Semantic Segmentation Using Deep Learning: Insights from the LICAID Dataset

Girma Tariku[1,*,†], Isabella Ghiglieno[2], Anna Simonetto[2], Gianni Gilioli[2] and Ivan Serina[1]

[1] University of Brescia – Department of Information Engineering (DII), –via Branze 38, 25123, Brescia, Italy, e-mail: g.tariku@unibs.it

[2] Agrofood Research Hub, University of Brescia – Department of Civil, Environmental, Architectural Engineering, and Mathematics – via Branze 43, 25123, Brescia, Italy

Abstract

Land cover mapping is critical for monitoring global land use patterns, assessing ecosystem health, and supporting conservation efforts. However, challenges persist in handling large satellite imagery datasets and acquiring specialized aerial datasets for deep-learning models. To address these challenges, this study introduces a methodology for semantic segmentation of land cover in agricultural regions, specifically tailored to the wine-growing region of Franciacorta, Italy. We present the "Land Cover Aerial Imagery" (LICAID) dataset and employ the advanced deep learning model DeepLabV3 with various pre-trained backbones (ResNet, DenseNet, and EfficientNet) for comparative analysis. The dataset comprises eleven land cover classes: grasslands, arable land, herb-dominated habitats, hedgerows, vineyards, tree-dominated man-made habitats, olive groves, wetlands, lines of planted trees, small anthropogenic forests, and others. Results demonstrate significant performance improvements in land cover classification using deep learning with pre-trained networks, providing a scalable and cost-effective approach to land cover mapping that supports environmental monitoring and conservation.

Keywords

Land cover mapping; Semantic segmentation; Deep learning, Satellite imagery, pre-trained backbone

1. Introduction

Accurate and timely land cover mapping is essential for understanding the complex interactions of land use patterns and their profound impact on global ecosystems [2]. This is particularly important in agricultural regions, where land management practices directly influence biodiversity, ecosystem services, and the overall sustainability of food production systems [3, 4, 5, 6, 7]. Inaccurate or outdated land cover data can have widespread impacts, affecting policy decisions related to resource allocation, environmental protection, and agricultural planning. Effective land management strategies rely on accurate land cover assessments to maximize resource utilization, minimize environmental risks, and promote sustainable agricultural practices. For example, accurate identification of various crop types enables targeted interventions to improve yields and reduce the need for chemical inputs. Similarly, accurate mapping of natural habitats within agricultural landscapes is crucial for biodiversity conservation and the maintenance of ecosystem services such as pollination and pest control.

Traditional land cover classification methods, such as support vector machines and random forests [8, 9, 10, 11, 12, 13], often struggle to effectively process the high-resolution images readily available from modern remote sensing platforms. These methods face challenges in handling the

large datasets generated by modern sensors, leading to computational limitations and increased processing times [14, 15, 16, 17]. Furthermore, the inherent complexity of agricultural landscapes characterized by significant spatial diversity and subtle spectral variations between land types often compromises the accuracy and reliability of traditional classification approaches. The limitations are particularly evident in areas with complex land use patterns and overlapping spectral signatures, resulting in misclassifications and inaccurate maps. This highlights the need for more advanced technologies capable of efficiently processing high-resolution images and large-scale datasets while maintaining high accuracy.

Recent advances in deep learning offer significant potential for improving the accuracy and efficiency of land cover mapping [18, 19, 20, 21, 22, 27, 30]. However, a major bottleneck remains: the scarcity of readily available, high-quality, and cost-effective datasets tailored to specific agricultural contexts [23, 24, 25, 31, 32, 33]. The development of robust and accurate deep learning models for land cover classification relies heavily on large, well-annotated datasets that accurately reflect the diversity and complexity of the target environment. The high costs and time investment associated with creating such datasets often hinder research and limit the widespread adoption of deep learning in agricultural applications. This data shortage has impeded progress in accurate agricultural practices and sustainable land management, underscoring the urgent need for a readily accessible and representative dataset.

This paper presents an approach to address the problem of data scarcity in land cover classification. We introduce a new semantic segmentation dataset, Land Cover Aerial Imagery (LICAID), specifically designed for the wine-growing region. LICAID includes eleven land cover classes representing the diverse landscapes surrounding vineyards: pastures, arable land, herb-dominated habitats, hedgerows, vineyards, tree-dominated man-made habitats, olive groves, wetlands, planted trees, small anthropogenic forests, and other. This detailed classification allows for a more precise understanding of the complex interactions within the vineyard ecosystem and its surrounding environment.

Our methodology focuses on cost-effective data acquisition and processing, making it a replicable approach for other agricultural regions. We evaluate the performance of the state-of-the-art deep learning model DeepLabV3, utilizing ResNet, DenseNet, and EfficientNet backbones, on the LICAID dataset. This comparative analysis demonstrates the potential of our approach to enhance the accuracy and efficiency of land cover mapping in agricultural settings, ultimately supporting sustainable land management and environmental monitoring efforts.

2. Materials and Methods

This study uses the deeplabV3 model to use four-step semantic segmentation methods in the wine region of Franciacorta wine region. The process includes remote sensing data collection, rigorous image preprocessing, expert validation, training and evaluation of DeepLabV3 models. The focus on DeepLabV3 allows for a deeper exploration of its implementation and optimization in this specific application context [34].

2.1. Study Area

The study focuses on the famous Italian wine-growing region of Franciacorta in Lombardy (Figure 1). Located in the picturesque province of Brescia, Franciacorta is renowned for its stunning landscapes, rich history, and world-class wine production.
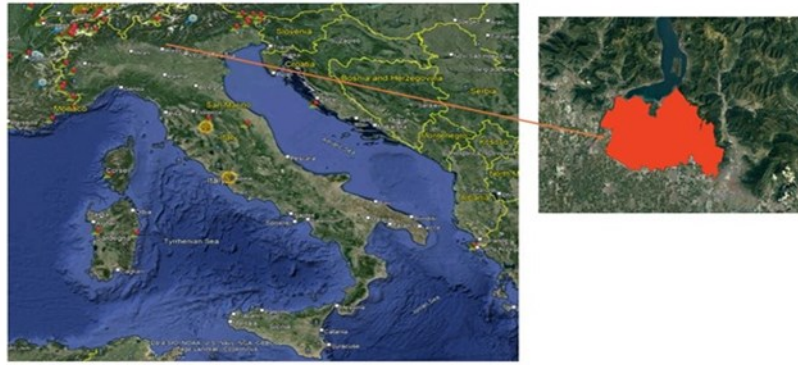
Figure 1. Geographical context of the Franciacorta study area in the Lombardy region of Italy. Left: Overview shows the position of the map of the Lombardy region in Italy. Right: Close-up view showing the detailed perspective of the Franciacorta research area.

2.2. Image Preprocessing, Segmentation, and Expert Validation

Satellite imagery for the study was acquired through Google Earth Pro, which provided high-resolution imagery of the area. The images were downloaded in .kmz format, which includes both the visual data and spatial attributes, forming the foundation for further analysis. Following acquisition, ArcGIS was used for georeferencing each image to ensure alignment with the geographical coordinates of the region, enabling precise spatial analyses. The study area's heterogeneous landscape posed a challenge for segmentation, as different cover types required fine-tuned segmentation techniques.

1. Image Tile Segmentation
   Multiresolution segmentation (MRS) was performed on the georeferenced images using eCognition software. This approach divides the imagery into smaller, homogeneous regions based on spectral and spatial characteristics, optimizing the image for more accurate classification. Key segmentation parameters included:

   - Scale parameter (100): Balances segment size to reflect distinct regions in the landscape.
   - Compactness (0.5): Adjusts the segment shape for clarity without excessive fragmentation.
   - Shape (0.4): Ensures the segments preserve geographic coherence while isolating major landscape features.

2. Expert Validation and Classification
   The segmented images were subsequently validated and classified by a plant expert using QGIS software. Each polygon within the shapefile was analyzed and assigned one of several classes, such as grasslands, vineyards, arboreal land, and olive groves. This validation ensured that each segment's classification was botanically accurate, allowing for reliable LULC representation.

The expert identified eleven cover types essential to Franciacorta's biodiversity and agricultural ecosystem:

1. Grassland: Supports biodiversity through pollination and pest control [3].
2. Arable Land: Cultivated land used for crop production.
3. Herb-Dominated Habitats: Supports herbivores, pollinators, and decomposers, contributing to soil health.

4. Hedgerows: Enhance ecosystem connectivity and provide habitat for wildlife.
5. Vineyards: The primary land cover type, supporting specialized flora and fauna.
6. Tree-Dominated Man-Made Habitats: Provide habitat in human-modified landscapes.
7. Olive Groves (Olea europaea): Support specialized flora and fauna in Mediterranean environments.
8. Wetlands: Areas characterized by saturated soil conditions.
9. Lines of Planted Trees: Linear plantings used for windbreaks, etc.
10. Small Anthropogenic Forests: Small, human-planted forest patches.
11. Other: A residual class for miscellaneous land cover types.

3. Patch Extraction and Dataset Preparation

Given the size of the original image pixels, preprocessing is required to optimize deep-learning model training. A Python script was developed to divide large images into smaller, manageable regions while maintaining spatial coherence between each region and its corresponding class labels. As illustrated in Figure 2, semantic segmentation was performed in three steps: 1) Images and corresponding masks were acquired from QGIS software; 2) Large training images and masks were divided into 128x128 pixel segments for training the DeepLabv3 semantic segmentation model; and 3) The trained DeepLabv3 model (saved for future use) was then used to predict and map the image.
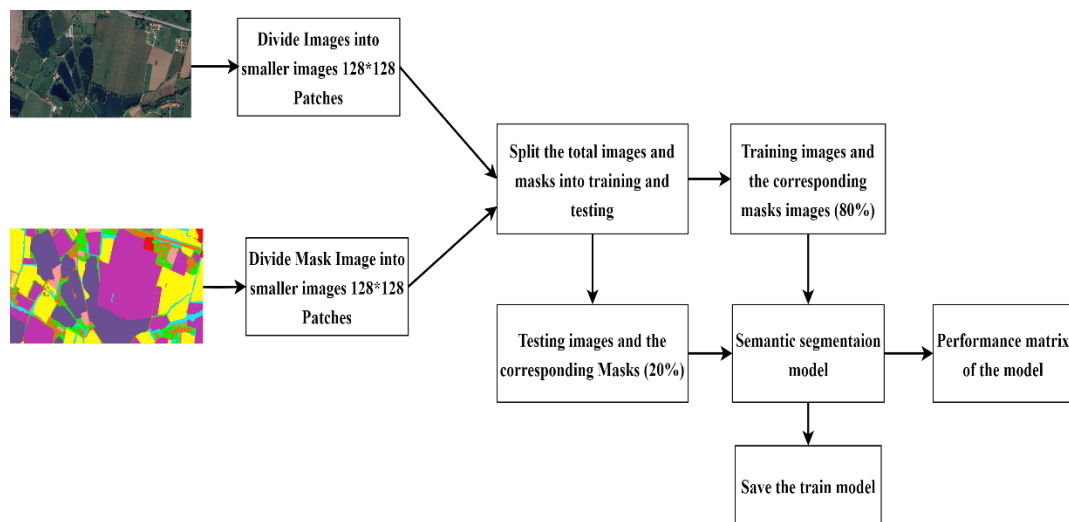

Figure 2. Illustration of the general flow of semantic segmentation images.

2.3. Image and Mask Processing

A Python script processed the image and mask files (Figure 3), resizing them to the nearest dimension divisible by the chosen patch size. This ensured the creation of non-overlapping patches using the patchy library, resulting in 921 training image-mask pairs and 230 validation image-mask pairs. This preprocessing step was crucial for efficient model training and prevented artifacts that could arise from overlapping patches.

The non-overlapping patch strategy, facilitated by the patchy library, provided a clean and consistent dataset for training the DeepLabv3 model. The resulting datasets 921 training and 230 validation image-mask pairs were balanced and suitable for effective model training and subsequent performance evaluation. This careful dataset preparation was a key factor in achieving the high accuracy reported in the results.
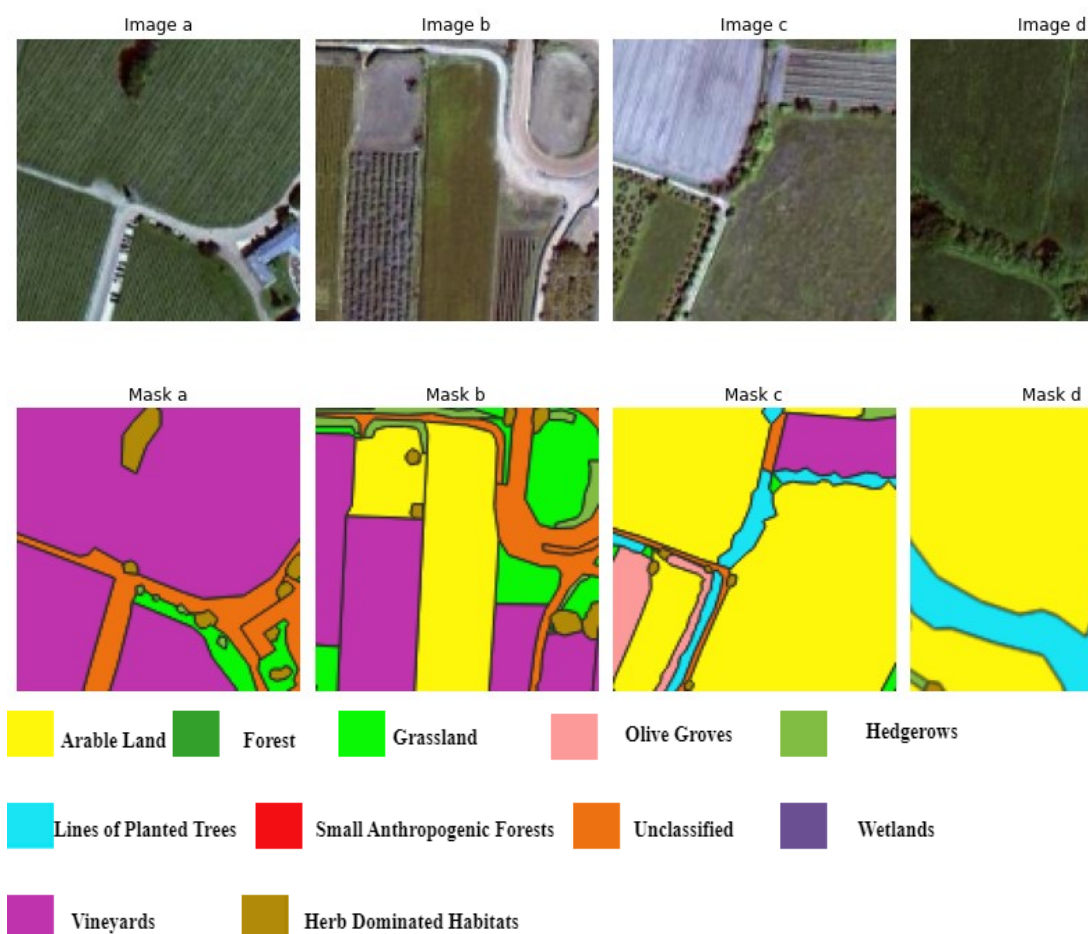
Figure 3. Illustration of four randomly selected image patches (A, B, C, D) of size 128x128 pixels, extracted from the original images along with their corresponding ground truth masks. These patches are used to train and evaluate a deep-learning model.

2.4. Data Split and Organization

Using TensorFlow's split folders library, the patches were split into training and validation datasets with an 80-20 split ratio. These splits facilitated model training and performance evaluation, allowing for robust testing across both seen and unseen data.

2.5. Semantic Segmentation Models and Backbone Architecture

The DeepLabv3 model was implemented using TensorFlow/Keras. The training involved minimizing loss functions (measuring the difference between predicted and ground truth masks) and evaluating performance using IoU (Intersection over Union), accuracy, and mean IoU. We compared semantic segmentation models with and without backbone architectures. The "no backbone" approach trained DeepLabv3 from scratch, relying solely on its intrinsic architecture for feature extraction. This contrasts with models incorporating four commonly used backbone architectures: ResNet-34 [35], a deep convolutional network with residual connections to mitigate vanishing gradients; InceptionV3 [36], a computationally efficient architecture using inception modules for multi-scale feature extraction; EfficientNet [37], distinguished by its compound scaling method for optimal resource utilization; and DenseNet [38], characterized by dense connections promoting feature reuse and efficient gradient propagation.

DeepLabV3[34] is an advanced convolutionary neural network architecture designed for semantic image segmentation. It uses the atrous convolution to efficiently capture multi-scale context information. The key features include the feature pyramid network for hierarchical feature extraction, the spatial pyramid pool for a multi-scale feature aggregation, and the efficient higher-

level sampling method for high-resolution segmentation maps. Our DeepLabV3+ model, shown in Figure 4, uses the Effi-cientNetB0 framework for feature extraction without fully connected layers as an encoder. The encoder processes the input image, extracts the functions, and transmits them to the decoder. The decoder consists of an Atrous Spatial Pyramid Pooling (ASPP) module followed by global average pooling and low-level function concatenation. The ASPP module captures multi-scale context information by using a convolutional layer with different dispersion rates. In addition, global average pooling is performed to capture global context information. Then the decoder combines ASPP output, global context, and low-level functionality through concatenation. Further convolutional layers refine the characteristics and then increase the sampling layers to restore spatial information. Finally, the 1x1 convolutional layer activated by SoftMax generates pixel predictions for semantic segmentation. The model was composed by Adam's optimizer, using categorical cross-entropy losses and accuracy as measures of evaluation.
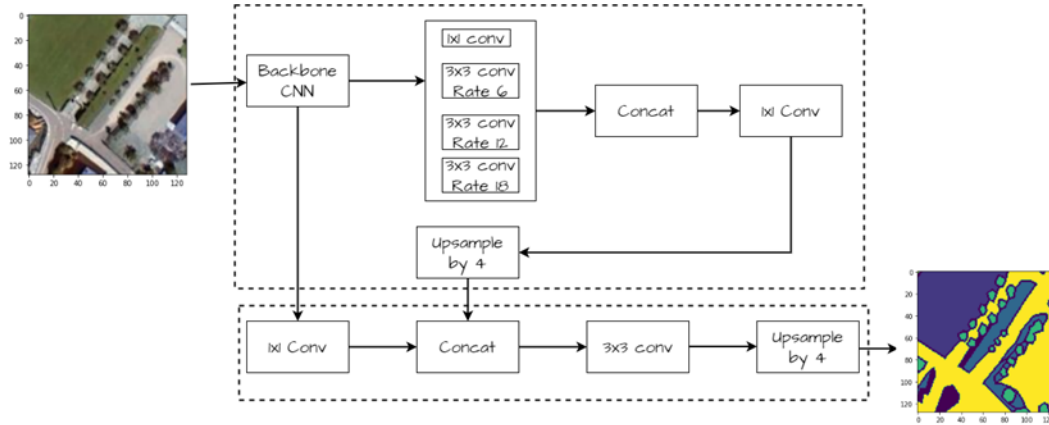

Figure 4. Block diagram of DeeplabV3 with backbone CNN as the encoder and decoder.

3. Results and Discussion

Our DeeplabV3 semantic segmentation model was trained on a dataset comprising 1152 image patches. These patches were derived from 15 satellite images by dividing them into smaller 128x128 pixel sections. To ensure rigorous evaluation, the dataset was divided into a training set (80%), and a test set (20%).

The training process began by loading and preprocessing the image-mask pairs from specified directories. Each image was annotated with one of eleven land cover classes, and the corresponding masks underwent label encoding before being split into training and testing datasets (80/20 split). The DeepLabv3 model was then configured, experimenting with four different backbone architectures (DenseNet121, ResNet34, InceptionV3, and EfficientNet) and employing a combined Dice and Categorical Focal loss function for optimization. Training proceeded for 100 epochs, with progress visualized through plots of training loss and Intersection over Union (IoU) scores. Upon completion, the trained model was saved.

Following training, the model's performance was rigorously evaluated on the held-out test dataset. A comprehensive suite of performance metrics, including accuracy, precision, recall, F1-score, Jaccard index, and mean IoU, were calculated and reported to provide a thorough assessment of the model's ability to accurately segment the eleven land cover classes. These results provided a quantitative measure of the model's effectiveness and informed the subsequent analysis and discussion.

Our methodology proves to be both scalable and cost-effective, with potential applications in other agricultural regions. The comparative analysis highlights the importance of model selection and backbone configuration in optimizing performance for specific land cover mapping tasks.

To address the class imbalance problem, where certain land cover classes were significantly under-represented in the dataset compared to others, a class-balanced data augmentation strategy was implemented. This involved selectively enhancing the training data by generating new image patches that focused specifically on the under-represented classes.

This augmentation technique modified existing masks to temporarily exclude the over-represented classes, effectively creating new masks where the under-represented classes became the dominant features. New image patches were then generated from these modified masks, increasing the number of training samples for the under-represented classes without altering the original dataset. This targeted approach ensured that the model received sufficient training examples for all classes, improving its ability to accurately segment even the less frequent land cover types.

In order to reduce over-fitting, we implemented an early stop. Training stopped when validation losses stopped declining for a certain period (e.g., 10 years), so that training continued only when the model improved its performance in invisible data. In addition, batch normalization has been used to stabilize and accelerate training, and loss-function class weight has helped to control class imbalances. These strategies collectively contributed to preventing over-adaptation and improving the generalization capacity of models.

The comparison of semantic segmentation models with different backbone architectures reveals nuanced differences in performance metrics (Table 1). Across all models, including ResNet34, EfficientNet, InceptionV3, DenseNet, and a model without a specified backbone, there is a notable consistency in accuracy, precision, recall, and F1 score, with variations typically within a range of 1-2 percentage points. However, when assessing metrics more tailored to semantic segmentation tasks, such as mean IOU and Jaccard score, subtle disparities emerge. EfficientNet and DenseNet exhibit slightly higher mean IOU and Jaccard scores compared to ResNet34 and InceptionV3, highlighting their marginally superior ability to accurately segment objects in images. For instance, EfficientNet achieves an accuracy of 85.6%, while DenseNet reaches 84.2%, both with corresponding mean IOU scores of 59.0% and 56.2%, respectively. These results underscore the importance of selecting an appropriate backbone architecture, as models with dedicated architectures designed for image segmentation tasks demonstrate enhanced performance, particularly in terms of mean IOU and Jaccard score, compared to models without a specified backbone.

Table 1

Performance Metrics of DeepLabV3 Semantic Segmentation Models with Different Backbone Architectures and without.

Backbone	Accuracy	Precision	Recall	F1 Score	Mean IOU	Jaccard Score
ResNet	79.0	78.8	78.3	71.1	49.5	62.1
EfficientNet B0	85.61	85.48	85.5	85.28	59.0	75.68
Inception V3	80.6	80.1	80.4	80.5	55.2	69.4
DenseNet	84.2	84.7	84.1	84.2	56.2	70.2
Without Backbone	75.0	75.9	75.8	75.5	49.9	63.2

Accuracy: This metric measures the overall correctness of the segmentation by calculating the ratio of correctly predicted pixels to the total number of pixels.

Precision: Precision quantifies the model's ability to correctly identify positive predictions among all predicted positives. It's calculated as the ratio of true positives to the sum of true positives and false positives.

Recall: Recall, also known as sensitivity, measures the ability of the model to detect all relevant instances of the class in the image. It's calculated as the ratio of true positives to the sum of true positives and false negatives.

F1 Score: The F1 score is the harmonic mean of precision and recall. It provides a balanced measure between precision and recall and is calculated as 2 * (precision * recall) / (precision + recall).

Mean IoU: Mean IoU calculates the average IoU across all classes. It's a popular metric for semantic segmentation tasks as it provides an overall measure of segmentation accuracy across different classes.

Jaccard Score (IoU): The Jaccard score, or Intersection over Union (IoU), measures the ratio of the intersection of the predicted and ground truth segmentation masks to their union. It evaluates the overlap between the predicted and ground truth regions.

Following the training phase, the optimized DeepLabv3 model was loaded for evaluation. A batch of test images and their corresponding ground truth masks were generated using the previously prepared validation data generator. This ensured that the model's performance was assessed on unseen data, providing a more robust and unbiased evaluation of its generalization capabilities. The loaded model then processed this batch of test images, generating predictions in the form of predicted segmentation masks. These predicted masks, initially represented in a categorical format, were subsequently converted to an integer format for easier visualization and compatibility with the Intersection over Union (IoU) calculation. This conversion simplifies the comparison between the predicted and ground truth masks, facilitating both qualitative and quantitative analysis of the model's performance.

To provide a qualitative assessment of the model's performance, a randomly selected test image was chosen for visualization alongside its corresponding ground truth mask and the model's predicted mask. This visual comparison, presented in Figure 5, allowed for a direct assessment of the model's ability to accurately segment the different land cover classes. The visual representation provides valuable insights into the model's strengths and weaknesses, highlighting areas where the model performed exceptionally well and areas where further improvement might be needed. This qualitative assessment complements the quantitative evaluation provided by the calculated performance metrics, offering a more comprehensive understanding of the model's overall performance. Visual comparison helps to identify potential biases or limitations in the model's predictions, informing future model improvements and dataset refinements. The selection of a random image ensures that the visualization is representative of the model's performance across the entire test dataset, avoiding potential bias towards specific image characteristics.
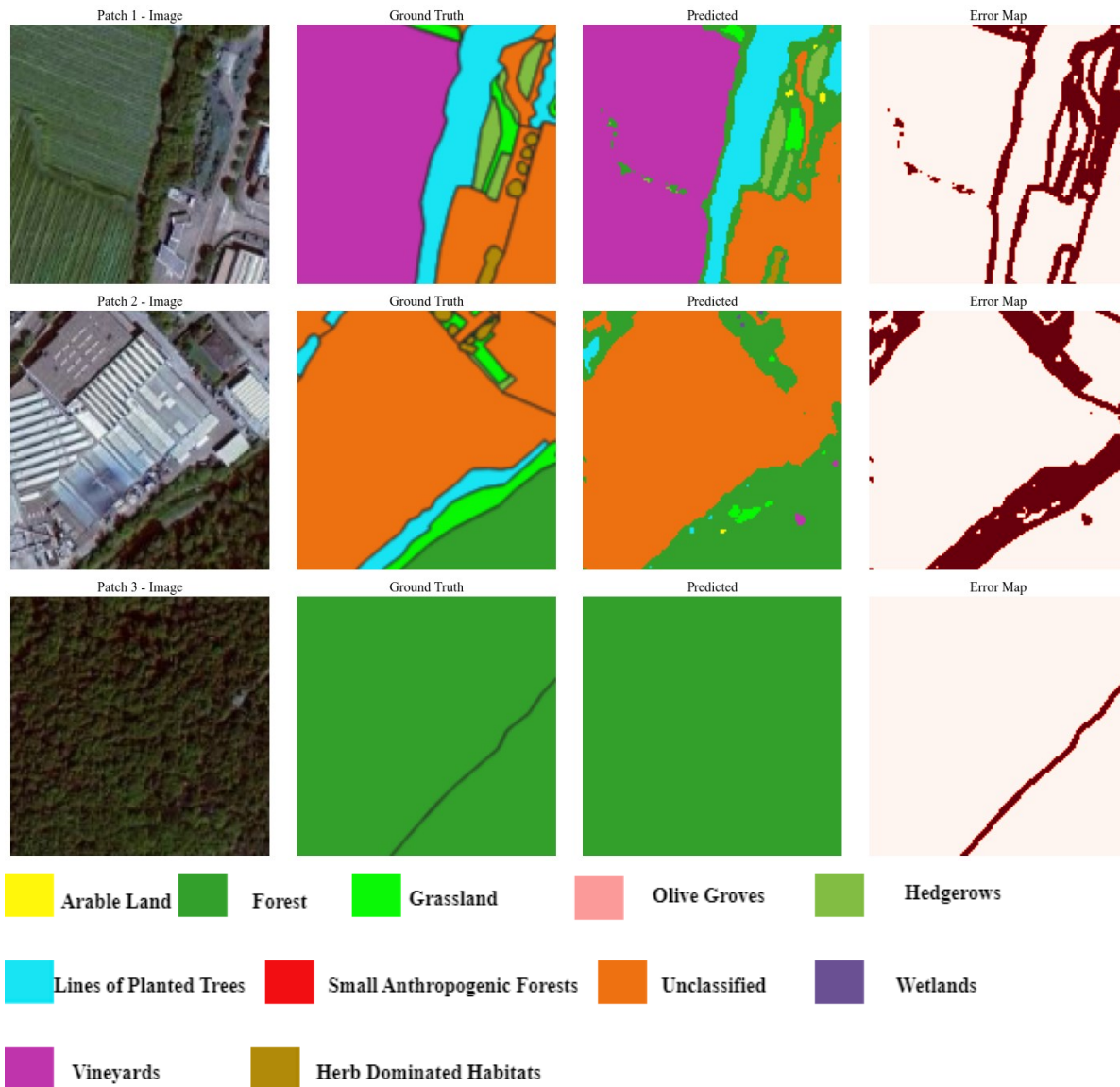
Figure 5. Visualization of a randomly selected test image, ground truth mask, and predicted mask. The figure illustrates the qualitative assessment of the model's performance, comparing the original test image to its corresponding ground truth and predicted masks. The legend describes the color representing each class.

After training, the optimized DeepLabv3 model was loaded to process a significantly larger image than those used during training and validation. Because the model was trained on smaller image patches, this large image was first segmented into a series of overlapping patches of the same size as those used during training. This tiling strategy ensured compatibility with the model's input requirements. The model then processed each patch individually, generating a predicted segmentation mask for each. A crucial step followed: these individual, patch-level predictions were carefully stitched together using an appropriate image stitching algorithm to reconstruct a complete, seamless predicted mask for the entire large input image. This process effectively scaled the model's application to images exceeding the size constraints of the training data, allowing for the analysis of larger scenes and demonstrating the model's ability to generalize to larger-scale applications (as shown in Figure 6). The overlapping patches helped to mitigate boundary artifacts that can sometimes occur during this stitching process, ensuring a more coherent and accurate final segmentation mask.
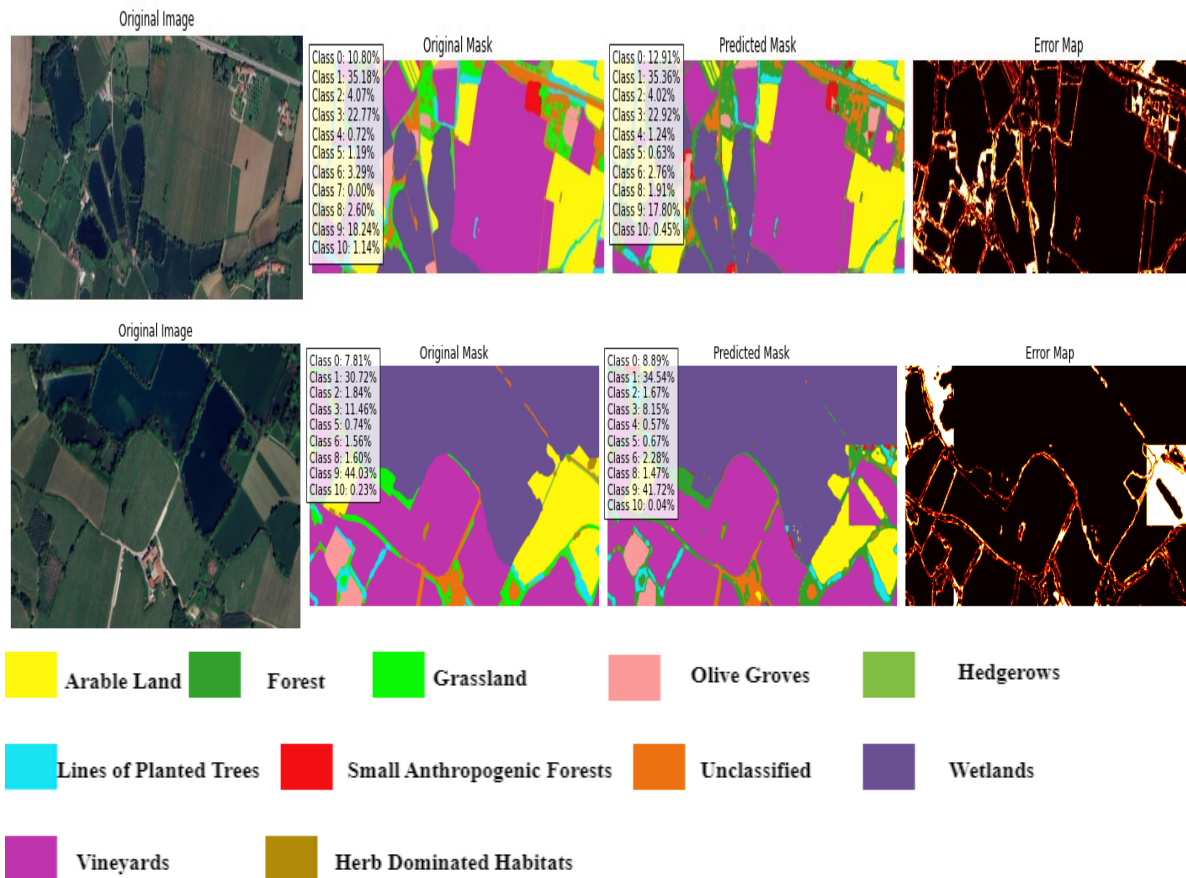
Figure 6. Illustrates the original image, ground truth mask, and the final segmented image mask obtained by applying a trained model. The model segments the large image into patches for processing, generates predictions for each patch, and seamlessly stitches together the resulting segmented patches to reconstruct the predicted mask for the entire large image. The legend describes the color representation of each class.

This study demonstrates land cover mapping in agricultural areas using semantic segmentation with the DeepLabv3 deep learning model and transfer learning. A new dataset of eleven agricultural land cover classes was created, addressing a critical gap in existing research. This dataset facilitates the development of mapping methods for the main land cover classes in Italy's Franciacorta region and provides a valuable resource for future research on biodiversity and sustainability. The study overcomes limitations of traditional methods that struggle with spectral similarities and class heterogeneity. Results show superior performance for DeepLabv3 with an EfficientNetB0 backbone, achieving significantly higher accuracy (0.866) and improved performance across other key metrics. DeepLabv3's advanced architecture, including atrous spatial pyramid pooling (ASPP), facilitates multi-scale context integration for improved detail capture. EfficientNetB0's efficient compound scaling method optimizes the balance between model complexity and accuracy. Future work should include expanding the dataset to incorporate more land cover classes and integrating ground truth data from field observations to further improve accuracy and reliability, particularly in complex agricultural landscapes.

4. Conclusion

This study successfully demonstrates the application of satellite imagery and deep learning, specifically the DeepLabv3 model, for accurate land cover mapping in the Franciacorta wine-growing area. A novel manually annotated dataset comprising eleven land cover classes provides a valuable

resource for agricultural research and future studies. Our results highlight the effectiveness of DeepLabv3 in accurately segmenting these classes from satellite imagery, showcasing the potential of this advanced method for sustainable land management, environmental monitoring, and informed agricultural decision-making. The analysis of diverse land cover types emphasizes the importance of understanding the contribution of surrounding habitats to vineyard sustainability, biodiversity, and ecosystem services. Future work should focus on enhancing the generalizability of the findings through dataset expansion (potentially including additional land cover classes and ground-truth data from field surveys) and exploration of more advanced deep learning architectures to further refine model performance and broaden applicability across diverse agricultural landscapes.

This study successfully demonstrates the application of deep learning, using the DeepLabv3 model, for accurate land cover mapping in the Franciacorta wine-growing region of Italy. Leveraging high-resolution satellite imagery, the study produced detailed land cover maps, providing valuable insights into the spatial distribution of eleven distinct land cover classes within this complex agricultural landscape. The creation of a novel, manually annotated dataset represents a significant contribution, offering a valuable resource for future research in agricultural applications and precision land management.

The results clearly show the effectiveness of DeepLabv3 in accurately segmenting these diverse land cover classes from satellite imagery. This success highlights the potential of advanced deep learning techniques for improving the accuracy and efficiency of land cover mapping, thereby supporting sustainable land management practices, environmental monitoring initiatives, and evidence-based agricultural decision-making. The detailed analysis of the various land cover types underscores the importance of understanding the intricate relationships between vineyard ecosystems and their surrounding habitats, emphasizing the contribution of these surrounding areas to overall vineyard sustainability, biodiversity, and the provision of essential ecosystem services.

Future research should focus on enhancing the generalizability and robustness of the findings. This can be achieved through expanding the dataset to include a wider range of land cover classes, incorporating ground-truth data collected through field surveys to further validate the model's accuracy, and exploring more advanced deep learning architectures or model optimization techniques to improve performance and broaden applicability across diverse agricultural settings. These enhancements will strengthen the model's capability for wider application and contribute to more comprehensive and reliable land cover mapping in various agricultural contexts.

References

[1] D. Aineto, R. De Benedictis, M. Maratea, M. Mittelmann, G. Monaco, E. Scala, L. Serafini, I. Serina, F. Spegni, E. Tosello, A. Umbrico, M. Vallati (Eds.), Proceedings of the International Workshop on Artificial Intelligence for Climate Change, the Italian workshop on Planning and Scheduling, the RCRA Workshop on Experimental evaluation of algorithms for solving problems with combinatorial explosion, and the Workshop on Strategies, Prediction, Interaction, and Reasoning in Italy (AI4CC-IPS-RCRA-SPIRIT 2024), co-located with 23rd International Conference of the Italian Association for Artificial Intelligence (AIxIA 2024), CEUR Workshop Proceedings, CEUR-WS.org, 2024.

[2] Olofsson, P.; Foody, G.M.; Herold, M.; Stehman, S.V.; Woodcock, C.E.; Wulder, M.A. Good practices for estimating area and assessing accuracy of land change. Remote Sens. Environ. 2014, 148, 42–57. https://doi.org/10.1016/j.rse.2014.02.015.

[3] Williams, J.N.; Morandé, J.A.; Vaghti, M.G.; Medellín-Azuara, J.; Viers, J.H. Ecosystem Services in Vineyard Landscapes: A Focus on Aboveground Carbon Storage and Accumulation. Carbon Balance Manag. 2020, 15, 23. https://doi.org/10.1186/s13021-020-00158-z.

[4] Giffard, B.; et al. Vineyard Management and Its Impacts on Soil Biodiversity, Functions, and Ecosystem Services. Front. Ecol. Evol. 2022, 10. https://doi.org/10.3389/fevo.2022.850272.

[5] The Regenerative Viticulture Foundation. Biodiversity. Available online: https://www.regenerativeviticulture.org/toolkit/biodiversity/ (accessed on 9 October 2024).

[6] Abad, J.; de Mendoza, I.H.; Marín, D.; Orcaray, L.; Santesteban, L.G. Cover crops in viticulture. A systematic review (1): Implications on soil characteristics and biodiversity in vineyard. OENO One 2021, 55, 1. https://doi.org/10.20870/oeno-one.2021.55.1.3599.

[7] Hurajová, E.; et al. Biodiversity and Vegetation Succession in Vineyards, Moravia (Czech Republic). Agriculture 2024, 14, 1036. https://doi.org/10.3390/agriculture14071036.

[8] Pal, M.; Mather, P.M. Support vector machines for classification in remote sensing. Int. J. Remote Sens. 2005, 26, 1007–1011. https://doi.org/10.1080/01431160512331314083.

[9] Pal, M.; Mather, P.M. An assessment of the effectiveness of decision tree methods for land cover classification. Remote Sens. Environ. 2003, 86, 554–565. https://doi.org/10.1016/S0034-4257(03)00132-9.

[10] Cutler, D.R.; Edwards, T.C.; Beard, K.H.; Cutler, A.; Hess, K.T.; Gibson, J.; Lawler, J.J. Random forests for classification in ecology. Ecology 2007, 88, 2783–2792. https://doi.org/10.1890/07-0539.1.

[11] Laban, N.; Abdellatif, B.; Ebeid, H.M.; Shedeed, H.A.; Tolba, M.F. Machine Learning for Enhancement Land Cover and Crop Types Classification. In Machine Learning Paradigms: Theory and Application; Hassanien, A.E., Ed.; Springer International Publishing: Cham, Switzerland, 2019; pp. 71–87. https://doi.org/10.1007/978-3-030-02357-7_4.

[12] Phiri, D.; Simwanda, M.; Salekin, S.; Nyirenda, V.R.; Murayama, Y.; Ranagalage, M. Sentinel-2 Data for Land Cover/Use Mapping: A Review. Remote Sens. 2020, 12, 2291. https://doi.org/10.3390/rs12142291.

[13] Gauci, A.; Abela, J.; Austad, M.; Cassar, L.F.; Zarb Adami, K. A Machine Learning approach for automatic land cover mapping from DSLR images over the Maltese Islands. Environ. Model. Softw. 2018, 99, 1–10. https://doi.org/10.1016/j.envsoft.2017.09.014.

[14] Mardani, M.; Mardani, H.; De Simone, L.; Varas, S.; Kita, N.; Saito, T. Integration of Machine Learning and Open Access Geospatial Data for Land Cover Mapping. Remote Sens. 2019, 11, 1907. https://doi.org/10.3390/rs11161907.

[15] Pelletier, C.; Valero, S.; Inglada, J.; Champion, N.; Dedieu, G. Assessing the robustness of Random Forests to map land cover with high resolution satellite image time series over large areas. Remote Sens. Environ. 2016, 187, 156–168. https://doi.org/10.1016/j.rse.2016.10.010.

[16] Machine Learning Algorithms for Satellite Image Classification Using Google Earth Engine and Landsat Satellite Data: Morocco Case Study. IEEE Journals & Magazine. Accessed: April 2, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/10177754.

[17] Han, R.; Liu, P.; Wang, G.; Zhang, H.; Wu, X. Advantage of Combining OBIA and Classifier Ensemble Method for Very High-Resolution Satellite Imagery Classification. J. Sens. 2020, 2020, e8855509. https://doi.org/10.1155/2020/8855509.

[18] Ma, L.; Liu, Y.; Zhang, X.; Ye, Y.; Yin, G.; Johnson, B.A. Deep learning in remote sensing applications: A meta-analysis and review. ISPRS J. Photogramm. Remote Sens. 2019, 152, 166–177. https://doi.org/10.1016/j.isprsjprs.2019.04.015.

[19] Zaabar, N.; Niculescu, S.; Kamel, M.M. Application of Convolutional Neural Networks With Object-Based Image Analysis for Land Cover and Land Use Mapping in Coastal Areas: A Case Study in Ain Témouchent, Algeria. IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens. 2022, 15, 5177–5189. https://doi.org/10.1109/JSTARS.2022.3185185.

[20] Zhang, X.; Han, L.; Han, L.; Zhu, L. How Well Do Deep Learning-Based Methods for Land Cover Classification and Object Detection Perform on High Resolution Remote Sensing Imagery? Remote Sens. 2020, 12, 417. https://doi.org/10.3390/rs12030417.

[21] Big Data for Remote Sensing: Challenges and Opportunities. IEEE Xplore. Accessed: July 24, 2024. [Online]. Available: https://ieeexplore.ieee.org/document/7565634.

[22] Ienco, D.; Gbodjo, Y.J.E.; Gaetano, R.; Interdonato, R. Weakly Supervised Learning for Land Cover Mapping of Satellite Image Time Series via Attention-Based CNN. IEEE Access 2020, 8, 179547–179560. https://doi.org/10.1109/ACCESS.2020.3024133.

[23] Yang, N.; Tang, H. Semantic Segmentation of Satellite Images: A Deep Learning Approach Integrated with Geospatial Hash Codes. Remote Sens. 2021, 13, 2723. https://doi.org/10.3390/rs13142723.

[24] Yuan, X.; Shi, J.; Gu, L. A review of deep learning methods for semantic segmentation of remote sensing imagery. Expert Syst. Appl. 2021, 169, 114417. https://doi.org/10.1016/j.eswa.2020.114417.

[25] Garcia-Garcia, A.; Orts-Escolano, S.; Oprea, S.; Villena-Martinez, V.; Garcia-Rodriguez, J. A Review on Deep Learning Techniques Applied to Semantic Segmentation. arXiv 2017, arXiv:1704.06857. https://doi.org/10.48550/arXiv.1704.06857.

[26] Marmanis, D.; Wegner, J.D.; Galliani, S.; Schindler, K.; Datcu, M.; Stilla, U. Semantic Segmentation of Aerial Images with an Ensemble of CNNs. ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci. 2016, III–3, 473–480. https://doi.org/10.5194/isprs-annals-III-3-473-2016.

[27] Li, R.; Zheng, S.; Duan, C.; Wang, L.; Zhang, C. Land Cover Classification from Remote Sensing Images Based on Multi-Scale Fully Convolutional Network. Geo-Spat. Inform. Sci. 2022, 25, 278–294. https://doi.org/10.1080/10095020.2021.2017237.

[28] Tzepkenlis, A.; Marthoglou, K.; Grammalidis, N. Efficient Deep Semantic Segmentation for Land Cover Classification Using Sentinel Imagery. Remote Sens. 2023, 15, 2027. https://doi.org/10.3390/rs15082027.

[29] Xu, R.; Wang, C.; Zhang, J.; Xu, S.; Meng, W.; Zhang, X. RSSFormer: Foreground Saliency Enhancement for Remote Sensing Land-Cover Segmentation. IEEE Trans. Image Process. 2023, 32, 1052–1064. https://doi.org/10.1109/TIP.2023.3238648.

[30] Längkvist, M.; Kiselev, A.; Alirezaie, M.; Loutfi, A. Classification and Segmentation of Satellite Orthoimagery Using Convolutional Neural Networks. Remote Sens. 2016, 8, 329. https://doi.org/10.3390/rs8040329.

[31] Vali, A.; Comai, S.; Matteucci, M. Deep Learning for Land Use and Land Cover Classification Based on Hyperspectral and Multispectral Earth Observation Data: A Review. Remote Sens. 2020, 12, 2495. https://doi.org/10.3390/rs12152495.

[32] Digra, M.; Dhir, R.; Sharma, N. Land Use Land Cover Classification of Remote Sensing Images Based on Deep Learning Approaches: A Statistical Analysis and Review. Arab. J. Geosci. 2022, 15, 1003. https://doi.org/10.1007/s12517-022-10246-8.

[33] Zhao, S.; Tu, K.; Ye, S.; Tang, H.; Hu, Y.; Xie, C. Land Use and Land Cover Classification Meets Deep Learning: A Review. Sensors 2023, 23, 18966. https://doi.org/10.3390/s23218966.

[34] Chen, L.-C.; Papandreou, G.; Kokkinos, I.; Murphy, K.; Yuille, A.L. DeepLab: Semantic Image Segmentation with Deep Convolutional Nets, Atrous Convolution, and Fully Connected CRFs. IEEE Trans. Pattern Anal. Mach. Intell. 2018, 40, 834–848. https://doi.org/10.1109/TPAMI.2017.2699184

[35] He, K.; Zhang, X.; Ren, S.; Sun, J. Deep Residual Learning for Image Recognition. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; IEEE: Las Vegas, NV, USA, 2016; pp. 770–778. Accessed: April 22, 2023. [Online]. Available: https://openaccess.thecvf.com/content_cvpr_2016/html/He_Deep_Residual_Learning_CVPR_2016_paper.html.

[36] Szegedy, C.; Vanhoucke, V.; Ioffe, S.; Shlens, J.; Wojna, Z. Rethinking the Inception Architecture for Computer Vision. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition; IEEE: Las Vegas, NV, USA, 2016; pp. 2818–2826. Accessed: April 22, 2023. [Online].

Available: https://www.cv-foundation.org/openaccess/content_cvpr_2016/html/Szegedy_Rethinking_the_Inception_CVPR_2016_paper.html.

[37] Tan, M.; Le, Q.V. EfficientNet: Rethinking Model Scaling for Convolutional Neural Networks. arXiv 2020, arXiv:1905.11946. https://doi.org/10.48550/arXiv.1905.11946.

[38] Weinberger, K.Q. Densely Connected Convolutional Networks. arXiv 2018, arXiv:1608.06993. https://doi.org/10.48550/arXiv.1608.06993.