Automatic Extraction and Multi-Class Instance Segmentation of Rural Road Networks from Orthoimagery using YOLOv11 and SAHI Sliced Inference for Cadastral Update

Marsia Sanità [1], Lindo Nepi [2], Eva Savina Malinverni [1], Adriano Mancini[2], Roberto Pierdicca [1], Artur Warchoł [3], Monika Balawejder [4]

[1] Dept. of Civil, Building and Architecture, Marche Polytechnic University, 60131 Ancona, Italy - m.sanita@pm.univpm.it, e.s.malinverni@staff.univpm.it, r.pierdicca@staff.univpm.it
[2] Department of Information Engineering (DII), Marche Polytechnic University, 60131 Ancona, Italy - l.nepi@pm.univpm.it, a.mancini@univpm.it
[3] Kielce University of Technology – Kielce, Poland – awarchol@tu.kielce.pl
[4] PANS State University of Applied Sciences in Jaroslaw, Poland – monika.balawejder@pansjar.edu.pl

Keywords: Road Extraction, SAHI, YOLOv11, Deep Learning, Orthoimagery, Multi-class instance segmentation.

Abstract

Extracting road networks from high-resolution imagery remains a significant challenge in geomatics, particularly in fragmented rural landscapes. The big difficulty is the spectral similarities between unpaved tracks and agricultural backgrounds that can lead to classification errors. This study proposes an automated geospatial pipeline based on the YOLOv11 architecture. Specifically, the approach is made on the optimization of the multi-class road detection in the rural areas of Kosina and Markowa, two villages in Poland. To reduce the computational effort, due to large-scale 9000x9000 px orthophotos and to improve the detection of small-scale features, Slicing Aided Hyper Inference (SAHI) strategy was integrated. High-resolution imagery has been decomposed into optimized tiles, ensuring feature continuity across boundaries and preventing GPU memory overhead. The instance segmentation model was trained on a custom-annotated dataset, with seven labels (categories) such as internal paved roads, rural tracks, and railway infrastructures. Therefore, a high level of robustness has been achieved reaching a mean Average Precision value (mAP@0.5) of 0.90. A confusion matrix reveals quantitatively that the pipeline effectively distinguishes between complex classes and low omission rates. As a result, the generated outputs are converted into interoperable GeoJSON format ensuring their integration into GIS environments. In conclusion, the experimental result demonstrates that the framework is valuable for emergency response logistics and urban planning. It offers a scalable and near real-time solution for updating national topographic databases.

1. Introduction

The progressive urban growth is not limited to the urban area, but it encloses the rural and the near-rural zones. This implies to an unplanned landscape transformation (Liu et al., 2024). A consequence of these new dynamics is the creation or modification of infrastructures, such as new rural roads, which are often not visible into the cadastral registry (Warchoł et al., 2022). As technology advances, methods for obtaining data for cadastre updates are becoming increasingly accurate; furthermore, due to rapid digitization, the availability of source materials obtained using modern technologies is also increasing. However, the inconsistency between the cadastre database and the actual situation is increasingly common issue (Buśko et al., 2023).

The land and building register, also known as the real estate cadastre, contains data on the quantity, dimensions, location, ownership of individual land plots, and their use (Apollo et al., 2023; Buśko et al., 2023; Maciuk et al., 2021). Therefore, the land and building register database is used, among other things, to distribute taxes, determine land ownership, manage land market prices, calculate agricultural subsidies, designate zones based on value and purpose of use, including roads, as well as to create the spatial data infrastructure of countries (Sobolewska-Mikulska et al., 2020; Dawidowicz et al., 2020; Balawejder et al., 2023).

The misalignment between the current and recorded status in the land cadastre can pose a big challenge in land registry level, in land consolidation process and in case of disaster management (Balawejder et al., 2018; Deng et al., 2025). Road detection is also fundamental for new urban planning or regeneration, and even for autonomous driving systems.

As shown in Figure 1, public roads can be divided into national roads, provincial roads, county roads, and municipal roads. National roads include, among others, highways, expressways, and international roads. Provincial roads are roads that connect cities. Roads connecting county seats are classified as county roads. Municipal roads are roads of local importance and do not fall into other categories, constituting a supplementary road network serving local needs. Private roads are not classified as municipal roads (Noga et al., 2017). In the study area, in the village of Markowa, agriculture is highly developed.

The road infrastructure network consists largely of access roads to agricultural plots and is often not recorded in the land and building register database, although it is present in the field.

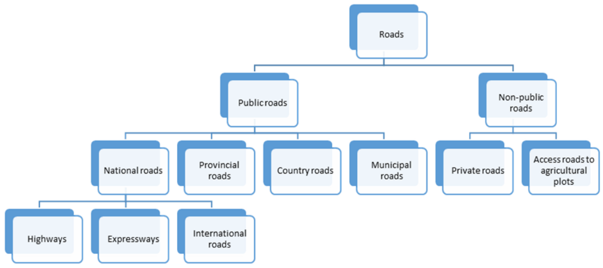

Figure 1. Division of public and non-public roads.

To address these discrepancies, this research introduces an automated pipeline specifically fine-tuned for the detection of rural and internal road networks. While road extraction is a widely debated topic in remote sensing, most existing models focus on structured urban highways or main arterial routes. Targeting the 'rural niche' represents a notable innovation, as these paths often lack standardized spectral signatures and are frequently confounded by surrounding vegetation and agricultural patterns. By developing a model specifically optimized for these challenging environments, we provide a tool that can bridge the gap between official cadastral records and the current ground reality following the framework for automated road vectorization proposed by Jiao et al. (2024).

The practical value of this approach extends beyond administrative updates; it is particularly critical for near real-time emergency response and post-disaster management. In the event of floods, landslide or seismic activities, having an up-to-date map of rural access roads, which are often the only lifelines for isolated communities, is essential for coordinating rescue logistics and assessing infrastructural damage. For similar purposes, other studies leverage DL techniques for post-disaster intervention (Linardos et al., 2022; Chaudhuri et al., 2020; Singla et al., 2025; Chicchon et al., 2024; Bhardwaj et al., 2025).

The paper is structured as follows: Section 1.1 reviews the related literature on deep learning for road extraction; Section 2 details the Materials and Methods, including the YOLOv11 architecture and the SAHI tiling strategy; Section 3 presents the experimental results and the quantitative analysis; and finally, Section 4 draws the conclusions and outlines future research directions.

1.1 State of the Art

Traditional topographic surveying requires a long period for acquisitions, while the application of Convolutional Neural Network (CNN) is a faster and more efficient methodology. Geomatics is widely used for the management and mapping of road infrastructure. The combination of both the disciplines is defined GeoAI (Pierdicca et al., 2022) and can automatically extract, classify, and analyze spatial information. This approach offers key advantages, including the ability to process large-scale and high-resolution datasets efficiently, reduce manual interpretation effort, ensure consistent and reproducible results, and capture complex spatial patterns that are difficult to identify using traditional methods (Pierdicca et al., 2023; Sanità et al., 2025). There are various Deep Learning (DL) algorithms used for road detection, as reported in this review (Liu et al., 2024). The approaches are mainly based on fully supervised, semi-supervised, and unsupervised learning (Mo et al., 2024). In the literature, there are studies demonstrating the use of the U-Net network for this type of analysis because it can retain spatial information and capturing contextual features. Very common is the use of Semantic Segmentation (SS) technique being a specific task in specialized fields such as disaster management and environmental monitoring. However, its drawback is the requirement for extensive datasets but to address this limitation, Few-Shot Semantic Segmentation (FSSS) has been introduced, requiring only a limited number of annotated samples (Petrov et al., 2026).

Accurate mapping is necessary for road network detection; however, the manual method is very time-consuming. Recent studies have leveraged GeoAI approaches to ensure a faster response. Çevikalp et al., proposed their method, using deep learning models like U-Net, comparing the accuracy assessment of road segmentation using high (Sentinel-2) and very high (Pléiades) spatial resolution imagery. They demonstrated that the accuracy decreases linearly to the spatial resolution reduction passing from F1 score (0.87) for Pléiades with spatial resolution of 30cm to F1 score (0.78) for Sentinel-2 imagery with spatial resolution of 10 m.

A multi-scale road segmentation model for UAV remote sensing imagery is proposed by Cui et al., focusing on the YOLO model limitations. They improved the YOLOv12 model focalizing to its use for UAV imagery named UAVYOLOv12.

The latest research is concentrated to generalize a model, furthermore some different methodologies have limitations especially for accuracy and real-time perform due to the heavy computing effort. The use of RGB-D data is often limited to an inadequate data fusion approach that can lead to a false representation in sensible region (Xue et al., 2025). To overcome this bottleneck, they proposed a data augmentation phase to increase the sensible region quantity useful for training the model. In addition, their methodology separates the road detection in two different lines and then they fused the results obtaining previously a result in each one of both subnets. Generalize a model is an actual challenge because most of the dataset annotated are composed by very high-resolution imagery in which the spatial resolution is around 1m. Then, the critical issue is the inference made on a dataset with lower spatial resolution, such as Sentinel-2 (10 m). This challenge has been investigated by Jangir et al. (2025); however, a key methodological distinction exists. While our approach utilizes a dataset natively annotated at high resolution for both training and inference, their research employs super-resolution techniques on Sentinel-2 imagery. Their objective is to generalize the model by artificially enhancing low-resolution data, whereas our pipeline leverages the intrinsic geometric fidelity of native high-resolution sources. Different approach with high accuracy and low Dice Loss are reported in the study by Yanuargi et al. (2025). They employed a Deep Learning (DL) Convolutional Neural Network (CNN) based on the U-Net architecture for road segmentation. They novelty was on the identification of new road paths evaluating two different encoders: Inception-ResNet-V2 and a pure U-Net encoder. Notably, the Inception-ResNet-V2 architecture achieves an accuracy of 91.3%, compared to 90.7% for the standard U-Net.

The development of a specific dataset for rural roads, as presented in this manuscript, is strongly required to bridge the gap between general-purpose foundation models and the specific requirements of rural infrastructure monitoring. As highlighted by Ameen et al. (2026), this process is essential for effective fine-tuning, which enables models to adapt their broad visual knowledge to the distinct semantic features of complex geospatial environments. The rapid evolution of computer vision led to the introduction of new foundation models leaving a gap between the natural images and geospatial data. Subrahmanya (2026) assumes that the standard architecture forgets the pre-trained features and degrade model generalization. As demonstrate Blushtein-Livnon et al. (2025), light fine-tuning with manual geometric annotation assures high precision segmentation.

The promising idea is to exploit cutting-edge model to achieve automated analysis of orthoimages capable of balancing high performance with lower training and execution times compared to other solutions based on Vision Transformer (ViT) architectures. A fast and efficient method for road detection is presented in (Sun et al., 2025) where a pre-trained ViT encoder has been considered to extract the features from RGB data. In this paper, CNN and ViT were considered for a multi class instance segmentation.

The novelty introduced in this paper is the innovative framework that exploits the sensing power of a DL model for a multiclass classification on the specific domain of rural network. Rural road networks present unique challenges, such as spectral similarity with the surrounding environment and irregular geometries (Mo et al., 2024). Furthermore, it is necessary to recognize the novelty of the multiclass approach as an essential element for the recognition of rural roads in country contests where the secondary or internal road can be confused with temporary passages of agricultural vehicles.

2. Material and method

The core of this research lies in the development of a fully automated pipeline designed to bridge the gap between deep learning inference and geospatial data integrity within a context of rural road networks. The proposed methodology is structured into five sequential stages detailed in Section 2.3 and represented in Figure 2. This sequential approach ensures the consistency of the road extraction process from raw data to final validation.

2.1 Study area

The proposed experimental framework of this study focuses on two rural settlements in Southeastern Poland: Kosina and Markowa, both located within the Podkarpackie Voivodeship (Figure 3). These areas were selected to evaluate the model's transferability across geographically proximal yet structurally distinct rural environments. These villages are situated within the transition zone between the Sandomierz Basin and the Subcarpathian Foothills. A predominantly flat to gently undulating topography characterizes this area, minimizing extreme topographic shadowing while introducing complexities related to varying soil moisture and agricultural reflectance.

Consistent phenological stages are ensured by the presence of a temperate continental climate. Within the study area, in the village of Markowa, agriculture is highly developed. The road infrastructure network consists largely of access roads to agricultural plots and is often not recorded in the land and building register database, although it is present in the field.

2.2 Dataset

The dataset consists of 80 RGB orthophotos in GeoTIFF format, acquired in a specific rural area in Poland and were retrieved from the Polish National Geoportal[1]. The ortophotos considered refers to Kosina and Markowa, urban-rural municipalities in the southeastern Poland. The image resolution is 0.25 m/px, acquired in 2022, the original size was 9000x9000 px in PL-1992 coordinate system (EPSG:2180).

2.3 Methodology

The data collection is reported in Section 2.2. To reduce the computational training cost, the original orthophotos have been cut and reduced in 1000x1000 tile size. Furthermore, before the training phase, data augmentation has been applied to increase the number of images in the input dataset. At the same time, working with smaller tiles simplify the data annotation made by domain experts. The data augmentation techniques have been performed making rotation, flipping, brightness changes, contrast, etc. The final input dataset results composed by 16000 images. To evaluate the model’s performance, another dataset referred to another Poland municipality has been considered. The input dataset has been split in 80/20 for training and validation.

The data annotation made by domain experts as reported in Figure 4 that is an example on how this step has been performed.

The supervised approach described and used in this work is a multiclass type based on seven road-classes: asphalt, exit, hard road, internal pavement, internal road, rural road, and railways.

“Asphalt” has been used to annotate the road made with asphalt that was out of private property. The “exit” were all the road parts that were at a cross point between the principal road and an internal pavement of a private or public road or property. “Hard

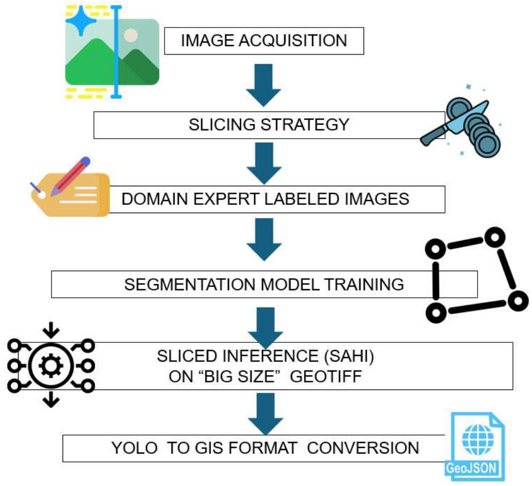

Figure 2. Methodology.

Figure 3. Location of the study area in Poland.

Figure 4. Example of the labelled image.

[1] https://www.geoportal.gov.pl/

road” roads that are difficult to travel. “Internal pavement” is for the pavement area inside a parking lot or inside a private pavement. “Internal Road” is a private road for residential use. “Rural road” are the roads in the vegetation area. To conclude, “railways” refers to the railway infrastructure.

The annotation phase has been conducted in Roboflow web platform[2] creating a polygon along the road boundary recognized by the experts. Roboflow is a platform for image annotation within the Machine Learning (ML) and computer vision contexts. It enables users to perform a dataset preparation, annotation and some model training in a unified workflow. It works very well with Python Notebook. Using the Python API is possible to send images to Roboflow models and obtain prediction within object detection and classification in short time. Some State-Of-The-Art models as YOLOv11-v26 and RF-DETR are integrated in Roboflow.

Instance segmentation aims to identify, localize, and delineate pixel-level boundaries for every individual object within an image. Unlike standard semantic segmentation, this task differentiates between multiple instances of the same class, assigning each a unique identity and a predefined category. This level of detail enables advanced spatial analysis, such as automated object counting and precise calculation of the pixel area for each detected feature. As an evolution of object detection, instance segmentation represents one of the most complex challenges in computer vision, requiring the model to simultaneously perform bounding box localization and classification. It occurs frequently in remote sensing projects.

Although tiling orthophotos during the training phase helps the model with small object detection, the problem resurfaces during the inference phase. Large-scale images undergo downsampling to prevent GPU memory overload and to comply with the input dimensions required by computer vision models.

The SAHI technique offers a simple and effective approach to aid in small object detection during inference on large-scale imagery. A slicing-aided inference is implemented where the image is divided into partially overlapping tiles each of which undergoes individual inference. This technique is considered model-agnostic because can be applied to any pre-trained model without any additional fine-tuning.

2.4 YOLO architecture

The YOLO model was first introduced in 2015 and has been continuously improved. YOLO has a single-sot detection architecture (SSD) based on CNNs, enable one-stage detection. Here, the prediction of both bounding box and class occurs in a single step. The widespread use of Ultralytics YOLO models in computer vision projects depends on their different approach respect to the traditional two-stage detection approach used in models like Faster R-CNN e Mask R-CNN. The SSD approach considers object detection as a single regression problem, making the process faster but less accurate than the two-stage detection approach. In the latter, the object detection process is slower but has higher precision because it is separated into two stages. The main goal is to develop models that maintain SSD speeds with high precision.

This has been achieved by introduction innovation in the subsequent versions such as C2PSA and C3k2 modules (Khanam et al., 2024). The former module implements an attention mechanism while the latter ensures more efficient feature map processing. These kinds of innovations allow YOLO models to achieve performances levels comparable to model based on Vision Transformer architectures.

The precise model YOLOv11, used to perform our research on road extraction, join the Ultralytics YOLO (You Only Look Once) family since 2024. Ultralytics YOLOv11 model is very robust for object detection, instance segmentation, pose estimation and classification. Other models have been tested with this input dataset, and the best performance has been achieved with the YOLOv11 medium-size model. It delivered optimal performance while effectively mitigating the risks of both overfitting and underfitting. An important contribution derived from the Slicing Aided Hyper Inference SAHI (Akyon et al., 2022), a tool specifically designed to handle large-scale inference projects and natively integrated with the Ultralytics software package. SAHI has the responsibility to manage the sliced inference even when performs on image bigger than those of training. Despite this, it successfully identifies small objects.

2.5 Inference pipeline

Figure 5 represents the pipeline representation of the automated processing script, illustrating the internal logic from data ingestion to the generation of the final road maps. As a first step, the script loads the 9000x9000-pixel GeoTIFF image from the test dataset. The model has never seen this image during the training and validation phases. It then extracts the coordinates, the resolution and the reference systems. The second step refers to the SAHI to reduce the computational effort.

Since the orthophoto exceeds GPU memory capacity, our algorithm employs the SAHI technique. Rather than processing the entire image in a single pass, inference is performed on individual tiles with dimensions like those used during the training phase. This prevents GPU memory overhead and ensures a low system memory footprint, allowing the inference to run efficiently even on hardware with limited resources without compromising processing speed. During sliced inference, the image is processed in 1000x1000 px slices (tile). An overlap ratio of 0.3 is maintained between adjacent tiles to ensure that objects bisected by tile boundaries are neither missed nor double-counted. In this step YOLO model analyses each individual slice, identifying regions of interest and generating pixel-based segmentation masks. The algorithm then aggregates these local results, reconstructing them to match the original dimensions of the full-scale image.

In the next step, the algorithm converts the pixel coordinates of the identified polygons into real-world geographical coordinates

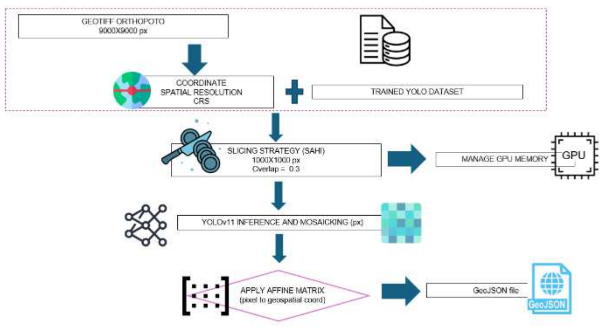

Figure 5. Inference pipeline.

---
[2] https://app.roboflow.com/

(expressed in metres or degrees). This is achieved by applying a mathematical transformation using an Affine Matrix, derived from the metadata previously extracted from the GeoTIFF file. The last step is the GeoJSON saving files. Simply the algorithm transforms the image in a geographic database letting the knowledge of the semantic and geographic information of the element. It can manage automatically the complex amount of georeferenced data.

The ultimate scope of the workflow developed in this paper takes archives of orthophotos as input and produces an output archive in an open format (GeoJSON) to manage a collection of spatial geometries enabling their use in GIS platforms. The second innovative aspect of this manuscript is the creation of a dedicate Python script named *Yolov11Seg2GeoJSON* [3] to generate automatically a GeoJSON file from the YOLOv11 inference on georeferenced imagery. The algorithm included in the Python script let to process GeoTIFF with YOLO-v11 based segmentation model without resolution loss. Since the YOLO architecture does not intrinsically handle the geographic headers found in georeferenced imagery, the *Yolov11Seg2GeoJSON* is necessary to bridge the gap between image-space and GIS environments for an immediate spatial analysis. After the model completes the inference in pixel space (x, y), the script translates these local coordinates into a global reference system.

3. Results

The use of the SAHI module for sliced inference enabled us to achieve strong metrics even when performing inference on orthophotos significantly larger than those used during the training phase, ensuring the successful identification of even the smallest objects. In addition to YOLO, training and inference tests were also conducted using the RF-DETR model, an object detection and instance segmentation framework developed by Roboflow based on Vision Transformer technology, which yielded results comparable to those of the YOLO model during the inference phase. Furthermore, our selected model supports the SAHI module for sliced inference without the need for additional custom code.

The training process was monitored through loss functions and accuracy metrics over 350 epochs (Figure 6 and Figure 7). Both training and validation losses for box, segmentation, and classification tasks showed a consistent downward trend, converging without signs of overfitting. This stability indicates that the YOLOV11 model effectively captured the underlying spatial patterns of the rural road network. Furthermore, the precision, recall, and mAP@0.5 metrics reached a stable plateau, confirming the robustness of the weights used for the subsequent inference phase on the Kosina and Markowa datasets.

The evaluation of the YOLOv11-based pipeline performance is made using the normalized confusion matrix shown in Figure 8. The results point to a high level of classification robustness, with diagonal elements consistently exceeding 0.88. This indicates a strong alignment between our predictions and the ground truth across all road categories.

In particular, the model performed exceptionally well on Internal Roads (0.96), Rural Roads (0.95), and Tramway/Railways (0.95). These figures are especially meaningful considering the highly fragmented landscape of the Kosina and Markowa study areas.

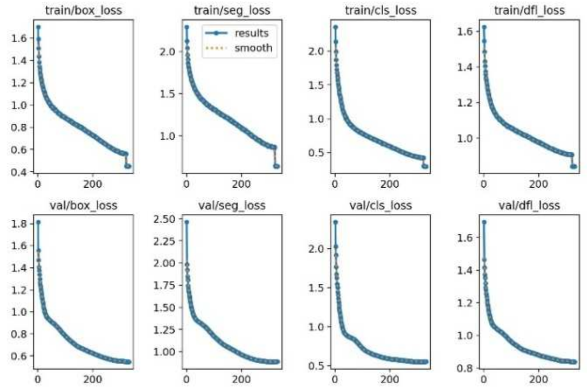

Figure 6. Training and validation metrics of the YOLOV11 model (loss function).

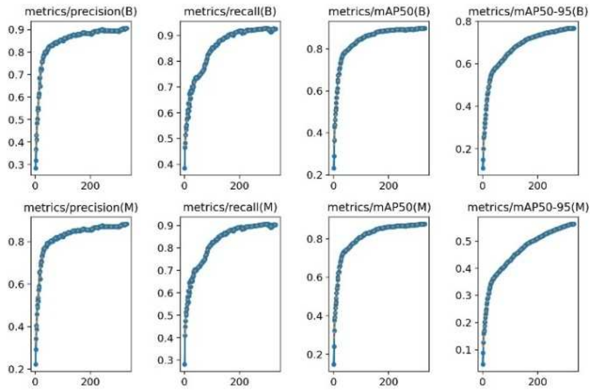

Figure 7. Training and validation metrics of the YOLOv11 model (precision and recall).

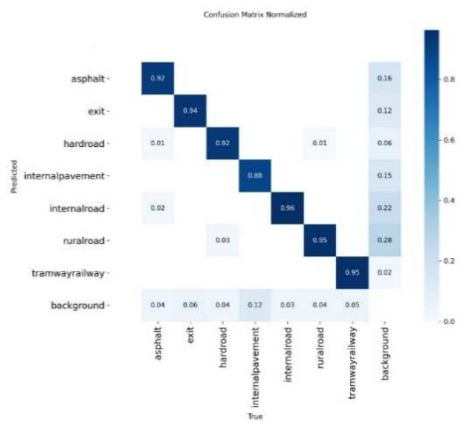

Figure 8. Confusion matrix.

The precision achieved for rural roads suggests the model successfully captured the specific geometric cues, such as linearity and typical widths that separate narrow, unpaved paths from the agricultural surroundings.

However, the data also highlights some expected challenges. It is evident a minor spectral overlap between Internal Pavement (0.88) and the Background (0.15). This likely stems from the similar radiometric signatures of concrete courtyards and certain roofing materials common in these villages. Similarly, the Asphalt class was misclassified as Background in 16% of cases. This is a typical issue in rural geomatics: older asphalt often weathers or becomes covered by dust and debris, making its

[3] https://github.com/lindonepi/Yolov11Seg2GeoJSON

spectral signature nearly indistinguishable from bare soil or dry vegetation.

Finally, while it is observed a 0.28 False Positive rate for rural roads against the background, this is a well-known hurdle in high-resolution road extraction. Features like linear farm tracks or irrigation furrows often mimic the topology of the actual road network. On the other hand, the remarkably low False Negative rates (between 0.03 and 0.12) across all classes are a key result. They confirm that the pipeline is highly reliable at maintaining network continuity, a feature that is essential both for enhancing emergency response logistics in rural areas and for supporting long-term urban planning and infrastructure management within GIS-based systems.

The Precision-Recall (PR) curve (Figure 9) further validates the model's effectiveness, showing an overall mAP@0.5 of 0.900. The training phase has been conducted over 350 epochs using an initial learning rate of 0.001 and a weight decay of 0.0002 for model regularization. The input image size, which is usually set to 640px in the Yolo configuration, was set to 1000 pixels to match the size of the images in the annotated dataset. This helps improve the model's ability to retain image detail and recognize small objects. An early stopping mechanism was implemented to terminate training before the onset of overfitting. A Focal Loss function has been used to enhance model performance. Model training and inference were performed on a Linux Ubuntu O.S. with an Intel i9 CPU and an Nvidia RTX 4070 Ti GPU.

Specifically, the model achieves near-optimal results for structured categories such as Exits (0.961) and Asphalt roads (0.937). While the Rural Road class shows a slightly lower performance (0.873) due to the inherent complexity of unpaved surfaces in agricultural zones, it still demonstrates remarkable stability. The lower mAP for Tramway/Railway (0.746) is likely linked to the lower frequency of these features within the study areas. Overall, the PR analysis confirms that the YOLOv11-based workflow is robust and well-balanced for automated road mapping. The sharp, rectangular profile of the curves for most categories indicates that the pipeline maintains high precision levels even at high recall thresholds.

Figure 10 presents the visual inference results on a grid of image tiles from the study area. The qualitative inspection confirms the model's high precision in detecting diverse road categories, with confidence scores consistently exceeding 0.85.

A key observation is the model's ability to maintain feature continuity across tile boundaries that is a direct result of the Sliced Aided Hyper Inference (SAHI) strategy. Even in densely

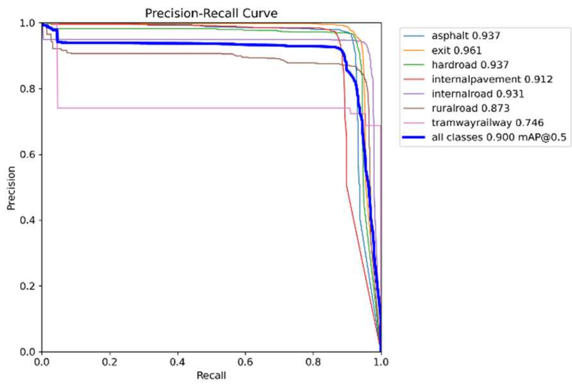

Figure 9. Precision-recall curve. The bold blue line represents the mean Average Precision (mAP) across all classes.

Figure 10. Class prediction.

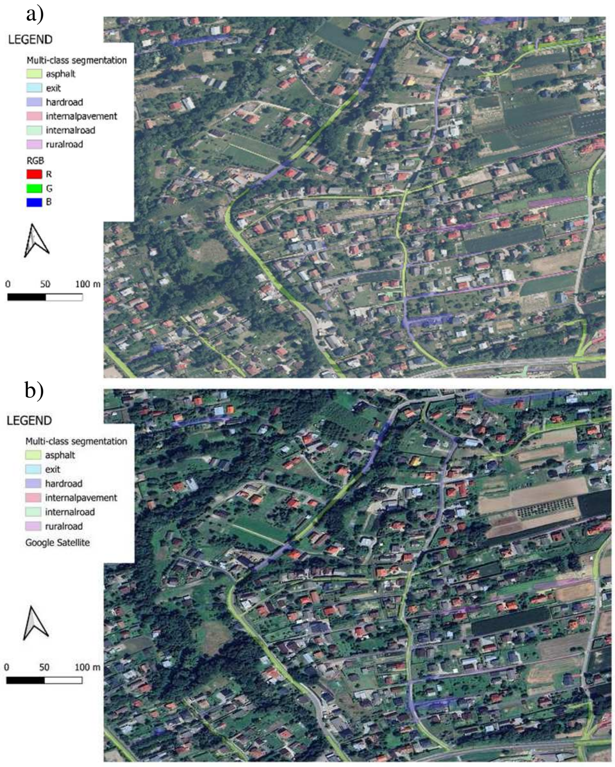

Figure 11. a) Overlapping between GeoJSON multi-class segmentation and raw ortophoto image. b) Overlapping between GeoJSON multi-class segmentation and google satellite wms.

structured rural zones, where internal roads and farm tracks are closely intertwined, the YOLOv11 framework successfully distinguishes between individual segments. The spatial consistency of these bounding boxes and masks is critical for the subsequent conversion into a topologically correct GeoJSON network, but we have our solution reported in Section 2.5.

The qualitative results are represented in Figure 11 in which the GeoJSON file, multiclass segmentation results, has been overlapped to its input orthoimage (Figure 11.a) and to the Google Satellite (Figure 11.b). It is evident how accurate are the results obtained from our experiment.

4. Conclusion

This research successfully implemented an automated geospatial pipeline based on the YOLOv11 architecture, specifically optimized for road extraction in complex rural landscapes. The annotation has been performed and provided by the authors in Roboflow environment, considering the 80% of the dataset for training and the 20% of the dataset for testing. To address the critical challenge of recognizing small rural roads, YOLO has been enhanced through sliced inference using Slicing Aided Hyper Inference (SAHI) cutting the large size images in 1000x1000 px tiles. Our approach gives good performance (Figure-8), reducing the time consuming and considering a multi-class supervised training. By integrating Slicing Aided Hyper Inference (SAHI), it has been possible to address the dual challenge of detecting small-scale features, such as unpaved rural tracks, while maintaining high computational efficiency on large-scale 9000x9000 px orthophotos. The efficiency of the proposed tiled-inference pipeline makes it particularly suitable for time-critical emergency response and post-disaster assessment (Figure-5). By providing rapid, automated road network extraction in near real-time, the system enables authorities to quickly identify accessible routes and logistical bottlenecks immediately following a catastrophic event, where manual mapping would be too slow to support life-saving operations.

The results demonstrate that this tiled-inference approach not only prevents GPU memory overhead but also significantly outperforms conventional manual or semi-automated mapping methods in terms of processing speed. The quantitative evaluation yielded a robust mAP@0.5 score of 0.90, confirming the model's high reliability across diverse road categories in the Kosina and Markowa study areas. Beyond the technical performance, the remarkably low omission rates ensure the network continuity required for practical GIS applications, ranging from emergency response logistics to long-term urban infrastructure planning. A significant technical contribution is made by the development of a python script for data type convertion. We designed a dedicated script to translate pixel-level masks into georeferenced GeoJSON structures ensuring full interoperability in GIS environments and for immediate integration into land register databases.

Future work will focus on expanding the diversity of the training dataset to include a wider range of geographical regions. Since the current study areas, Kosina and Markowa, share similar phenological and topographic characteristics, a key objective will be to test the model's generalization capabilities in significantly different environments. Finally, exploring additional instance segmentation models will help further refine the detection of overlapping features, ultimately providing a scalable and interoperable framework for the rapid update of national topographic databases. Moreover, future research will focus on compare our YOLOv11-SAHI framework on large-scale benchmarks as for example the WHU-CR dataset. It assesses to perform our model across diverse geographic regions. A comparative analysis with the Narrow-Road-Aware Network (NANet) could be conducted to benchmark our approach against state-of-the-art architectures specifically designed for narrow rural road extraction (Wang et al., 2026).

References

Akyon, F. C., Onur Altinuc, S., Temizel, A., 2022: Slicing Aided Hyper Inference and Fine-Tuning for Small Object Detection, 2022 IEEE International Conference on Image Processing (ICIP), Bordeaux, France, 2022, pp. 966-970, doi.org/10.1109/ICIP46576.2022.9897990.

Ameen, H., Soilán, M., 2026: A review of large language models in geomatics: integrating multimodal data, addressing challenges, and exploring synergies, *Spat. Inf. Res.*, 34, 16. doi.org/10.1007/s41324-026-00670-3.

Apollo, M., Jakubiak, M., Nistor, S., Lewińska, P., Krawczyk, A., Borowski, L., Specht, M., Krzykowska-Piotrowska, K., Marchel, Ł., Pęska-Siwik, A., Kardoš, M. Maciuk, K., 2023: Geodata in science - a review of selected scientific fields. *Acta Scientiarum Polonorum. Formatio Circumiectus*, 22(2), 17–40. doi.org/10.15576/ASP.FC/2023.22.2.02

Balawejder, M., Banaś, M., Matkowska, K., Warchoł, A., 2018: The use of geomatic tools and modern technologies to road network and registered plot analysis in Poland, 17th edition National Technical-Scientific Conference Modern Technologies for the 3rd Millennium, Oradea, Romania, 2018, pp. 9-14.

Balawejder, M., Warchoł, A., Konttinen, K., 2023: Energy Efficiency in Agricultural Production—Experience from Land Consolidation in Poland and Finland, *Energies*, 16(22), 7598. doi.org/10.3390/en16227598.

Bhardwaj, D., Nagabhooshanam, N., Singh, A. et al., 2025: Enhanced satellite imagery analysis for post-disaster building damage assessment using integrated ResNet-U-Net model, *Multimed Tools Appl*, 84, 2689–2714. doi.org/10.1007/s11042-024-20300-0.

Blushtein-Livnon, R., Rafaeli, O., Ioffe, D., Boger, A., Esquenazi, K. S., Svoray, T., 2025: On the Effectiveness of Textual Prompting with Lightweight Fine-Tuning for SAM3 Remote Sensing Segmentation. arXiv preprint arXiv:2512.15564v1. doi.org/10.48550/arXiv.2512.15564.

Busko, M., Balawejder, M., Kovalyshyn, O., Apollo, M., 2023: Do geographic location and historical conditions affect the quality and availability of open cadastral data? From early cadastral maps till now, *Reports on Geodesy and Geoinformatics*, 116(1), 23-38. doi.org/10.2478/rgg-2023-0008.

Çevikalp, M. R., Mutlu, B., Yanalak, M., and Musaoğlu, N.: Comparative Analysis of Multi-Resolution Remote Sensing Data for Accurate Road Segmentation in Urban Environments, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-M-6-2025, 103–108, https://doi.org/10.5194/isprs-archives-XLVIII-M-6-2025-103-2025, 2025.

Chaudhuri, N., Indranil, B., 2020: Exploring the role of deep neural networks for post-disaster decision support, *Decision Support Systems*, 130, 113234. doi.org/10.1016/j.dss.2019.113234.

Chicchon, M., Malinverni, E. S., Sanità, M., Pierdicca, R., Colosi, F., & Trujillo, F. J. L., 2024: Building Semantic Segmentation Using UNet Convolutional Network on SpaceNet Public Data Sets for Monitoring Surrounding Area of Chan Chan (Peru), *Geomatics and Environmental Engineering*, 18(3), 25–43. doi.org/10.7494/geom.2024.18.3.25.

Cui, B., Liu, Z., Yang, Q., 2025: UAV-YOLO12: A Multi-Scale Road Segmentation Model for UAV Remote Sensing Imagery, *Drones*, 9(8), 533. doi.org/10.3390/drones9080533.

Dawidowicz, A., Kulawiak, M., Zysk, E., Kocur-Bera, K., 2020: System architecture of an INSPIRE-compliant green cadastre system for the EU Member State of Poland. *Remote Sensing Applications: Society and Environment*, 20:100362, doi:10.1016/j.rsase.2020.100362.

Deng, L., et al., 2025: GLD-Road: A global–local decoding road network extraction model for remote sensing images, *ISPRS Journal of Photogrammetry and Remote Sensing*, 228, 741-755. doi.org/10.1016/j.isprsjprs.2025.07.026.

Jangir, S. K., Henry, C., Merkle, N., 2025: Investigating the Potential of Super-Resolution for Road Segmentation in Sentinel-2 Images, IGARSS 2025 - 2025 IEEE International Geoscience and Remote Sensing Symposium, Brisbane, Australia, 2025, pp. 1762-1766. doi.org/10.1109/IGARSS55030.2025.11243039.

Jiao, C., Heitzler, M., Hurni, L., 2024: A novel framework for road vectorization and classification from historical maps based on deep learning and symbol painting, *Computers, Environment and Urban Systems*, 108, 102060. doi.org/10.1016/j.compenvurbsys.2023.102060

Khanam, R., Hussain, M., 2024: YOLOv11: An Overview of the Key Architectural Enhancements. arXiv preprint arXiv:2410.17725v1. doi.org/10.48550/arXiv.2410.17725.

Linardos, V., Drakaki, M., Tzionas, P., Karnavas, Y. L., 2022: Machine Learning in Disaster Management: Recent Developments in Methods and Applications, *Machine Learning and Knowledge Extraction*, 4(2), 446-473. doi.org/10.3390/make4020020.

Liu, R., et al., 2024: A Review of Deep Learning-Based Methods for Road Extraction from High-Resolution Remote Sensing Images, *Remote Sensing*, 16(12), 2056. doi.org/10.3390/rs16122056.

Maciuk, K., Peska-Siwik, A., El-Mowafy, A., Borowski, L., Apollo, M., 2021: Crustal Deformation Across and beyond Central Europe and Its Impact on Land Boundaries. *Resources*, 10(2), 15. https://doi.org/10.3390/resources10020015

Mo, S., Shi, Y., Yuan, Q., Li, M., 2024: A Survey of Deep Learning Road Extraction Algorithms Using High-Resolution Remote Sensing Images, *Sensors*, 24(5), 1708. doi.org/10.3390/s24051708.

Noga, K., Balawejder, M., Matkowska, K., 2017: Dimensions of destruction of road network providing access to cadastral parcels resulting from motorway construction, *Geomatics and Environmental Engineering*, 11(4), 65-81. dx.doi.org/10.7494/geom.2017.11.4.65.

Petrov, M., Pandilova, E., Dimitrovski, I., Trajanov, D., Spasev, V., Kitanovski, I., 2026: Few-Shot Semantic Segmentation in Remote Sensing: A Review on Definitions, Methods, Datasets, Advances and Future Trends, *Remote Sensing*, 18(4), 637. doi.org/10.3390/rs18040637.

Pierdicca, R., Nepi, L., Mancini, A., Malinverni, E. S., and Balestra, M.: UAV4TREE: DEEP LEARNING-BASED SYSTEM FOR AUTOMATIC CLASSIFICATION OF TREE SPECIES USING RGB OPTICAL IMAGES OBTAINED BY AN UNMANNED AERIAL VEHICLE, ISPRS Ann. Photogramm. Remote Sens. Spatial Inf. Sci., X-1/W1-2023, 1089–1096, https://doi.org/10.5194/isprs-annals-X-1-W1-2023-1089-2023, 2023.

Pierdicca, R., Paolanti, M., 2022: GeoAI: a review of artificial intelligence approaches for the interpretation of complex geomatics data, *Geoscientific Instrumentation, Methods and Data Systems*, 11(1), 195-218. doi.org/10.5194/gi-11-195-2022.

Sanità, M., Malinverni, E. S., Pierdicca, R., Mancini, A., Glowienka, E., and Nepi, L.: AI4EO hyperview challenge: combination of machine learning methods on hyperspectral images to predict the soil parameters, Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci., XLVIII-G-2025, 1315–1322, https://doi.org/10.5194/isprs-archives-XLVIII-G-2025-1315-2025, 2025.

Singla, S., Priyanshu, Rohit, Swami, A., Singh, L., Singla, P., 2025: Post-Disaster Damage Analysis Using Deep Learning Techniques and Convolutional Neural Networks (CNNs), 2025 Seventh International Conference on Computational Intelligence and Communication Technologies (CCICT), Sonepat, India, 2025, pp. 227-234, doi.org/10.1109/CCICT65753.2025.00044.

Sobolewska-Mikulska, K. and Cienciała, A., 2020: Legal surveying documentation in the registration of rights to properties of the treasury and of territorial self-government units– case study of Poland. *Real Estate Management and Valuation*, 28(2):76–88, doi:10.1515/remav-2020-0017.

Subrahmanya, P. P., 2026: An Experimental Review of Foundation Models for Remote Sensing Image, *The GIS-IDEAS Journal*, 2(1), 1-14. doi.org/10.82667/gisideas.2.1_1

Sun, T., et al., 2025: ROD: RGB-Only Fast and Efficient Off-Road Freespace Detection, 2025 IEEE International Conference on Robotics and Automation (ICRA), Atlanta, GA, USA, 2025, pp. 9787-9793. doi.org/10.1109/ICRA55743.2025.11127491.

Wang, N., et al., 2026: Identifying rural roads in remote sensing imagery: From benchmark dataset to coarse-to-fine extraction network—A case study in China, *ISPRS Journal of Photogrammetry and Remote Sensing*, 231, 487-506. doi.org/10.1016/j.isprsjprs.2025.11.010

Warchoł, A., Balawejder, M., 2022: The Use of Orthophotomaps to Verify the Network of Agricultural Transport Roads in the Land Consolidation Project, FIG Congress 2022 Volunteering for the Future - Geospatial excellence for a better living, Warsaw, Poland, 2022.

Xue, F., Chang, Y., Xu, W., Liang, W., Sheng, F., Ming, A., 2025: Evidence-Based Real-Time Road Segmentation With RGB-D Data Augmentation, *IEEE Transactions on Intelligent Transportation Systems*, 26(2), pp. 1482-1493, 2025. doi.org/10.1109/TITS.2024.3509140.

Yanuargi, B., Utami, E., 2025: Inception-ResNet-V2 The U-Net Encoder for Road Segmentation using Sentinel 2A, *ComTech: Computer, Mathematics and Engineering Applications*, 16(2), 127–138. doi.org/10.21512/comtech.v16i2.12089.