convolutional networks for biomedical image segmentation. Lecture Notes in Computer Science in Medical Image Computing and Computer-Assisted Intervention, International Conference on Medical Image Computing and Computer-Assisted Intervention; Springer: Cham, Switzerland, 9351, pp. 234–241.

Sarra, K., & Boulmerka, A., 2022. Semantic segmentation of remote sensing images using U-net and its variants: Conference New Technologies of Information and Communication (NTIC 2022). 2nd IEEE International Conference on New Technologies of Information and Communication, NTIC 2022 - Proceeding, Ntic, pp. 1–5.

Solórzano, J. V., Mas, J. F., Gao, Y., & Gallardo-Cruz, J. A., 2021. Land use land cover classification with U-net: Advantages of combining Sentinel-1 and Sentinel-2 imagery. *Remote Sensing*, 13(18).

Tzepkenlis, A., Marthoglou, K., & Grammalidis, N., 2023. Efficient Deep Semantic Segmentation for Land Cover Classification Using Sentinel Imagery. *Remote Sensing*, 15(8).

Xu, L., Liu, Y., Shi, S., Zhang, H., & Wang, D., 2023. Land-Cover Classification with High-Resolution Remote Sensing Images Using Interactive Segmentation. *IEEE Access*, 11(January), pp. 6735–6747.

Yuan, Q., Shen, H., Li, T., Li, Z., Li, S., & Jiang, Y., 2020. Remote Sensing of Environment Deep learning in environmental remote sensing: Achievements and challenges. *Remote Sensing of Environment*, 241(January), 111716.