İSTANBUL UNIVERSITY PRESS

International Journal of Environment and Geoinformatics
2025, 12 (2): 34–44

https://doi.org/10.26650/ijegeo.1644726

Submitted: 21.02.2025
Revision Requested: 28.02.2025
Last Revision Received: 08.03.2025
Accepted: 25.03.2025

International Journal of Environment and Geoinformatics

Research Article

Open Access

CORINE Level-2 Land Cover Classification from Sentinel-2 Images Using U-Net Deep Learning Model

Emin Atabey Peker [1], Murat Çalışkan [1], Eser Bora [1], Çiğdem İnan [1] & M. Erkan Uçaner [1]

[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye

Abstract This study implements and evaluates an advanced U-Net deep learning architecture for semantic segmentation of multi-spectral Sentinel-2 imagery to classify CORINE (Coordination of Information on the Environment) Level 2 land cover classes in Türkiye. We implemented an advanced U-Net architecture incorporating spatial attention mechanisms and residual connections, utilizing multi-spectral Sentinel-2 imagery (Red, Green, Blue [RGB], Near-Infrared [NIR], and Short-Wave Infrared [SWIR] bands). A total of 92 working units were identified across Türkiye to represent all 15 classes of Level 2, covering 12,000 square kilometers (approximately 1.5% of the country's total area). The imagery was collected between June and September 2018, corresponding to the original CORINE 2018 production timeframe. The model was trained on 2,815 patches of 256x256 pixels generated from these working units. Our architecture integrates an Atrous Spatial Pyramid Pooling (ASPP) block for multi-scale feature extraction and employs a hybrid loss function combining Categorical Cross-Entropy and Jaccard Loss. The model achieved a mean Intersection over Union (IoU) of 0.5225 across all classes, with particularly strong performance in water bodies (Marine Waters: 0.8884 IoU, 0.9291 accuracy; Inland Waters: 0.8113 IoU, 0.8791 accuracy) and permanent crops (0.7118 IoU, 0.8389 accuracy). While the model excelled in classifying spectrally distinct classes, it faced challenges with classes sharing similar characteristics, such as artificial non-agricultural vegetated areas (0.0896 IoU, 0.1017 accuracy) and pastures (0.2549 IoU, 0.3504 accuracy). This research demonstrates the potential of deep learning approaches for automated CORINE land cover mapping, while highlighting areas for future improvement in distinguishing between spectrally similar classes.

Keywords CORINE · Deep Learning · U-Net · Remote Sensing · Sentinel-2

Citation: Peker, E. A., Çalışkan, M., Bora, E., İnan, Ç. & Uçaner, M. E. (2025). CORINE level-2 land cover classification from sentinel-2 images using U-Net deep learning model. *International Journal of Environment and Geoinformatics*, *12*(2), 34-44. https://doi.org/10.26650/ijegeo.1644726

This work is licensed under Creative Commons Attribution-NonCommercial 4.0 International License.

© 2025. Peker, E. A., Çalışkan, M., Bora, E., İnan, Ç. & Uçaner, M. E.

Corresponding author: Emin Atabey Peker eminatabeypeker@gmail.com

Introduction

Land cover and land use classification systems play a vital role in environmental monitoring, urban planning, and the development of sustainable policies. Land cover affects environmental conditions in many ways, by influencing water quality, watershed hydrology, habitat and species composition, climate, and carbon storage, while also affecting the mass and energy exchanges between the surface and the atmosphere that impact weather and climate patterns (U.S. EPA, 2024). Accurate and timely land use and land cover maps are essential for a variety of applications including natural resources and environmental management, disaster and hazard monitoring, and food security. These maps may help tackle significant large-scale challenges such as global warming, accelerating loss of species habitat, unprecedented population migration, increasing urbanization, and growing inequalities within and between nations (Zhang & Li, 2022). Among these classification systems, CORINE (Coordination of Information on the Environment) stands as a significant example, produced by the European Environment Agency every six years. The project, which started in the early 1980s with the aim of monitoring land cover at a continental scale, has produced datasets in 1990, 2000, 2006, 2012, and 2018. This inventory hierarchically categorizes land cover into five classes at Level 1, 15 classes at Level 2, and a total of 44 different land cover classes at Level 3. CORINE uses a Minimum Mapping Unit of 25 hectares for areal phenomena and a minimum width criterion of 100 meters for linear phenomena. For more sensitive monitoring of land cover changes, a more detailed Minimum Mapping Unit of 5 hectares has been adopted in the change layers (Copernicus Land Monitoring Service (CLMS), 2025). Traditional CORINE mapping relies heavily on visual interpretation by experts, which is both time-consuming and potentially subject to inconsistencies. Furthermore, the six-year interval between updates limits the dataset's ability to capture rapid land cover changes. These limitations have stimulated interest in developing automated approaches for land cover classification. Recent advancements in deep learning have shown promising results for automating this process, with several studies exploring various architectural approaches. To provide context for our approach, we highlight key studies that have explored automated CORINE classification techniques. Balado et al. (2018) evaluated the potential of airborne LIDAR data for automatic CORINE land cover classification using a ResNet-50 architecture. Their methodology involved extracting height difference, average intensity, and number of returns from point clouds and achieved 73.5% accuracy and a 59.8% Cohen Kappa coefficient at Level 1 CORINE classification. Interestingly, this LIDAR-based approach outperformed classification using aerial orthophotos on the same area (68.6% accuracy), though both remained below human interpreter accuracy (84.1%). Their research also identified specific strengths in classifying continuous urban fabric (84.9% accuracy), pastures (88.1%), agricultural lands (76.2%), and broad-leaved forests (83.3%), while highlighting challenges in distinguishing between spectrally similar classes like discontinuous urban fabric and agricultural areas. Demir & Musaoglu (2023) employed U-Net architecture with different ResNet backbones for automatic CORINE land cover classification, achieving impressive IoU results for Level 2 classes such as 0.73 for urban fabric and 0.99 for sea areas using Sentinel-2 imagery. Their study demonstrated the effectiveness of using different band combinations (RGB and NRG) and the advantages of incorporating multi-seasonal data (April and September) in the training process, showing that models trained in one region (Italy) could be successfully applied to other regions (Çanakkale, Turkey). The authors noted that "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." Varga et al. (2021) analyzed the spectral characteristics of the CORINE Land Cover (CLC) 2018 dataset using three different satellite images (Landsat-8, Sentinel-2, and PlanetScope). The study examined Level 1 land cover categories using statistical tests and machine learning classification algorithms (Linear Discriminant Analysis and Random Forest), demonstrating that these categories could be classified with approximately 70% accuracy. The research revealed that using CORINE data directly as a reference can be misleading while highlighting that high-resolution satellite imagery and the Random Forest algorithm could potentially improve classification performance. Stara & Halounova (2022) evaluated the classification of CORINE Land Cover classes using Sentinel-2 satellite imagery and the Random Forest algorithm with multiple input sources (optical, NDVI, topographic features). The study revealed that permanently irrigated arable lands were classified almost perfectly with F1 scores between 96-99%, while pastures and natural grasslands showed classification performance between 77-85%. The research highlighted the limitations of the CORINE dataset in classification accuracy, emphasizing the importance of topographic features and multitemporal approaches in land cover classification across different locations. Balzter et al. (2015) investigated the potential of Sentinel-1A SAR imagery for CORINE land cover mapping in Thuringia, Germany, using a Random Forest classification

approach. By combining radar backscatter data from May and December 2014 with Shuttle Radar Topography Mission (SRTM) elevation data, the study achieved a classification accuracy of 68.4% across 27 hybrid CORINE land cover classes. The research demonstrated that integrating geomorphometric information significantly improved land cover classification, particularly in distinguishing between agricultural crops and forest cover types. This pioneering study highlighted Sentinel-1's utility for land cover mapping, especially in cloud-covered regions, and proposed a methodological framework for future large-scale land cover monitoring efforts. LULC classification techniques have evolved significantly over time. Initially, visual interpretation classification was the primary method, offering high accuracy but suffering from poor repeatability and timeliness. As computer technology advanced, traditional machine learning approaches became widespread, including unsupervised methods like K-Means and Expectation Maximization, and supervised techniques such as Support Vector Machines (SVM), decision trees, and maximum likelihood classification. While these methods improved repeatability and processing time compared to visual interpretation, they often showed reduced accuracy when applied to new study areas or different datasets (Zhao et al., 2023). In recent years, deep learning-based semantic segmentation, particularly using Convolutional Neural Networks (CNNs), has emerged as a powerful technique for pixel-wise classification in remote sensing imagery. These approaches can potentially reduce the production time and improve the consistency of land cover maps. Among various CNN architectures, several specialized segmentation models have demonstrated remarkable performance in remote sensing applications. In this study, we propose and evaluate an advanced U-Net architecture for semantic segmentation of multi-spectral Sentinel-2 imagery. Our primary objective is to develop a robust deep learning approach for land cover classification, using CORINE Level 2 categories as our reference dataset. Through this approach, we aim to contribute to the methodological advancement of deep learning-based semantic segmentation for remote sensing applications.

Materials and Methods

Preprocessing

In the automatic classification study of Corine level-2 classes (Table 1), 92 working units were initially identified across the country (Figure 1). During this identification, the aim was to include representative samples of each of the 15 CORINE level-2 classes. These 92 working units correspond to a total area of 12,000 square kilometers, and the class-based distribution is shown in Table 2 Sentinel-2 images of these working units were downloaded from the Google Earth Engine platform (Gorelick, et al., 2017), adhering to the dates used during the Corine 2018 project production. During image download, NIR, SWIR 1, and SWIR 2 bands were downloaded along with RGB bands, and all bands were subsequently normalized. Multi-spectral satellite imagery provides significant advantages for land cover classification by offering meaningful spectral information that cannot be obtained from visible band data alone. While RGB bands capture visible surface characteristics essential for distinguishing artificial structures and basic land cover types, the near-infrared (NIR) band delivers crucial complementary information. NIR images are particularly valuable for land cover classification as they provide additional data about vegetation characteristics such as types, health, and water content that are not visible in RGB spectrum. The spectral range of NIR (750-1100 nm) falls within the solar radiation peak, allowing for high-quality data with good resolution and signal-to-noise ratio. The combination of RGB and NIR bands enhances discrimination between different land cover classes, especially for vegetation-related categories, making this multi-spectral approach widely used in the fields of agriculture, forestry, and geology/natural resources (Baek, Lee, & Jung, 2024). Short wave infrared (SWIR) bands are particularly valuable for land cover classification as they allow scientists to estimate water content in plants and soil through their unique absorption properties. This capability makes SWIR bands essential for characterizing different vegetation types and wetland areas, as variations in moisture content create distinctive spectral signatures that can be detected in the SWIR wavelength range. Additionally, SWIR imagery is especially effective for detecting and classifying mining areas, as different mineral compositions can be distinguished based on their unique spectral responses in the SWIR range (1.2 to 2.33 µm), providing critical information for land cover classification that is not available in the visible spectrum alone (Johnson & Koperski, 2017).

For labels, areas corresponding to the 92 working units from the digitized vector format Corine 2018 public dataset were used. These ground truth data were not collected through field validation in this study, but were obtained from the official CORINE Land Cover dataset that has been verified and validated by the European Environment Agency before being published on their official website. This dataset represents the authoritative reference data for CORINE land cover classifications throughout Europe, providing validated ground truth information. In this dataset, level 3 codes were mapped to their corresponding level 2 categories by using the first two digits of the level 3 codes (e.g., level 3 code “123” was mapped

to level 2 category "12"). In total, 15 classes were included in the study. In the next phase, the vector data to be used as labels were converted to raster format and divided into 256x256 patches. The total of 2,815 patches obtained were split into train, validation, and test datasets at the ratios shown in Table 3. Preprocessing operations were performed using GDAL (3.9.2) for raster data, while NumPy (1.26.4) and Pandas (2.2.3) were used for numerical operations and data manipulation. Vector data processing and geometric operations were handled using GeoPandas (0.14.4) and Shapely (2.0.6) libraries.

Table 1. Corine Level 1 and Level 2 Classes

Level 1	Level 2
1. Artificial Surfaces	1.1 Urban fabric
	1.2 Industrial, commercial and transport units
	1.3 Mine, dump and construction sites
	1.4 Artificial, non-agricultural vegetated areas
2. Agricultural Areas	2.1 Arable land
	2.2 Permanent crops
	2.3 Pastures
	2.4 Heterogeneous agricultural areas
3. Forest and Semi Natural Areas	3.1 Forest
	3.2 Scrub and/or herbaceous vegetation associations
	3.3 Open spaces with little or no vegetation
4. Wetlands	4.1 Inland wetlands
	4.2 Coastal wetlands
5. Water Bodies	5.1 Inland waters
	5.2 Marine waters

Table 2. Distribution of Corine Level-2 Land Cover Classes Across 92 Working Units (in percentages)

Corine Level-2 Class Distributions Used in the Study	%
Urban fabric	8.77
Industrial, comercial and transport units	3.27
Mine, dump and construction sites	0.95
Artificial, non-agricultural vegetated areas	0.48
Arable land	20.99
Permanent crops	5.52
Pastures	2.54
Heterogeneous agricultural areas	9.71
Forest	7.31
Shrub and/or herbaceous vegetation associations	9.54
Open spaces with little or no vegetation	8.63
Inland wetlands	3.30
Coastal wetlands	4.72
Inland waters	6.36
Marine waters	7.91

Table 3. Distribution of Dataset Splits for Model Training and Evaluation

Initial data dimensions	**Images shape: (2815, 256, 256, 6)**
	Labels shape: (2815, 256, 256, 15)
Dimensions after dataset split	Train set
	X_train shape: (1801, 256, 256, 6)
	y_train shape: (1801, 256, 256, 15)
	Validation set
	X_val shape: (451, 256, 256, 6)
	y_val shape: (451, 256, 256, 15)
	Test set
	X_test shape: (563, 256, 256, 6)
	y_test shape: (563, 256, 256, 15)
Dataset split ratios	Train set: 63.98%
	Validation set: 16.02%
	Test set: 20.00%

Proposed Model

The proposed model is based on the U-Net architecture, which was first introduced in 2015 for biomedical image segmentation, particularly for analyzing microscopy images. The U-Net architecture consists of two main components: an encoder (contracting path) and a decoder (expansive path). In the encoder part, the input image goes through multiple convolution and pooling operations, where the spatial dimensions are progressively reduced while the number of feature channels increases (Özturk et al., 2020). This process allows the network to capture and learn important features and context from the input image. In the decoder part, the network performs upsampling operations to restore the

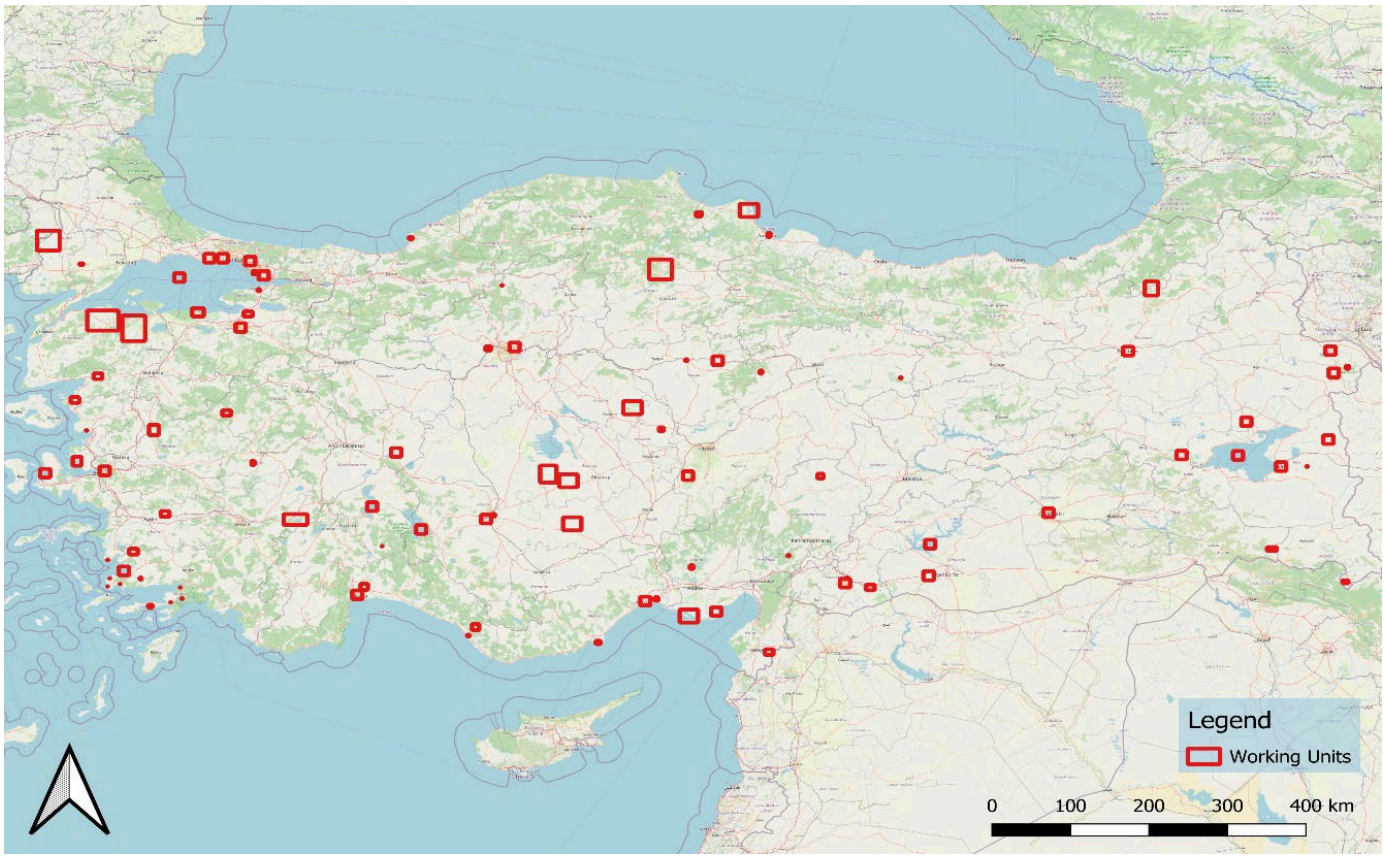


Figure 1. Distribution of Working Units Used for Training Data Collection

spatial dimensions while decreasing the number of feature channels, utilizing up-convolutions and concatenations with corresponding encoder features through skip connections. This setup enables the network to reconstruct detailed segmentation maps by combining the high-level features learned in the encoder with the spatial precision preserved through skip connections (Ronneberger, Fischer, & Brox, 2015). In this study, our advanced U-Net architecture is implemented for image segmentation (Figure 2). The model takes an input image of size (256, 256, 6) and utilizes an encoder structure to process it. In the encoder part, convolutional blocks with residual connections and spatial attention mechanisms are employed. At each encoder level, the spatial dimensions of the image are halved using MaxPooling, while the number of channels is increased (starting from 64 up to 512). In the residual convolutional blocks, regularization techniques such as batch normalization and dropout are used to improve the model's generalization capability. At the deepest point of the encoder, an ASPP (Atrous Spatial Pyramid Pooling) block is present. This block extracts multi-scale features using convolutions with different dilation rates and global pooling. This approach allows the model to better perceive objects of various sizes. The ASPP block employs 1x1, 3x3 convolutions (with dilation rates of 6, 12, and 18), and global average pooling to create a rich feature representation. In the decoder part, the features obtained from the encoder are progressively upsampled to higher resolutions. At each decoder level, the spatial dimensions are doubled using Conv2DTranspose, and skip connections from the encoder are merged. These skip connections help preserve low-level details. Residual convolutional blocks are also used in the decoder to further process the features. Finally, the model produces a probability distribution among 15 CORINE land cover classes for each pixel using a 1x1 convolution layer followed by a softmax activation. This ensures that the total probability for each pixel sums up to 1 across the 15 different classes. The spatial attention mechanisms in the model enable it to focus on important features, while the residual connections enhance gradient flow and allow for a deeper model architecture.

Model Parameters

A comprehensive parameter configuration was used for model training and optimization. The AdamW optimizer was chosen due to its superior generalization performance compared to Adam (Loshchilov & Hutter, 2017), with a weight decay value of 1e-5. A combination of Categorical Cross-Entropy and Jaccard Loss with equal weighting (0.5-0.5) was employed as the loss function (Equation 1,2,3). The Jaccard Loss function

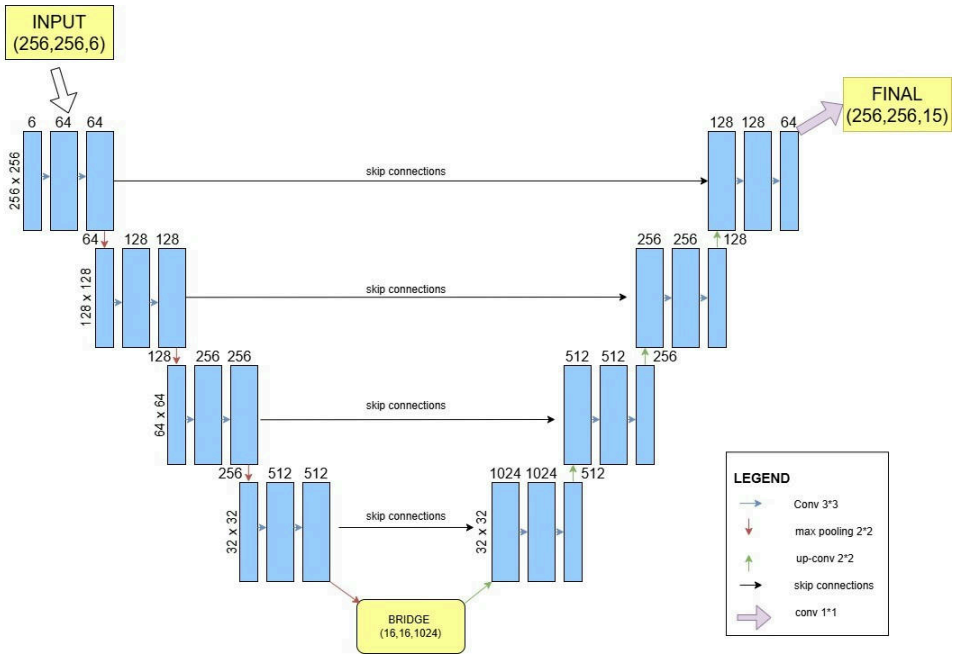


Figure 2. The U-Net Architecture of the Proposed Model

was customized to be calculated separately for each class and averaged, with a smooth parameter of 1e-7 added for numerical stability. To evaluate model performance, a customized MeanIoU (Equation 4,5) metric was used, modified to handle one-hot encoded predictions. Additionally, class-wise Accuracy (Equation 6) was calculated to assess the proportion of correctly classified pixels for each class. The learning rate strategy was managed with a custom scheduler for a total of 200 epochs. The initial learning rate was set to 5e-4, with the first 5 epochs designated as a warmup period. Rapid learning was achieved during the first 60 epochs with learning rates around 1e-4, followed by a gradual decrease from 1e-4 to 1.88e-5 between epochs 60 and 100. At the 100th epoch, the learning rate was reduced to 9e-6, and it continued to decrease linearly to 1e-6 over the next 100 epochs. Several callback mechanisms were utilized during model training: Early stopping was implemented to halt training if the validation loss did not improve for 25 consecutive epochs, with the best weights being restored. The ModelCheckpoint callback was used to save model weights at the point of minimum validation loss. Additionally, the ReduceLROnPlateau callback was employed to reduce the learning rate by a factor of 0.5 if the validation loss did not improve for 10 epochs, with a minimum learning rate of 1e-6. The model was trained using 8 CPU cores, 64GB RAM, and an NVIDIA L40S GPU with 48GB memory. The training process took approximately 2 hours to complete.

$$\text{Categorical Cross-Entropy Loss (CCE)}(\mathrm{y},\,\hat{\mathrm{y}}) = -\sum_{i=1}^{C} y_i\, log(\hat{\mathrm{y}}_i) \tag{1}$$

y_i : True Class Label (one-hot encoded)

$\hat{\mathrm{y}}_i$: Predicted Probability for Class i

CCE: Number of Classes (15 for CORINE Level 2)

$$\text{Jaccard Loss} = 1 - |A \text{ sect } B|/|A \text{ union } B| = 1 - ((\sum_{i=1}^{n} y_i\,\hat{\mathrm{y}}_i + \varepsilon)/(\sum_{i=1}^{n} y_i + \sum_{i=1}^{n} \hat{\mathrm{y}}_i - \sum_{i=1}^{n} y_i\,\hat{\mathrm{y}}_i + \varepsilon)) \tag{2}$$

y_i : True Class Values

$\hat{\mathrm{y}}_i$: Predicted Class Values

ϵ : Small Value Added for Numerical Stability

$$\text{Combined Loss Function} = 0.5 \text{ x CCE } (\mathrm{y}, \hat{\mathrm{y}}_i) + 0.5 \text{ x Jaccard Loss } (\mathrm{y}, \hat{\mathrm{y}}_i) \tag{3}$$

$$\text{IOU} = TP/(TP + FP + FN) \tag{4}$$

TP: True Positive

FP: False Positive

FN: False Negative

$$\text{Mean IOU} = 1/C \sum_{i=1}^{C} IOU_i \tag{5}$$

C: Number of Classes

$$\text{Accuracy} = (TP + TN)/(TP + TN + FP + FN) \tag{6}$$

TN: True Negative

Results

Training Results

The model training was planned for 200 epochs but terminated early due to the early stopping mechanism. As shown in Figure 3, during the training process, the model initially showed rapid learning, followed by a gradually decreasing learning rate strategy. Starting at 5e-4, the learning rate was reduced to 1e-4 at the 60th epoch, 9e-6 at the 100th epoch, and eventually decreased to 1e-6 in the final epochs.

The best model performance was achieved at the 172nd epoch, with metrics presented in Table 4: training accuracy of 0.7440, validation accuracy of 0.7100, training mean IoU of 0.5441, validation mean IoU of 0.4928, training loss of 0.7579, and validation loss of 0.8467. The gap observed between the training and validation metrics indicates some degree of overfitting. Although the model demonstrated acceptable performance for the 15-class CORINE level 2 land cover classification, additional regularization techniques or enhanced data augmentation strategies could be applied to reduce overfitting.

Table 4. Best Performance Metrics

Metric	Value
Training Loss	0.7579
Validation Loss	0.8467
Training Accuracy	0.7440
Validation Accuracy	0.7100
Training IoU	0.5441
Validation IoU	0.4928
Best IOU Performance (Epoch 172)	

Test Results

Test results are shared in Table 5. According to these results, the Marine Waters class achieved the highest IoU value of 0.8884 with a class accuracy of 0.9291, followed closely by Inland Waters with an IoU of 0.8113 and accuracy of 0.8791. The class with the lowest performance is Artificial Non-Agricultural Vegetated Areas, with an IoU value of 0.0896 and class accuracy of 0.1017, followed by Heterogeneous Agricultural Areas with an IoU of 0.2347 and accuracy of 0.3269. Among the other classes, Urban Fabric performed well with an IoU of 0.7308 and class accuracy of 0.8506, followed by Permanent Crops (IoU: 0.7118, Accuracy: 0.8389), and Arable Land (IoU: 0.6456, Accuracy: 0.8189). The Mean IoU, which is the simple average of all class IoU values, was calculated by summing the IoU values and dividing by the number of classes, resulting in 0.5225. Most classes showed notably higher class accuracy compared to their IoU values, particularly evident in vegetation classes such as Forest (IoU: 0.5921, Accuracy: 0.7893) and Open Spaces with Little or No Vegetation (IoU: 0.4406, Accuracy: 0.6146), indicating the model's better performance in identifying class presence while struggling with precise spatial delineation.

Table 5. Test Data Class Based Performance

Class	IoU	Accuracy
Urban fabric	0,7308	0,8506
Industrial, comercial and transport units	0,5647	0,6881
Mine, dump and construction sites	0,3391	0,523
Artificial, non-agricultural vegetated areas	0,0896	0,1017
Arable land	0,6456	0,8189
Permanent crops	0,7118	0,8389
Pastures	0,2549	0,3504
Heterogeneous agricultural areas	0,2347	0,3269
Forest	0,5921	0,7893
Shrub and/or herbaceous vegetation associations	0,3836	0,5792
Open spaces with little or no vegetation	0,4406	0,6146
Inland wetlands	0,5566	0,7808
Coastal wetlands	0,5933	0,7634
Inland waters	0,8113	0,8791
Marine waters	0,8884	0,9291

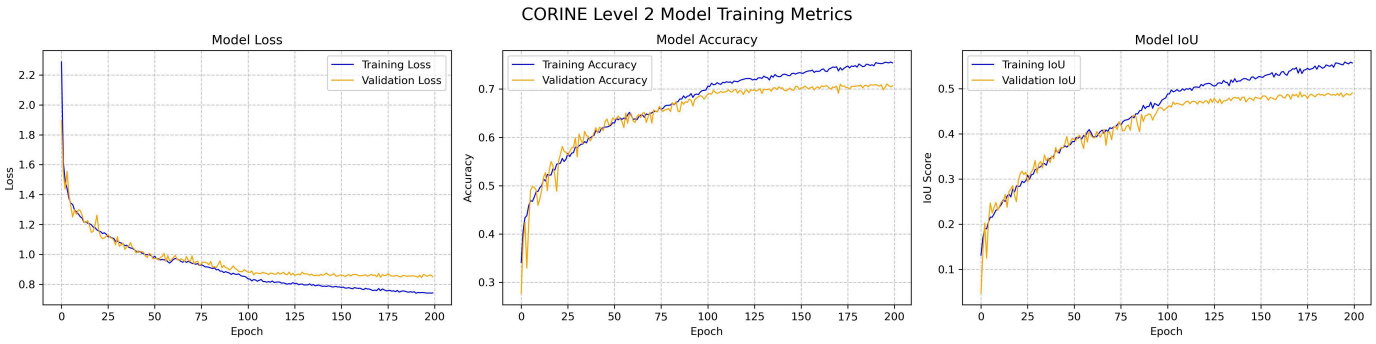


Figure 3. Training and Validation Metrics Over Epochs

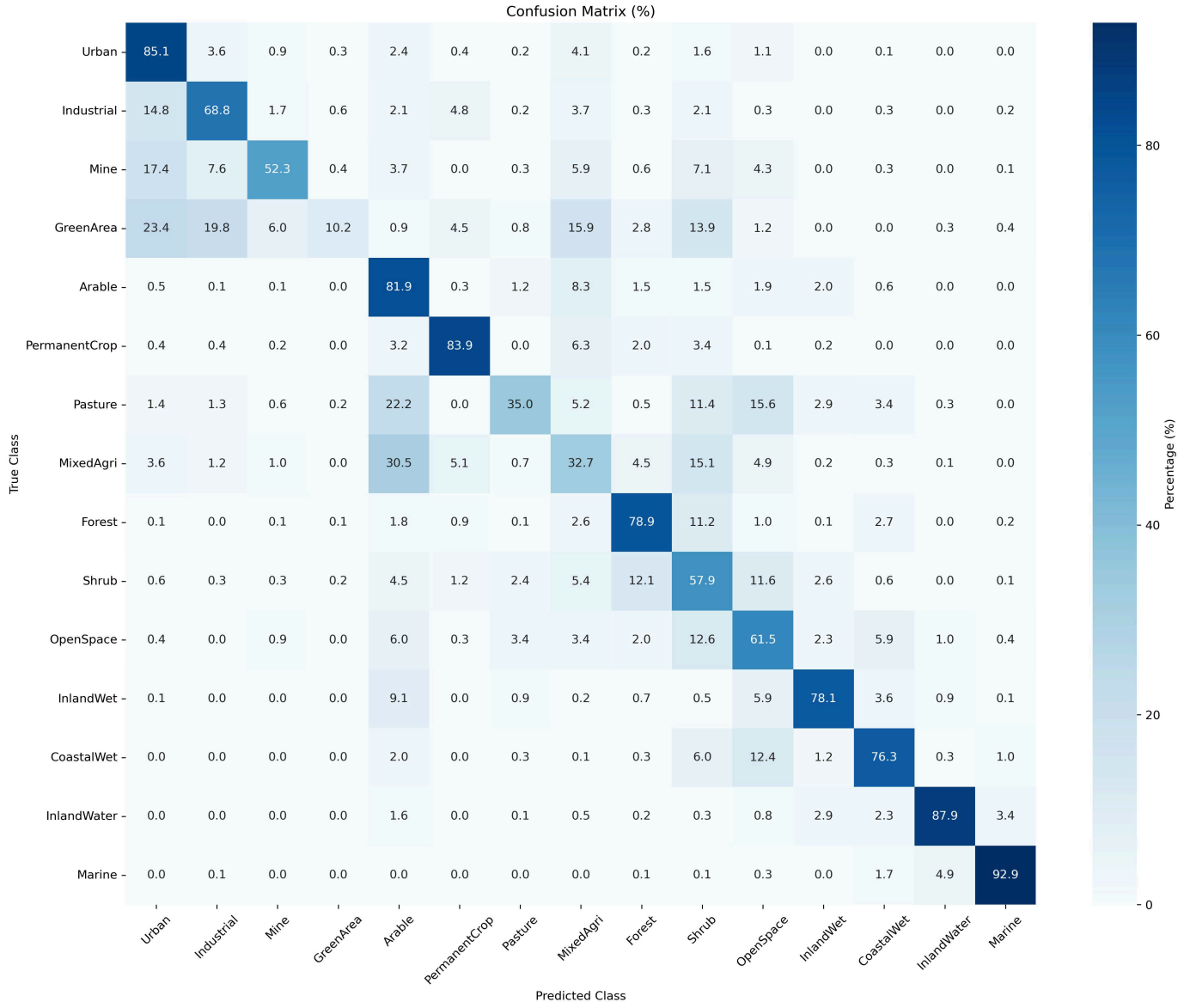


Figure 4. Test Confusion Matrix

Confusion Matrix

A confusion matrix is a fundamental evaluation tool in classification tasks, offering a comprehensive overview of a model's performance by illustrating the frequency with which each class is accurately predicted or misclassified as other classes. This detailed representation enables a nuanced understanding of the model's successes and the specific error patterns it exhibits across different classes. The Confusion matrix reveals significant insights into the model's classification performance across different land cover classes. As shown in Figure 4, Water bodies demonstrate the strongest performance, with Marine Waters achieving an accuracy of 92.9% and showing minimal confusion with other classes except for a slight confusion with Inland Waters (4.9%). Similarly, Inland Waters performs excellently with 87.9% accuracy, primarily showing minor confusion with Marine Waters (3.4%). Among terrestrial classes, Urban Fabric and Permanent Crops show robust performance with accuracies of 85.1% and 83.9% respectively, though there is notable mutual confusion with other classes. The model struggles most with Artificial Non-Agricultural Vegetated Areas and Pastures, achieving accuracies of only 10.2% and 35.0% respectively. Pastures show significant confusion with multiple classes, including Arable Land (22.2%) and Open Spaces (15.6%). Industrial areas achieve a moderate accuracy of 68.8%, with notable confusion with Urban areas (14.8%). The wetland classes show varying performance, with Inland Wetlands achieving 78.1% accuracy and Coastal Wetlands 76.3%, showing some mutual confusion and overlap with Open Spaces. This pattern suggests that the model performs best with

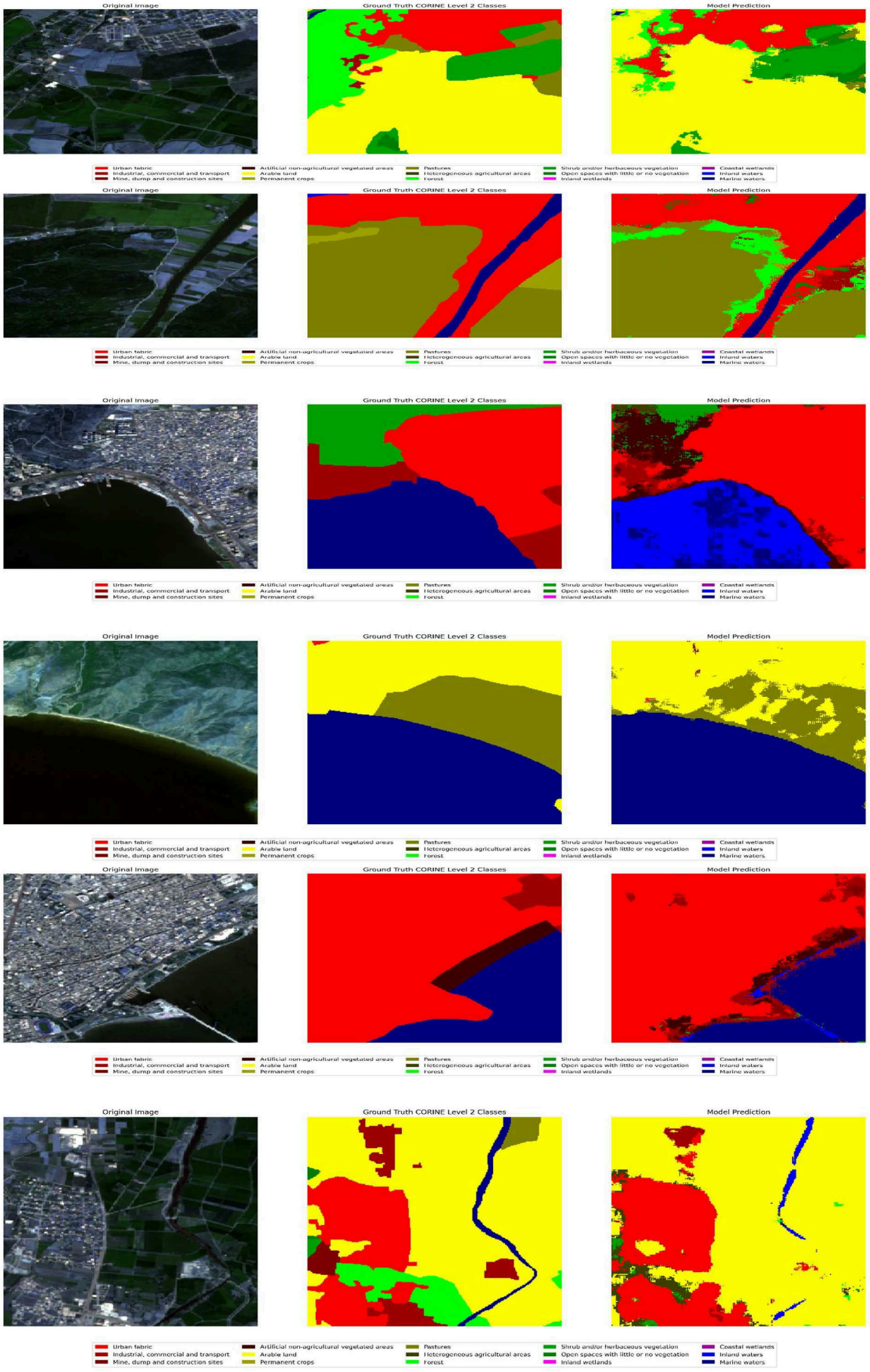


Figure 5. Visual Examples of Model Predictions from the Test Dataset

spectrally distinct classes like water bodies but struggles with classes that have similar spectral and textural characteristics, particularly in mixed vegetation and artificial surfaces.

Test Predictions

To visually assess the model's performance, we present a set of representative examples from the test dataset. Figure 5 shows the comparison between ground truth labels and model predictions for selected test regions. Each row displays the original RGB image (left), ground truth labels (middle), and model predictions (right). These examples were carefully selected to demonstrate both the model's capabilities and limitations across different land cover classes and challenging scenarios. The visual results complement the quantitative metrics presented earlier and provide insights into the model's behavior in real-world classification scenarios.

Discussion and Conclusion

This study delves into the complex realm of automated land cover classification, employing an advanced U-Net deep learning architecture to address the challenges of CORINE Level 2 land cover mapping. Our research transcends simple classification, aiming to demonstrate the potential of semantic segmentation techniques using multi-spectral Sentinel-2 imagery. By developing a sophisticated approach that integrates spatial attention mechanisms and multi-scale feature extraction, we sought to push the boundaries of automated land cover mapping beyond traditional methodological limitations. The results reveal a nuanced landscape of machine learning capabilities in environmental monitoring. Our model demonstrated exceptional performance in spectrally distinct classes, achieving impressive Intersection over Union (IoU) values for water bodies like Marine Waters (0.8884) and Inland Waters (0.8113), as well as urban fabric (0.7308). However, the study also illuminated significant challenges in distinguishing between spectrally similar land cover classes, with particularly low performance in Artificial Non-Agricultural Vegetated Areas (IoU: 0.0896) and Heterogeneous Agricultural Areas (IoU: 0.2347). These variations underscore the complexity of automated land cover classification and highlight the critical need for more sophisticated approaches. When directly comparing our results with existing literature, notable insights emerge. Demir & Musaoglu (2023) achieved an IoU of 0.73 for Urban Fabric and 0.99 for Sea Areas, while our study obtained IoU values of 0.7308 for Urban Fabric and 0.8884 for Marine Waters. Other studies in the field presented varied approaches: Balado et al. (2018) used LIDAR data with 73.5% accuracy, Varga et al. (2021) reached approximately 70% accuracy with multiple satellite images, and Stara & Halounova (2022) achieved 96-99% performance for specific classes using Sentinel-2 imagery and Random Forest algorithms. The research contributes meaningfully to the field of environmental monitoring and remote sensing by showcasing the potential of advanced deep learning techniques in land cover mapping. Our approach goes beyond mere technical implementation, offering insights into the strengths and limitations of semantic segmentation models when applied to complex environmental data. The methodology demonstrates how sophisticated machine learning architectures can be leveraged to address critical challenges in land use and land cover monitoring, which is increasingly important in the context of climate change, urbanization, and environmental conservation. Looking forward, this study points to several promising avenues for future research. The variations in model performance across different land cover classes suggest the need for more advanced feature extraction techniques, potentially incorporating additional data sources or exploring hybrid classification approaches. The challenges encountered with spectrally similar classes indicate that future work should focus on developing more nuanced machine learning models that can better distinguish subtle differences in land cover characteristics. Moreover, the research underscores the importance of continued innovation in remote sensing and machine learning techniques to support more accurate and timely environmental monitoring. Ultimately, this study represents a significant step towards more automated, efficient, and precise land cover mapping. By comprehensively addressing the challenges of CORINE classification through multi-band Sentinel-2 imagery and an advanced deep learning architecture, we provide crucial insights for future research. While the results demonstrate both the potential and the current limitations of deep learning approaches, they also provide a clear roadmap for future advancements in the field. The ability to rapidly and accurately classify land cover has profound implications for environmental management, urban planning, and climate change research, making continued research in this area both necessary and promising.

Peer Review: Externally peer-reviewed.
Author Contributions: Conception/Design of Study: E.A.P., M.Ç.; Data Acquisition- E.A.P., M.Ç, E.B., Ç.İ.; Data Analysis/Interpretation E.A.P., M.Ç, E.B., Ç.İ.; Drafting Manuscript- E.A.P.; Critical Revision of Manuscript- M.Ç.; Final Approval and Accountability- E.A.P., M.Ç., E.B., İ.Ç., M.E.U.
Conflict of Interest: Authors declared no conflict of interest.
Financial Disclosure: Authors declared no financial support.

Acknowledgements The authors would like to thank Amazon Web Services (AWS) for providing access to Amazon SageMaker, a comprehensive machine learning platform that significantly enhanced our research capabilities through its robust computational resources, streamlined workflow management, and advanced model training capabilities (Vinay, 2024). We also extend our gratitude to the employees of Deloitte Türkiye for their valuable support throughout this study, and to Bilge Miraç Atıcı for his guidance on Amazon SageMaker implementation. We would especially like to express our gratitude to Yasin Hakkı Kocaman, Deputy General Director of Information Technologies at the Ministry of Agriculture and Forestry, for his support throughout all processes.

The implementation of the model architecture and all code used in this study is publicly available at: https://github.com/eapeker/corine-unet-model

Author Details

Emin Atabey Peker
[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye
0000-0002-8014-7250 eminatabeypeker@gmail.com

Murat Çalışkan
[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye
0009-0001-8359-1904

Eser Bora
[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye
0000-0002-4228-1491

Çiğdem İnan
[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye
0009-0006-9895-6167

M. Erkan Uçaner
[1] Ministry of Agriculture and Forestry, Department of Information Technologies, Ankara, Türkiye
0009-0005-2388-3728

REFERENCES

Balado, J., Arias, P., Vilariño, L. D., & González-deSantos, L. M. (2018). Automatic CORINE land cover classification from airborne LIDAR data. *Procedia Computer Science*, 186-194.

Balzter, H., Cole, B., Thiel, C., & Schmullius, C. (2015). Mapping CORINE Land Cover from Sentinel-1A SAR and SRTM Digital Elevation Model Data using Random Forests. *Remote Sensing*, 14876-14898.

Copernicus Land Monitoring Service (CLMS). (2025). *Corine land cover*. Retrieved from https://land.copernicus.eu/en/products/corine-land-cover

Demir, D. B., & Musaoglu, N. (2023). Automatic Classification Of Selected Corine Classes Using Deep Learning Based Semantic Segmentation. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Volume XLVIII-M-3-2023). Denver, Colorado: ASPRS.

Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., & Moore, R. (2017). Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sensing of Environment, 202*(3), 18-27.

Johnson, K. E., & Koperski, K. (2017). WorldView-3 SWIR Land Use-Land Cover Mineral Classification: Cuprite, Nevada. In *Pecora 20 - Observing a Changing Earth; Science for Decisions— Monitoring, Assessment, and Projection*. Sioux Falls, SD.

Loshchilov, I., & Hutter, F. (2017). Decoupled Weight Decay Regularization. In *Proceedings of the International Conference on Learning Representations*. https://doi.org/10.48550/arXiv.1711.05101

Macarringue, L. S., Bolfe, É. L., & Pereira, P. R. (2022). Developments in Land Use and Land Cover Classification Techniques in Remote Sensing: A Review. *Journal of Geographic Information System, 14*, 1-28.

Özturk, O., Sarıtürk, B., & Seker, D. Z. (2020). Comparison of fully convolutional networks (FCN) and U-Net for road segmentation from high resolution imageries. *International journal of environment and Geoinformatics*, 7(3), 272-279.

Ronneberger, O., Fischer, P., & Brox, T. (2015). U-Net: Convolutional Networks for Biomedical Image Segmentation. In N. Navab, J. Hornegger, W. Wells, & A. Frangi (Eds.), *Medical Image Computing and Computer-Assisted Intervention* (pp. 234-241). Springer.

Solórzano, J. V., Mas, J. F., Gao, Y., & Cruz, J. G. (2021). Land Use Land Cover Classification with U-Net: Advantages of Combining Sentinel-1 and Sentinel-2 Imagery. *Remote Sensing, 13*(18), 3600.

Stara, L., & Halounova, L. (2022). Classification Of Selected Corine Classes Using Sentinel-2 Data. In *The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences* (Volume XLIII-B3). Nice, France: XXIV ISPRS Congress.

U.S. Environmental Protection Agency (EPA). (2024). *Report on the Environment (ROE): What are the trends in land cover and their effects on human health and the environment?*

Varga, O. G., Kovács, Z., Bekő, L., Burai, P., Szabó, Z. C., Holb, I., & Szabó, S. (2021). Validation of Visually Interpreted Corine Land Cover Classes with Spectral Values of Satellite Images and Machine Learning. *Remote Sensing, 13*(5), 857.

Vinay, S. B. (2024). AI and Machine Learning Integration with Aws Sagemaker: Current Trends and Future Prospects. *International Journal of Artificial Intelligence Tools (IJAIT)*, 1-24.

Zhang, C., & Li, X. (2022). Land Use and Land Cover Mapping in the Era of Big Data. *Land, 11*, 1692.

Zhao, S., Tu, K., Ye, S., Tang, H., Hu, Y., & Xie, C. (2023). Land Use and Land Cover Classification Meets Deep Learning: A Review. *Sensors, 23*(21), 8966.