



**HAL**  
open science

# UAV-based Localization of Removable Urban Pavement Elements Through Deep Object Detection Methods

Iason Katsamenis, Grégory Andreoli, Margarita Skamantzari, Nikolaos Bakalos, Franziska Schmidt, Thierry Sedran, Nikolaos Doulamis, Eftychios Protopapadakis, Dimitris Kalogeras

## ► To cite this version:

Iason Katsamenis, Grégory Andreoli, Margarita Skamantzari, Nikolaos Bakalos, Franziska Schmidt, et al.. UAV-based Localization of Removable Urban Pavement Elements Through Deep Object Detection Methods. PETRA '24: 17th International Conference on PErvasive Technologies Related to Assistive Environments, Jun 2024, Crète, Greece. pp.440-448, 10.1145/3652037.3663934 . hal-04631663

**HAL Id: hal-04631663**

**<https://hal.science/hal-04631663v1>**

Submitted on 2 Jul 2024

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.



Distributed under a Creative Commons Attribution 4.0 International License



# UAV-based Localization of Removable Urban Pavement Elements Through Deep Object Detection Methods

Iason Katsamenis  
Institute of Communications and  
Computer Systems  
Athens, Greece  
iasonkatsamenis@mail.ntua.gr

Grégory Andreoli  
Gustave Eiffel University  
Nantes, France  
gregory.andreoli@univ-eiffel.fr

Margarita Skamantzari  
Institute of Communications and  
Computer Systems  
Athens, Greece  
mskamantz@central.ntua.gr

Nikolaos Bakalos  
Institute of Communications and  
Computer Systems  
Athens, Greece  
bakalosnik@mail.ntua.gr

Franziska Schmidt  
Gustave Eiffel University  
Marne-la-Vallée, France  
franziska.schmidt@univ-eiffel.fr

Thierry Sedran  
Gustave Eiffel University  
Nantes, France  
thierry.sedran@univ-eiffel.fr

Nikolaos Doulamis  
Institute of Communications and  
Computer Systems  
Athens, Greece  
ndoulam@cs.ntua.gr

Eftychios Protopapadakis  
University of Macedonia  
Thessaloniki, Greece  
eftprot@uom.edu.gr

Dimitris Kalogeras  
Institute of Communications and  
Computer Systems  
Athens, Greece  
d.kalogeras@noc.ntua.gr

## ABSTRACT

We introduce a deep learning framework leveraging YOLOv8 architecture to automate the localization of Removable Urban Pavements (RUPs) using UAV imagery. The core idea behind RUPs is to provide pavements that can be quickly opened and closed using lightweight on-site equipment. This approach aims to efficiently restore the street’s original appearance and functionalities within a short timeframe, typically just a few hours. Our study explores the feasibility of autonomously localizing RUP elements, paving the way for robotic-driven replacement with prefabricated, fully functional components. Moreover, the integration of UAV data enhances safety and accessibility to challenging areas. Experimental results underscore the efficacy of our approach in achieving precise localization and thereby enabling proactive maintenance efforts.

## CCS CONCEPTS

• **Computing methodologies** → **Object detection; Scene understanding; Vision for robotics.**

## KEYWORDS

Object detection, Removable Urban Pavements, UAV imagery, Road infrastructure monitoring, Computer vision

## ACM Reference Format:

Iason Katsamenis, Grégory Andreoli, Margarita Skamantzari, Nikolaos Bakalos, Franziska Schmidt, Thierry Sedran, Nikolaos Doulamis, Eftychios Protopapadakis, and Dimitris Kalogeras. 2024. UAV-based Localization of Removable Urban Pavement Elements Through Deep Object Detection Methods. In *The Pervasive Technologies Related to Assistive Environments (PETRA) conference (PETRA '24), June 26–28, 2024, Crete, Greece*. ACM, New York, NY, USA, 9 pages. <https://doi.org/10.1145/3652037.3663934>

## 1 INTRODUCTION

In a strategy aimed at managing efficiently risks with limited budgets, it is crucial to deal with the hazards while deepening our understanding of transportation infrastructure. The goal is to execute maintenance measures that are corrective, preventative, and/or predictive, reducing risks to a level that is deemed acceptable. Thus, the monitoring of pavements, especially in a preventive capacity, plays a pivotal role in the approach to risk analysis. It becomes imperative in this scenario to gain a better grasp of their structural behavior and acquire dependable data for robust risk assessments.

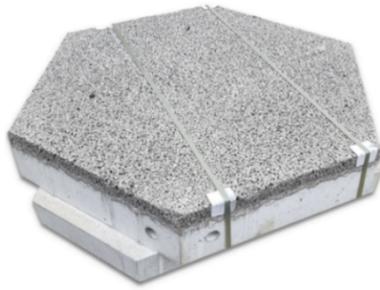
With this aim, the European project HERON strives to merge repair and/or replacement actions with automated identification of potentially defective areas [19]. Embedded within this context is the Forever Open Road (FOR) principle formulated by the Forum of European National Highway Research Laboratories (FEHRL). The French declination of FOR is called “Route de 5ème generation” (meaning 5th generation road) [6], and has led to various research projects, among which the development and optimization of the RUP (Removable Urban Pavements) slabs, which are illustrated in Fig. 1.

As a matter of fact, the H2020 project HERON answers to the three FOR principles: resilience, automation, and adaptability. More specifically, HERON seeks to develop an integrated automated system for road network maintenance and rehabilitation. This includes use cases such as crack sealing, pothole patching, rejuvenation of



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

PETRA '24, June 26–28, 2024, Crete, Greece  
© 2024 Copyright held by the owner/author(s).  
ACM ISBN 979-8-4007-1760-4/24/06  
<https://doi.org/10.1145/3652037.3663934>



**Figure 1: RUP (hexagonal slab).**

asphalt pavement, detection and removal of traffic obstacles, identification of marking defects, autonomous handling of elements in RUPs, and more, with a particular focus on this latter innovative technology. To complete these objectives, the HERON system incorporates an instrumented autonomous terrestrial robotic vehicle (Unmanned Ground Vehicle - UGV) and aerial drones (Unmanned Aerial Vehicle - UAV) for coordinating maintenance and intervention phases. It also includes a detection interface for enhanced structural condition monitoring, augmented reality visualization tools, data processing mechanisms, and workflow coordination utilizing deep learning techniques, such as the You Only Look Once (YOLO) object detection method [20, 42].

Ultimately, HERON facilitates a modular design of the operational system, enhancing capabilities and adaptability for transportation infrastructures [7]. This approach reduces human risks, maintenance expenses, and traffic disturbances, thereby enhancing network efficiency [21, 33]. One of the innovative structures under scrutiny within the project is RUP. Gustave Eiffel University is actively engaged in designing, implementing, testing, and validating RUP on various test sites (see Fig. 2) in France and various load conditions. This facilitates comprehensive trials where the HERON project integrates with a multi-technique Non-Destructive Techniques (NDT) framework.

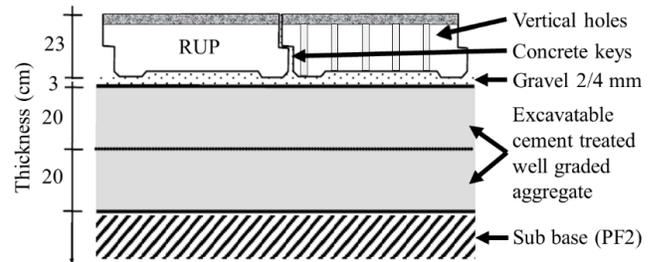


**Figure 2: RUP test site.**

On pavement, anomalies often result from the degradation of underground networks (sewage, electricity, telecommunications, etc.), requiring project owners to intervene. By coordinating all work, we can avoid presenting an inconsistent image of interventions

by different companies and save users from unnecessary expenses. However, experience shows that the best coordination is limited by the policies and time constraints of each stakeholder, each responding to real objectives that are not easily reconcilable with the overall goal. The result of this observation is a pavement that is quickly deformed by repeated openings and closures of trenches, despite the quality of the compaction implementation. Sometimes imperfect repairs no longer ensure complete waterproofing and homogeneity of the entire structure. Consequently, these multiple maintenance interventions have a significant impact on the operating budget.

To resolve this, the RUP emerges as a solution. It is a pavement that can be quickly opened and closed, using lightweight equipment, and allows easy access to the sub-base and the underground networks. This original design is based on hexagonal prefabricated concrete slabs, supported by an original excavatable and draining cement-treated base, all on a supporting platform reminding the multilayer structures of a standard pavement [11]. A straight RUP structure comprises various elements: whole hexagons in the center of the pavement, half-hexagons arranged at the edges, all interconnected with a "concrete key" (shoulder located at the base of the slab), and pavers surrounding the entire structure. Sanded joints seal the gaps between the slabs. Note that the presence of vertical holes allows water drainage (see Fig. 1 and Fig. 3).



**Figure 3: Implementation of RUP on a structure.**

Civil engineering structures, including road infrastructures, are traditionally examined by technicians using rope and harness equipment, along with construction machinery like lifts and cranes. These conventional methods of inspection not only present safety hazards, potentially resulting in worker accidents and injuries, but are also expensive and time-intensive [28]. Moreover, they necessitate the use of bulky machinery, causing disruptions in road and waterway traffic. Damage identification is typically done visually, which could lead to undetected damage in hard-to-reach areas of the structures. As such, it is essential to implement innovative inspection techniques that enhance the efficiency of defect detection while simultaneously ensuring the safety of workers [5, 24, 30].

UAVs present numerous benefits in operations involving the collection of remote sensing data. Specifically, the use of drone technology allows for the safe and remote acquisition of data from areas that are otherwise virtually or physically inaccessible. In parallel, drones enable the efficient collection of timely and on-demand imagery [10], circumventing the need for short-term traffic arrangements that necessitate lengthy permit processes and often result in traffic congestion, shutdowns, accidents, and CO<sub>2</sub> emissions.

Therefore, it is emphasized that UAVs are emerging as a viable and cost-effective means of obtaining high-quality image data, which includes crucial spatial, textural, and chromatic information about the structure being inspected [3, 44, 55].

The deployment of drones will facilitate the surveillance of hard-to-reach road infrastructures. In terms of road maintenance, drones can provide an up-to-date model of defects, enabling the planning of automated actions for subsequent maintenance tasks [29]. This eliminates the need for personnel to conduct visual inspections (which often involve driving vehicles and walking on the road), thereby reducing the risk of accidents [48]. Drone technology has the potential to decrease the overall expense of these costly operations. As a result, the utilization of aerial drones can offer a comprehensive view of the area undergoing maintenance or upgrade interventions.

Simultaneously, the efficient examination of road defects is a vital responsibility for engineers as it contributes to the safety and durability of road infrastructures [41]. Conventional road inspection methods can be laborious, costly, and often lack precision, making it challenging to detect defects before they escalate into significant issues. However, the emergence of deep learning algorithms has introduced a novel and more efficient method for road defect inspection [26]. By processing vast amounts of data, these deep networks can pinpoint, categorize, and localize even the minutest defects, such as cracks or potholes, facilitating focused repairs and more strategic maintenance plans [4, 23, 25, 47]. In this context, this paper assesses the efficacy of a YOLO object detector that leverages UAV imagery to establish an image-based automated system for effective point of interest detection and classification in road networks [35, 36, 39].

The remainder of the paper is organized as follows. In section 2, we provide an overview of related work on object detection related to road infrastructures as well as our contribution to the field. In section 3, we describe the proposed method in detail, as well as its components and architecture. In section 4, we analyze the utilized dataset, training procedure, and experimental results as well as present the performance of our proposed method. Finally, we conclude the paper in section 5 and discuss potential future work in the field of computer vision systems related to the concept of RUPs.

## 2 PREVIOUS WORK

The literature encompasses a variety of notable research endeavors on road infrastructure monitoring through the utilization of deep learning and in particular object detection methodologies. For instance, in [38], a real-time YOLO object detection system applied to the pre-processed radar range-Doppler-angle power spectrum was utilized for single-shot detection and classification of road users.

Another relevant aspect pertains to obstacle recognition in road images [22], where obstacle detection and avoidance for driverless cars were implemented using Convolutional Neural Networks (CNNs) [43]. Additionally, [40] employed a Faster Region-based convolutional neural network to detect and classify on-road obstacles such as vehicles, pedestrians, and animals. Furthermore, in the work of [14] a deep learning model was developed to simulate

driving behaviors by integrating dynamic vehicle information with image data to enhance the performance of self-driving vehicles.

In [46], road surface monitoring with smartphones equipped with GPS and inertial sensors was investigated, leveraging wavelet decomposition analysis for processing inertial sensor signals and Support Vector Machine (SVM) for anomaly detection and classification. Moreover, [32] proposed a system for autonomous and comprehensive road infrastructure condition monitoring, suggesting methodologies that could incorporate automatic ground truth data collection for supervised machine learning. Additionally, [49] focused on a segment-based spatial stratified heterogeneity approach to assess the holistic impacts of variables such as vehicles, climate, road properties, and socioeconomic conditions on pavement infrastructure performance.

Lastly, [8] and [13] delved into the deployment of various Machine Learning (ML) algorithms, including Support Vector Machine, Random Forest, Naïve Bayes, Artificial Neural Networks, and Convolutional Neural Networks, for ML-based pavement evaluation and inspection of image contents for vehicle safety systems purposes, respectively. Nevertheless, the literature lacks any studies focused on the localization and inspection of the novel concept of RUP structures.

## 2.1 Our Contribution

Inspired by the above research work, in the present study, we present the development of an object detection deep learning framework based on the YOLOv8 architecture for the automated localization of RUP slabs by utilizing UAV images. The central idea behind the RUPs is to provide pavements that can be quickly opened and closed using lightweight on-site equipment [11]. This approach aims to efficiently restore the street's original appearance and functionalities within a short timeframe, typically just a few hours. Autonomous localization and inspection of deteriorated RUP elements through state-of-the-art deep learning frameworks can enable robotic-driven autonomous replacement of the slabs with full-functioned components. For instance, when a defected RUP is detected, it could be easily removed by a UGV using a standardized robotic procedure and be replaced by another pre-fabricated slab [34]. Lastly, the integration of artificial intelligence models with UAV imagery not only enhances safety but also provides access to areas that are typically challenging to reach, both virtually and physically [31, 52]. This synergy enables efficient localization and inspection of RUP elements, paving the way for more effective maintenance and monitoring strategies.

## 3 SYSTEM ARCHITECTURE

### 3.1 The RUP concept

In 2020, as part of a mechanical and hydraulic design study on RUP, the Gustave Eiffel University (Nantes, France) constructed a test section measuring 8.51 meters in length and 2.31 meters in width. This section comprises slabs with a 46 cm edge and 23 cm thickness, including a surface layer of 4 cm porous concrete for the 22 double-layer hexagonal slabs with a weight of around 280 kg, 4 cm hydraulic concrete for the 15 half-edge slabs, and 2 single-layer quarter-slabs (see Fig. 4).



**Figure 4: Construction of the test structure at Gustave Eiffel University.**

These slabs have undergone an accelerated test with a fatigue carousel. Since 1984, Gustave Eiffel University has had a system to test pavements under heavy vehicle loads at speeds up to 100 km/h (see Fig. 5). This helps validate new structures by testing for fatigue before moving to larger road tests. So, from November 2020 to February 2021, the RUPs structure was subjected to 200,000 cycles of loading by a half-axle with dual wheels loaded at 65 kN (which corresponds on average to traffic between 45 and 91 buses/day/direction for 20 years). One of the current objectives is to identify potential structural anomalies of RUPs through multi-technique methods on the surface and subsurface [45].



**Figure 5: Fatigue Carousel (Gustave Eiffel University – Nantes, France).**

### 3.2 UAV imagery

The DJI Air 2S (see Fig. 6) represents a notable advancement in consumer drone technology, offering engineer experts an auxiliary tool for effective infrastructure inspection using aerial imagery. More specifically, equipped with a 1-inch CMOS sensor, the drone is capable of capturing high-resolution 20-megapixel stills and recording 5.4K video at 30 frames per second [37, 50]. It is also noted that

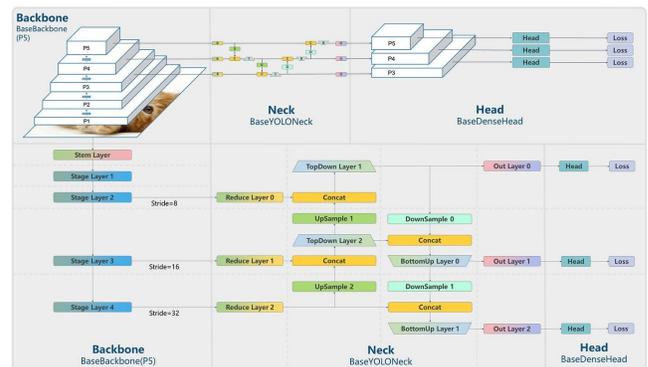
its compact and lightweight design enhances portability, enabling users to easily transport it to various inspection sites. In parallel, the DJI Air 2S offers a range of intelligent flight features and safety functions aimed at improving the overall flying experience. These include advanced obstacle sensors and intelligent flight modes such as Spotlight 2.0, ActiveTrack 4.0, and Point of Interest 3.0, enabling precise tracking of subjects. Lastly, features like AirSense, which alerts users to nearby aircraft, and Return-to-Home ensure safe and reliable operation during flights.



**Figure 6: DJI Air 2S.**

### 3.3 YOLOv8

YOLOv8 was introduced by Ultralytics on January 10th, 2023 [16]. Building upon the groundwork laid by the successful YOLOv5 network, YOLOv8 represents a significant advancement in the fields of object detection, image classification, and instance segmentation [27]. Mirroring its predecessors, YOLOv8 maintains the three primary architectural components characteristic of the YOLO framework, the (i) backbone, (ii) neck, and (iii) head [12, 17, 22]. As one can observe in Fig. 7, these three components collaborate seamlessly, and in particular, the backbone extracts features from the images, the neck combines these features, and the head makes predictions regarding object locations and classes [9].



**Figure 7: YOLOv8 architecture.**

One notable aspect of YOLOv8 is its departure from the anchor-box approach seen in previous YOLO models, instead adopting an anchor-free model [53]. The specific innovative shift enables direct

prediction of an object’s center, addressing challenges associated with anchor boxes such as limited generalization and difficulties in handling irregularities. By reducing the number of box predictions, YOLOv8 improves the speed of the Non-Maximum Suppression (NMS) process, a critical post-processing step responsible for refining candidate detections following inference [15, 27].

Additionally, YOLOv8 introduces architectural enhancements related to convolutions, which are the core components of neural networks. The integration of C2f in place of C3, coupled with substituting the initial  $6 \times 6$  convolution in the stem with a  $3 \times 3$  convolution, has led to a more efficient and adaptable model structure [2, 27]. In this updated design, outputs from the Bottleneck, consisting of two  $3 \times 3$  convolutions with residual connections, are concatenated, departing from the C3 setup where only the output of the final Bottleneck was utilized [1]. This adjustment bolsters the model’s structural robustness, pushing the boundaries of YOLOv8’s capabilities in computer vision tasks.

Continuing its commitment to addressing objects of diverse scales, YOLOv8 integrates the Spatial Pyramid Pooling Feature (SPPF) [54]. Additionally, the model employs an online image augmentation strategy during the training procedure, including mosaic augmentation. This process involves stitching together four images, allowing the network to learn from objects in new locations, partial occlusion scenarios, and against varying surrounding pixels [51].

Acknowledging the requirements of the various different computer vision applications, YOLOv8, like its predecessors, is available in multiple versions that are presented in Table 1 below. In harnessing the capabilities of a drone for various applications, integrating the small version of YOLOv8 presents a compelling proposition [18]. By leveraging the lightweight design and optimized computational efficiency of YOLOv8s, the system can strike a balance between accuracy and computational complexity, making it particularly suitable for real-time applications. This integration enables the drone to swiftly process imagery captured during flight, facilitating rapid object detection and analysis without compromising accuracy.

Model	Size (pixels)	Params (M)	FLOPs (B)
YOLOv8n	640	3.2	8.7
YOLOv8s	640	11.2	28.6
YOLOv8m	640	25.9	78.9
YOLOv8l	640	43.7	165.2
YOLOv8x	640	68.2	257.8

**Table 1: Details of the various YOLOv8 variants.**

## 4 EXPERIMENTAL EVALUATION

### 4.1 Dataset Description

To train and evaluate the proposed methodology, a DJI Air 2S (see Fig. 6), which was described in section 3.2, was utilized, in combination with the DJI Smart Controller (see Fig. 8). More specifically, the experiments took place in the surrounding area of Gustave Eiffel University premises in Bouguenais, near Nantes, France.

To properly acquire the necessary RGB data manual flights were planned and executed, while the image overlap (forward and side), camera angle and flight altitude were properly adjusted during the



**Figure 8: DJI Smart Controller for Air2S.**

flights according to the needs and complexity of the wider area. During the data acquisition the UAV performed three flight strips at 5m, 10m and 15m above the ground covering an area of  $25m \times 10m$ . The RGB data exhibit an aspect ratio of 16:9 and in particular a pixel resolution of  $5,472 \times 3,078$ . An indicative photographic example is illustrated in Fig. 9 below.

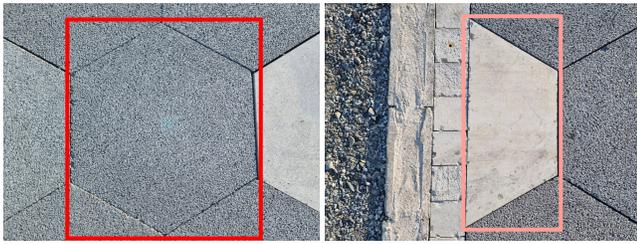


**Figure 9: Indicative image from the utilized dataset.**

It is also highlighted that the images were captured under adverse weather conditions, characterized by heavy rain and strong winds with gusts. To this end, it is worth mentioning that the data acquisition process using the UAV can be challenging or even impossible due to several restrictions. The weather conditions may not make it possible to plan and execute a flight, the vegetation and terrain may also be a restriction since the aircraft is not able to go near any obstacles. Other parameters that should always be taken into consideration are the flight restrictions and regulations of each area (no-fly zones) as well as the flight time limitation of the UAV (approximately 30 minutes).

After the data acquisition procedure, 70 images that depict the RUP structure were manually annotated and verified by engineer experts under the framework of the H2020 HERON project [19]. The dataset was partitioned into three distinct sets for training, validation, and testing, with 70%, 15%, and 15% of the data allocated to each, respectively. During the annotation process, two classes

were considered, based on the shape, material, and color of the slab, as well as its location on the RUP structure. In particular, (i) Slab\_H represents the central slabs characterized by a hexagonal shape and dark gray color (see the red bounding box in Fig. 10a), while (ii) Slab\_T refers to the slabs positioned on the sides of the RUP structure, featuring a trapezoidal shape and light gray color (see the pink bounding box in Fig. 10b). The dataset of this paper is available online at: <https://github.com/ikatsamenis/RUP-UAV/>



(a) Hexagonal slab (Slab\_H). (b) Trapezoidal slab (Slab\_T).

**Figure 10: The two different slab classes.**

### 4.2 Experimental Setup - Model Training

As already analyzed in Section 3, in this study, we leverage the compact architecture of the small version of YOLOv8 (see Table 1) to address the inherent trade-off between accuracy and computational efficiency, particularly in the context of drone-based applications. By incorporating this lightweight network, our system aims to achieve real-time object detection capabilities while ensuring minimal computational burden. This strategic integration enables the drone to efficiently process captured imagery during flight operations, facilitating rapid analysis of the environment without compromising on the accuracy of detected objects. Towards this direction, the deep model takes up less than 25 MB of storage and thus can be easily embedded in low-memory digital devices or systems, including drones and microcontrollers. The YOLOv8s detector was trained and evaluated using an NVIDIA Tesla T4 GPU with 12 GB of memory, provided by Google Colab. We trained the network, using batches of size 16, for 100 epochs, and set the input image resolution to  $640 \times 640$  pixels. Indicative outputs of the utilized YOLOv8s model are illustrated in Fig. 11.

### 4.3 Performance Metrics

The evaluation of the implemented architecture involves analyzing various well-established metrics, detailed in subsequent paragraphs. Before introducing each metric, it's essential to note the use of four key components for assessing deep learning model performance: true positives (TP) for correctly classified positive instances, true negatives (TN) for correctly classified negatives, false positives (FP) for incorrectly classified positives, and false negatives (FN) for incorrectly classified negatives. These parameters form the basis for defining performance metrics, providing insight into the effectiveness of the utilized deep network.

Precision (Prec) measures the proportion of true positive predictions out of all positive predictions made by the model. It assesses

the accuracy of positive identifications, with higher precision indicating fewer false positive predictions. Precision is crucial in scenarios where minimizing false alarms is vital and is calculated as follows:

$$Prec = \frac{TP}{TP + FP} \tag{1}$$

Recall (Rec), also known as sensitivity, quantifies the model's ability to identify all relevant instances, measuring the proportion of true positive predictions out of all actual positives in the dataset. A higher recall indicates that the model captures a greater portion of relevant instances. In applications where missing positive instances is costly, high recall is essential. Its formula is expressed as:

$$Rec = \frac{TP}{TP + FN} \tag{2}$$

F1-score (F1) is the harmonic mean of precision and recall, providing a single metric that balances both precision and recall. It is particularly useful when there is a need to strike a balance between minimizing false positives and false negatives. F1-score is calculated as follows:

$$F1 = \frac{2 \cdot Prec \cdot Rec}{Prec + Rec} = \frac{2 \cdot TP}{TP + \frac{1}{2}(FP + FN)} \tag{3}$$

Average Precision (AP) is a pivotal metric in the domain of object detection, serving as a gauge of a model's precision-recall trade-off for individual object classes. It measures the area under the precision-recall curve, reflecting how effectively a model balances precision and recall across various confidence thresholds. A higher AP score indicates superior performance, showcasing the model's capability to precisely identify objects while minimizing false positives. If we represent the precision-recall curve as  $p(r)$ , then the AP can be computed as follows:

$$AP = \int_0^1 p(r)dr \tag{4}$$

Mean Average Precision (mAP) extends the assessment of AP by computing the average of AP scores across multiple object classes. By aggregating the AP scores, mAP offers a comprehensive evaluation of a model's overall detection performance across diverse classes. A higher mAP signifies a model's proficiency in accurately detecting objects across the entire spectrum of classes, contributing to a more holistic understanding of its capabilities. Therefore, in this study, we compute the mean Average Precision (mAP) for each localization scenario using the following formula, where  $N$  denotes the total number of slab categories:

$$mAP = \frac{1}{N} \sum_{k=1}^N AP_k \tag{5}$$

Lastly, it is noted that in the present paper, we use mAP50 and mAP50:95 which are variants of the mAP metric. These metrics focus on different aspects of object detection performance. On the one hand, mAP50 calculates the average precision at a single Intersection over Union (IoU) threshold of 0.5, providing insight into a model's performance under less strict localization criteria. On the other hand, mAP50:95 computes the average precision over a range



Figure 11: Indicative outputs of the YOLOv8s model.

of IoU thresholds from 0.5 to 0.95, offering a more nuanced assessment of a model’s detection accuracy across a broader spectrum of localization thresholds.

#### 4.4 Experimental Results

As already mentioned, Fig. 11 depicts representative detection outputs of the YOLO network utilized for localizing RUP elements. In parallel, as can be seen in Table 2 the results of YOLOv8s showcase remarkable precision and recall values with average scores over 99%. In particular, average precision is reported at 99.8%, suggesting that nearly all positive predictions made by the model are indeed correct. Similarly, the average recall rate stands at 99.1%, indicating that the network effectively identifies the vast majority of positive instances in the dataset. These high precision and recall scores are further corroborated by mAP metrics, with both mAP50 and mAP50-95 values exceeding 98%. The consistency across different evaluation metrics underscores the robustness of the model in accurately detecting slabs across the entire dataset.

Class	Precision	Recall	F1-score	mAP50	mAP50:95
all	99.8%	99.1%	99.4%	99.5%	98.0%
Slab_H	99.9%	100%	100%	99.5%	98.0%
Slab_T	99.7%	98.2%	98.9%	99.5%	98.1%

Table 2: Performance metrics of the YOLOv8s model.

Further analysis reveals even more impressive performance for individual slab classes. The Slab\_H class exhibits near-perfect precision of 99.9%, coupled with flawless recall of 100%. This implies that the model not only identifies all instances of the hexagonal slabs but also minimizes false positive predictions to an exceptional degree. Similarly, the Slab\_T class demonstrates strong precision of 99.7% and respectable recall of 98.2%, highlighting the detector’s capability to accurately detect instances of this class while maintaining a low false positive rate. The mAP metrics for both classes are consistent with the overall high performance, indicating the model’s ability to maintain precision and recall at different confidence thresholds.

In conclusion, the reported results that are demonstrated in Table 2, as well as the indicative object detection outputs that are depicted in Fig. 11, reflect the high accuracy and reliability of the YOLOv8s model, with exceptional F1-scores across both slab classes. The model’s ability to achieve near-perfect performance for individual classes underscores its effectiveness in accurately identifying objects of interest while minimizing false positives. These results provide confidence in the model’s capability to generalize well to unseen data and showcase its potential for real-world applications in localization tasks that are related to the RUP scenario.

## 5 CONCLUSIONS

In conclusion, our study introduces a novel approach for the automated localization of RUP elements using a deep learning framework based on YOLOv8s architecture and UAV imagery. By leveraging this technology, we aim to streamline the process of effectively localizing the slabs leading to more efficient maintenance and monitoring strategies in the future. Our experimental results demonstrate the feasibility and efficacy of this approach, showcasing its potential to significantly enhance RUP structure inspection and enhance bolster street infrastructure resilience.

Further analysis, research and work will be carried out in order to produce point clouds from the acquired digital images by exploiting Structure from Motion (SfM) and Multi-view Stereo (MVS) reconstruction algorithms to compare them with LiDAR products that will also be acquired in the project's lifetime in order to annotate the necessary data and detect the RUP slabs in the 3D point clouds. Lastly, exploring various state-of-the-art instance segmentation models could present an opportunity to advance the object detection task towards pixel-wise segmentation of the slabs, thereby enhancing the granularity and accuracy of our analysis.

## ACKNOWLEDGMENTS

This work has received funding from the European Union's Horizon 2020 Research and Innovation Programme under grant agreement No 955356 (Improved Robotic Platform to perform Maintenance and Upgrading Roadworks: The HERON Approach)

## REFERENCES

- [1] Armstrong Aboah, Bin Wang, Ulas Bagci, and Yaw Adu-Gyamfi. 2023. Real-time multi-class helmet violation detection using few-shot data sampling technique and yolov8. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5349–5357.
- [2] Sania Akhtar, Muhammad Hanif, and Hamidi Malih. 2023. Automatic Urine Sediment Detection and Classification Based on YoloV8. In *International Conference on Computational Science and Its Applications*. Springer, 269–279.
- [3] Nassim Ammour, Haikel Alhichri, Yakoub Bazi, Bilel Benjdira, Naif Alajlan, and Mansour Zuair. 2017. Deep learning approach for car detection in UAV imagery. *Remote Sensing* 9, 4 (2017), 312.
- [4] Sukhad Anand, Saksham Gupta, Vaibhav Darbari, and Shivam Kohli. 2018. Crackpot: Autonomous road crack and pothole detection. In *2018 digital image computing: techniques and applications (DICTA)*. IEEE, 1–6.
- [5] Deeksha Arya, Hiroya Maeda, Sanjay Kumar Ghosh, Durga Toshniwal, Alexander Mraz, Takehiro Kashiyama, and Yoshihide Sekimoto. 2021. Deep learning-based road damage detection and classification for multiple countries. *Automation in Construction* 132 (2021), 103935.
- [6] Natasha Bahrani, Franziska Schmidt, and Nicolas Hautiere. 2023. A comprehensive study & outlook on integrated structural health monitoring and C-intelligent transport system for pavement condition assessment. *Transportation Research Procedia* 72 (2023), 4231–4238.
- [7] Nikolaos Bakalos, Iason Katsamenis, Eftychios Protopapadakis, Carlos Martín-Portugués Montoliu, Yannis Handanos, Franziska Schmidt, Olov Andersson, Helen Oleynikova, Miquel Cantero, Ilias Gkotsis, et al. 2024. Robotics-enabled roadwork maintenance and upgrading. In *Robotics and Automation Solutions for Inspection and Maintenance in Critical Infrastructures*. 264–286.
- [8] Saül Cano-Ortiz, Pablo Pascual-Muñoz, and Daniel Castro-Fresno. 2022. Machine learning algorithms for monitoring pavement performance. *Automation in Construction* 139 (2022), 104309.
- [9] Edmundo Casas, Leo Ramos, Eduardo Bendek, and Francklin Rivas-Echeverría. 2023. Assessing the effectiveness of YOLO architectures for smoke and wildfire detection. *IEEE Access* (2023).
- [10] Ismael Colomina and Pere Molina. 2014. Unmanned aerial systems for photogrammetry and remote sensing: A review. *ISPRS Journal of photogrammetry and remote sensing* 92 (2014), 79–97.
- [11] François De Larrard, Thierry Sedran, and Jean-Maurice Balay. 2013. Removable urban pavements: an innovative, sustainable technology. *International Journal of Pavement Engineering* 14, 1 (2013), 1–11.
- [12] Gege Guo and Zhenyu Zhang. 2022. Road damage detection algorithm for improved YOLOv5. *Scientific reports* 12, 1 (2022), 15523.
- [13] Malik Haris and Adam Glowacz. 2021. Road object detection: A comparative study of deep learning-based algorithms. *Electronics* 10, 16 (2021), 1932.
- [14] Tsung-Ming Hsu, Cheng-Hsien Wang, and Yu-Rui Chen. 2018. End-to-end deep learning for autonomous longitudinal and lateral control based on vehicle dynamics. In *Proceedings of the 2018 International Conference on Artificial Intelligence and Virtual Reality*. 111–114.
- [15] Muhammad Hussain, Hussain Al-Aqrabi, Muhammad Munawar, Richard Hill, and Tariq Alsbou. 2022. Domain feature mapping with YOLOv7 for automated edge-based pallet racking inspections. *Sensors* 22, 18 (2022), 6927.
- [16] Glenn Joher, Ayush Chaurasia, and Jing Qiu. 2023. YOLO by Ultralytics. (2023).
- [17] Glenn Joher, K Nishimura, T Mineeva, and R Vilarinho. 2020. yolov5. *Code repository* (2020).
- [18] Iason Katsamenis, Nikolaos Bakalos, Eftychios Protopapadakis, Eleni Eirini Karolou, Georgios Kopsiaftis, and Athanasios Voulodimos. 2023. Real time road defect monitoring from UAV visual data sources. In *Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments*. 603–609.
- [19] Iason Katsamenis, Matthaïos Bimpas, Eftychios Protopapadakis, Charalampos Zafeiropoulos, Dimitris Kalogeras, Anastasios Doulamis, Nikolaos Doulamis, Carlos Martín-Portugués Montoliu, Yannis Handanos, Franziska Schmidt, et al. 2022. Robotic maintenance of road infrastructures: The heron project. In *Proceedings of the 15th International Conference on Pervasive Technologies Related to Assistive Environments*. 628–635.
- [20] Iason Katsamenis, Agapi Davradou, Eleni Eirini Karolou, Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, and Dimitris Kalogeras. 2022. Evaluating YOLO transferability limitation for road infrastructures monitoring. In *Novel & Intelligent Digital Systems Conferences*. Springer, 349–358.
- [21] Iason Katsamenis, Nikolaos Doulamis, Anastasios Doulamis, Eftychios Protopapadakis, and Athanasios Voulodimos. 2022. Simultaneous Precise Localization and Classification of metal rust defects for robotic-driven maintenance and pre-fabrication using residual attention U-Net. *Automation in Construction* 137 (2022), 104182.
- [22] Iason Katsamenis, Eleni Eirini Karolou, Agapi Davradou, Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, and Dimitris Kalogeras. 2022. TraCon: A novel dataset for real-time traffic cones detection using deep learning. In *Novel & Intelligent Digital Systems Conferences*. Springer, 382–391.
- [23] Iason Katsamenis, Eftychios Protopapadakis, Nikolaos Bakalos, Andreas Varvarigos, Anastasios Doulamis, Nikolaos Doulamis, and Athanasios Voulodimos. 2023. A Few-Shot Attention Recurrent Residual U-Net for Crack Segmentation. In *International Symposium on Visual Computing*. Springer, 199–209.
- [24] Iason Katsamenis, Eftychios Protopapadakis, Anastasios Doulamis, Nikolaos Doulamis, and Athanasios Voulodimos. 2020. Pixel-level corrosion detection on metal constructions by fusion of deep learning semantic and contour segmentation. In *International Symposium on Visual Computing*. Springer, 160–169.
- [25] Iason Katsamenis, Athanasios Sakelliou, Nikolaos Bakalos, Eftychios Protopapadakis, Christos Klaridopoulos, Nikolaos Frangakis, Matthaïos Bimpas, and Dimitris Kalogeras. 2023. Deep transformer networks for precise pothole segmentation tasks. In *Proceedings of the 16th International Conference on Pervasive Technologies Related to Assistive Environments*. 596–602.
- [26] Rana Khallaf and Mohamed Khallaf. 2021. Classification and analysis of deep learning applications in construction: A systematic literature review. *Automation in construction* 129 (2021), 103760.
- [27] Jun-Hwa Kim, Namho Kim, and Chee Sun Won. 2023. High-speed drone detection based on yolo-v8. In *ICASSP 2023-2023 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)*. IEEE, 1–2.
- [28] Xiao Li, Wen Yi, Hung-Lin Chi, Xiangyu Wang, and Albert PC Chan. 2018. A critical review of virtual and augmented reality (VR/AR) applications in construction safety. *Automation in construction* 86 (2018), 150–162.
- [29] Han Liang, Seong-Cheol Lee, Woosung Bae, Jeongyun Kim, and Suyoung Seo. 2023. Towards UAVs in construction: advancements, challenges, and future directions for monitoring and inspection. *Drones* 7, 3 (2023), 202.
- [30] Jiajing Liu, Hanbin Luo, and Henry Liu. 2022. Deep learning-based data analytics for safety in construction. *Automation in Construction* 140 (2022), 104302.
- [31] Jhonattan G Martinez, Gilles Albeaino, Masoud Gheisari, Raja RA Issa, and Luis F Alarcón. 2021. iSafeUAS: An unmanned aerial system for construction safety inspection. *Automation in Construction* 125 (2021), 103595.
- [32] Johannes Masino, Jakob Thumm, Michael Frey, and Frank Gauterin. 2017. Learning from the crowd: Road infrastructure monitoring system. *Journal of Traffic and Transportation Engineering (English Edition)* 4, 5 (2017), 451–463.
- [33] Qipei Mei and Mustafa Gül. 2020. A cost effective solution for pavement crack inspection using cameras and deep neural networks. *Construction and Building Materials* 256 (2020), 119397.
- [34] Nathan Melenbrink, Justin Werfel, and Achim Menges. 2020. On-site autonomous construction robots: Towards unsupervised building. *Automation in construction* 119 (2020), 103312.

- [35] Nicoletta Nappo, Olga Mavrouli, Francesco Nex, Cees van Westen, Roberto Gambillara, and Alessandro Maria Michetti. 2021. Use of UAV-based photogrammetry products for semi-automatic detection and classification of asphalt road damage in landslide-affected areas. *Engineering geology* 294 (2021), 106363.
- [36] Mingxin Nie and Cheng Wang. 2019. Pavement Crack Detection based on yolo v3. In *2019 2nd international conference on safety produce informatization (IICSPI)*. IEEE, 327–330.
- [37] Janet Mayowa Nwaogu, Yang Yang, Albert PC Chan, and Hung-lin Chi. 2023. Application of drones in the architecture, engineering, and construction (AEC) industry. *Automation in Construction* 150 (2023), 104827.
- [38] Rodrigo Pérez, Falk Schubert, Ralph Rasshofer, and Erwin Biebl. 2019. Deep learning radar object detection and classification for urban automotive scenarios. In *2019 Kleinheubach Conference*. IEEE, 1–4.
- [39] Ping Ping, Xiaohui Yang, and Zeyu Gao. 2020. A deep learning approach for street pothole detection. In *2020 IEEE Sixth International Conference on Big Data Computing Service and Applications (BigDataService)*. IEEE, 198–204.
- [40] Gowdham Prabhakar, Binsu Kailath, Sudha Natarajan, and Rajesh Kumar. 2017. Obstacle detection and classification using deep learning for tracking in high-speed autonomous driving. In *2017 IEEE region 10 symposium (TENSYMP)*. IEEE, 1–6.
- [41] Eftychios Protopapadakis, Iason Katsamenis, and Anastasios Doulamis. 2020. Multi-label deep learning models for continuous monitoring of road infrastructures. In *Proceedings of the 13th ACM International Conference on PErvasive Technologies Related to Assistive Environments*. 1–7.
- [42] Joseph Redmon, Santosh Divvala, Ross Girshick, and Ali Farhadi. 2016. You only look once: Unified, real-time object detection. In *Proceedings of the IEEE conference on computer vision and pattern recognition*. 779–788.
- [43] Nischal Sanil, V Rakesh, Rishab Mallapur, Mohammed Riyaz Ahmed, et al. 2020. Deep learning techniques for obstacle detection and avoidance in driverless cars. In *2020 International Conference on Artificial Intelligence and Signal Processing (AISP)*. IEEE, 1–4.
- [44] S Sankarasrinivasan, E Balasubramanian, K Karthik, U Chandrasekar, and Rishi Gupta. 2015. Health monitoring of civil structures with integrated UAV and image processing system. *Procedia Computer Science* 54 (2015), 508–515.
- [45] Thierry Sedran, Eric Genesseeux, Julien Waligora, Philippe Klein, Mai Lan Nguyen, Julien Cesbron, and Christophe Ropert. 2023. Development of a permeable removable urban pavement. In *ISCR 2013-14th International Symposium on Concrete Roads*.
- [46] Fatjon Seraj, Berend Jan Van Der Zwaag, Arta Dilo, Tamara Luarasi, and Paul Havinga. 2014. RoADS: A road pavement monitoring system for anomaly detection using smart phones. In *International Workshop on Modeling Social Media*. Springer, 128–146.
- [47] Nima Sholevar, Amir Golroo, and Sahand Roghani Esfahani. 2022. Machine learning techniques for pavement condition evaluation. *Automation in Construction* 136 (2022), 104190.
- [48] Luís Augusto Silva, Héctor Sanchez San Blas, David Peral García, André Sales Mendes, and Gabriel Villarubia González. 2020. An architectural multi-agent system for a pavement monitoring system with pothole recognition in UAV images. *Sensors* 20, 21 (2020), 6205.
- [49] Yongze Song, Graeme Wright, Peng Wu, Dominique Thatcher, Tom McHugh, Qindong Li, Shuk Jin Li, and Xiangyu Wang. 2018. Segment-based spatial analysis for assessing road infrastructure performance using monitoring observations and remote sensing data. *Remote Sensing* 10, 11 (2018), 1696.
- [50] Susrutha Babu Sukhavasi, Suparshya Babu Sukhavasi, Khaled Elleithy, Shakour Abuzneid, and Abdelrahman Elleithy. 2021. CMOS image sensors in surveillance system applications. *Sensors* 21, 2 (2021), 488.
- [51] Sanjog Tamang, Biswaraj Sen, Ashis Pradhan, Kalpana Sharma, and Vikash Kumar Singh. 2023. Enhancing covid-19 safety: Exploring yolov8 object detection for accurate face mask classification. *International Journal of Intelligent Systems and Applications in Engineering* 11, 2 (2023), 892–897.
- [52] Chee How Tan, M Ng, Danial Sufiyan Bin Shaiful, Shane Kyi Hla Win, WJ Ang, Sai Kit Yeung, HB Lim, Minhn N Do, and S Foong. 2018. A smart unmanned aerial vehicle (UAV) based imaging system for inspection of deep hazardous tunnels. *Water Practice & Technology* 13, 4 (2018), 991–1000.
- [53] Arpita Vats and David C Anastasiu. 2023. Enhancing retail checkout through video inpainting, yolov8 detection, and deepsort tracking. In *Proceedings of the IEEE/CVF conference on computer vision and pattern recognition*. 5529–5536.
- [54] Kewen Xia, Zhongliang Lv, Chuande Zhou, Guojun Gu, Zhiqiang Zhao, Kang Liu, and Zelun Li. 2023. Mixed receptive fields augmented YOLO with multi-path spatial pyramid pooling for steel surface defect detection. *Sensors* 23, 11 (2023), 5114.
- [55] Sizeng Zhao, Fei Kang, Junjie Li, and Chuanbo Ma. 2021. Structural health monitoring and inspection of dams based on UAV photogrammetry with image 3D reconstruction. *Automation in Construction* 130 (2021), 103832.