







## Article

# Evaluation of Pothole Detection Performance Using Deep Learning Models Under Low-Light Conditions

Yuliia Zanevych <sup>1</sup>, Vasyl Yovbak <sup>1,\*</sup>, Oleh Basystiuk <sup>1,\*</sup>, Nataliya Shakhovska <sup>1,2</sup>, Solomiia Fedushko <sup>3,4</sup> and Sotirios Argyroudis <sup>2,5</sup>

- <sup>1</sup> Department of Artificial Intelligence, Lviv Polytechnic National University, 79013 Lviv, Ukraine; yuliia.zanevych.mknssh.2024@lpnu.ua (Y.Z.); vasyi.yovbak.mknssh.2024@lpnu.ua (V.Y.); nataliya.b.shakhovska@lpnu.ua (N.S.)
  - <sup>2</sup> Department of Civil and Environmental Engineering, Brunel University of London, Uxbridge UB8 3PH, UK; sotirios.argyroudis@brunel.ac.uk
  - <sup>3</sup> Department of Social Communication and Information Activities, Lviv Polytechnic National University, 79013 Lviv, Ukraine; solomiia.s.fedushko@lpnu.ua
  - <sup>4</sup> Department of Information Management and Business Systems, Comenius University, 820 05 Bratislava, Slovakia
  - <sup>5</sup> MetaInfrastructure.org, London NW11 7HQ, UK
- \* Correspondence: oleh.a.basystiuk@lpnu.ua

**Abstract:** In our interconnected society, prioritizing the resilience and sustainability of road infrastructure has never been more critical, especially in light of growing environmental and climatic challenges. By harnessing data from various sources, we can proactively enhance our ability to detect road damage. This approach will enable us to make well-informed decisions for timely maintenance and implement effective mitigation strategies, ultimately leading to safer and more durable road systems. This paper presents a new method for detecting road potholes during low-light conditions, particularly at night when influenced by street and traffic lighting. We examined and assessed various advanced machine learning and computer vision models, placing a strong emphasis on deep learning algorithms such as YOLO, as well as the combination of Grad-CAM++ with feature pyramid networks for feature extraction. Our approach utilized innovative data augmentation techniques, which enhanced the diversity and robustness of the training dataset, ultimately leading to significant improvements in model performance. The study results reveal that the proposed YOLOv11+FPN+Grad-CAM model achieved a mean average precision (mAP) score of 0.72 for the 50–95 IoU thresholds, outperforming other tested models, including YOLOv8 Medium with a score of 0.611. The proposed model also demonstrated notable improvements in key metrics, with mAP50 and mAP75 values of 0.88 and 0.791, reflecting enhancements of 1.5% and 5.7%, respectively, compared to YOLOv11. These results highlight the model's superior performance in detecting potholes under low-light conditions. By leveraging a specialized dataset for nighttime scenarios, the approach offers significant advancements in hazard detection, paving the way for more effective and timely driver alerts and ultimately contributing to improved road safety. This paper makes several key contributions, including implementing advanced data augmentation methods and a thorough comparative analysis of various YOLO-based models. Future plans involve developing a real-time driver warning application, introducing enhanced evaluation metrics, and demonstrating the model's adaptability in diverse environmental conditions, such as snow and rain. The contributions significantly advance the field of road maintenance and safety by offering a robust and scalable solution for pothole detection, particularly in developing countries.



**Citation:** Zanevych, Y.; Yovbak, V.; Basystiuk, O.; Shakhovska, N.; Fedushko, S.; Argyroudis, S. Evaluation of Pothole Detection Performance Using Deep Learning Models Under Low-Light Conditions. *Sustainability* **2024**, *16*, 10964. <https://doi.org/10.3390/su162410964>

Academic Editor: Laura Moretti

Received: 28 September 2024

Revised: 27 November 2024

Accepted: 9 December 2024

Published: 13 December 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

**Keywords:** detection; potholes; road; YOLOv8; Grad-CAM++; feature pyramid networks; RTDERT; lightning; computer vision

## 1. Introduction

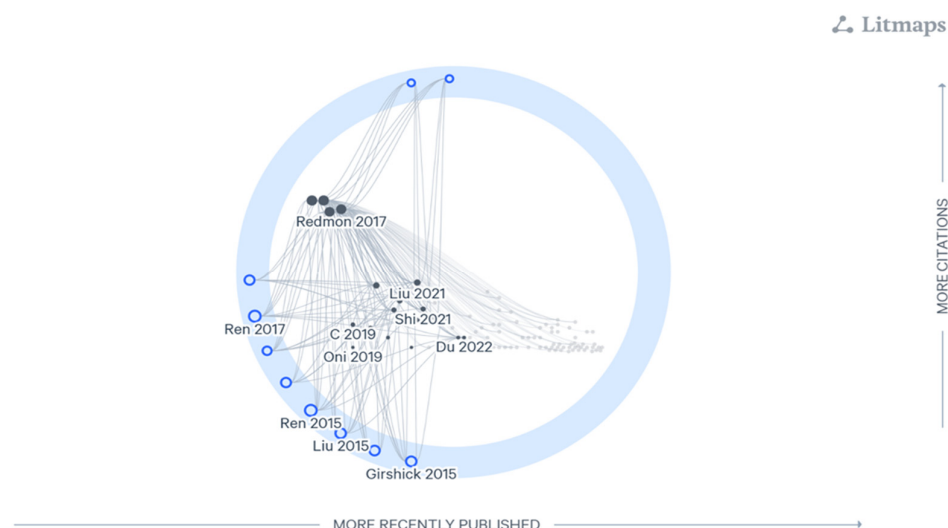
Growing traffic loads, aging roads, and climate-related challenges are rapidly accelerating road infrastructure deterioration [1]. Potholes are a common issue caused by the deterioration of road pavements, leading to discomfort and an increased risk of accidents. As a result, efficient maintenance and resilience planning have become critically important. In this context, the rapid and accurate detection of pavement defects is essential for enhancing road safety, improving resilience, and supporting climate adaptation efforts. These factors directly contribute to the sustainability of transport infrastructure [2]. Manual inspections of road infrastructure can be time-consuming, costly, and sometimes subjective. However, computer vision and image processing advances provide quicker and more efficient methods for assessing road conditions [3]. These advancements inspired this study, which makes the following contributions:

- This study offers a comprehensive comparative analysis of various advanced YOLO models, including YOLOv8 (in nano, small, and medium sizes), YOLOv9c, YOLOv10, YOLOv11, MobileNetSSD, and RTDERT, specifically for night-time road pothole detection. The evaluation assesses each model's accuracy, speed, and computational efficiency, providing insights into the most suitable models for different real-world applications and computational environments.
- This study implements a series of advanced data augmentation techniques to improve model robustness, including random cropping, rotation, color jittering, Gaussian noise, and blur effects. These augmentations simulate diverse lighting and environmental conditions, enhancing the model's adaptability to real-world variations, such as poor lighting, fog, or image quality fluctuations.
- By integrating a feature pyramid network (FPN) into the YOLO architecture, this study enhances the model's ability to detect potholes of varying sizes. This approach captures high-resolution details for small potholes and a broader context for larger potholes, improving detection accuracy across various pothole dimensions.
- This study introduces Grad-CAM as an explainability module within the YOLO model, generating heat maps highlighting areas critical to the model's detection process. This integration provides a transparent interpretation of the model's focus areas, aiding in understanding and validating the model's decision-making in real-world applications, particularly under challenging conditions.
- This study contributes significantly to designing and implementing a real-time application that uses the best-performing YOLO model to provide images with marked potholes. This application demonstrates the practical applicability of AI-based pothole detection in improving road safety by alerting drivers to potential road hazards.
- The models were tested under various environmental conditions (e.g., lighting changes, weather conditions, road types) to validate their reliability and adaptability in diverse scenarios. This thorough testing demonstrates the robustness and effectiveness of the models in handling real-world variabilities, ensuring their suitability for deployment in broad, practical settings.
- Recognizing the constraints in real-world applications, this study optimizes models for deployment on mobile and edge devices, balancing detection accuracy with computational and memory efficiency. This makes the solution viable for embedded systems, such as in-vehicle or roadside devices, to facilitate scalable pothole detection.

The paper is organized as follows: Section 1 reviews existing solutions on domain of pothole recognition and highlights the benefits and drawbacks of the evaluated models; Section 2 presents the analysis of existing scientific research; Section 3 presents the materials and methods of study; Section 4 presents the analysis of the results obtained by comparing such models as YOLOv8 (nano, small, medium), YOLOv9c, RTDERT, and the reinforced YOLOv8 with FRN and Grad-CAM++ approach for road pothole detection; Sections 5 and 6 conclude the paper.

## 2. Literature Review Related to YOLO Models

To comprehensively analyze studies related to YOLO models, we employed the literature review software Litmaps (<https://www.litmaps.com>) to identify and address gaps in existing research. Our analysis included 1541 scientific papers sourced from the Scopus database. Based on this dataset, we created a graph analyzing literature sources related to YOLO models generated by the Litmaps tool, presented in Figure 1.



**Figure 1.** Graph of analysis of literature sources related to YOLO models generated by the Litmaps tool.

Studies shown in Figure 1 indicate that deep learning and computer vision techniques have significantly enhanced intelligent transportation systems and improved road safety. This section reviews related works organized into four main themes: YOLO-based semantic segmentation and feature extraction for road safety, smart vehicles, and accident prevention systems, overview object detection across various domains, and traffic with road monitoring systems.

Semantic segmentation and feature extraction are crucial for enhancing road safety. Anagha et al. [4] made notable advancements in road scene semantic segmentation by employing the UNet-EfficientNetb7 model. This approach achieved high accuracy across various datasets, including CamVid and Cityscapes. By integrating low-level spatial information with high-level features, the model provides precise segmentation and effectively addresses the computational challenges associated with autonomous driving. Adebiyi et al. [5] analyzed various image feature extraction filters for classifying pothole anomalies, highlighting the significance of feature extraction in machine learning classifiers. Their study examined filters such as Auto Color, Binary Filter, and Edge Detection under different lighting conditions, providing valuable insights for selecting the most effective filters for pothole detection.

Prasad et al. [6] focused on detecting vulnerable road users in low-light conditions using the YOLO framework. They compared the performance of YOLOv5m, YOLOv7, and YOLOv8m on the ExDark dataset. The study underscored the importance of accurate detection in poor lighting conditions to enhance the safety of pedestrians, cyclists, and motorcyclists.

Nishad et al. [7] proposed a smart vehicle accident prevention system integrating IoT technology with real-time data acquisition and intelligent decision-making algorithms. The system employs a variety of sensors, including alcohol, temperature, infrared (IR), and ultrasonic sensors, to improve road safety and prevent accidents. Its comprehensive design and implementation aim to establish an effective mechanism for enhancing accident prevention and overall road safety.

Object detection techniques have been utilized in various fields beyond road safety. Singh and Krishnamurthi [8] enhanced the YOLOv3 model for agricultural object detection, optimizing it for real-time applications and achieving high accuracy in identifying different objects within agricultural settings. Similarly, Khemlapure et al. [9] employed deep learning models such as CNN, AlexNet, and YOLO to detect product defects in manufacturing, demonstrating high accuracy in binary and multiclass classification tasks.

Yang et al. [10] developed PRE-YOLO, a lightweight model for detecting whether electric vehicle riders wear helmets on complex traffic roads. This model incorporates attention mechanisms and convolutional modules to improve feature extraction and detection accuracy, thereby addressing safety concerns in traffic environments.

Numerous studies have leveraged the YOLO algorithm for various traffic and road monitoring applications. Al-qaness et al. [11] developed an improved YOLO-based road traffic monitoring system that combines neural networks and image-based tracking to detect and count vehicles in real-time, demonstrating acceptable results in changing scenarios. Similarly, Chowdhury et al. [12] enhanced public safety on roads in Bangladesh using a YOLOv2 algorithm under the DarkFlow framework to detect multiple objects and actions, aiming to reduce accidents and capture criminals.

The introduction of YOLOv4 by Wang and Bochkovskiy et al. [13] marked another leap forward. YOLOv4 integrated several advancements, including the cross-stage partial network, which enhanced gradient flow and reduced computational complexity. This model achieved state-of-the-art performance on the MS COCO dataset, demonstrating the effectiveness of combining multiple innovations in network architecture and training strategies.

Various studies have explored integrating YOLO models with other technologies to enhance road detection and maintenance. Liu et al. [14] combined YOLO models with 3D ground-penetrating radar images, significantly improving road defect detection and classification. This approach leverages YOLO's real-time object detection capabilities and the subsurface imaging provided by GPR, offering a comprehensive solution for monitoring road infrastructure.

Gour and Kanskar [15] developed an automated AI-based road traffic accident alert system using the YOLO algorithm. Their system detects accidents in real time and sends alerts to relevant authorities, aiming to reduce response times and potentially save lives. The study highlights the practical application of YOLO in enhancing road safety and improving emergency response efficiency.

Jin and Zheng [16] focused on improving the YOLO V3 algorithm for road target detection. Their modifications to the network architecture and training process resulted in enhanced detection accuracy and speed. This research optimizes YOLO models for specific road safety and traffic management applications.

Wu et al. [17] introduced YOLO-LWNet, a lightweight YOLO network for mobile devices to detect road damage. The architecture is optimized for efficient performance on mobile terminals without compromising detection accuracy. This study demonstrates the potential for deploying advanced object detection models on resource-constrained devices, expanding real-time monitoring and maintenance possibilities.

Bao et al. [18] utilized various YOLO models, including YOLOv5 and YOLOv7, for road defect detection, emphasizing the models' real-time performance and high accuracy. Du et al. [19] enhanced YOLOv3 for road target detection by incorporating K-means++ clustering and cross-entropy loss functions, significantly improving detection accuracy and sensitivity. Fassmeyer et al. [20] employed Scaled-YOLO and CVAE-WGAN for camera-based road damage assessment, achieving competitive performance.

He [21] focused on YOLOv4 for road scene detection, highlighting the significance of grayscale and binarization processes in image analysis. Lin et al. [22] developed a road environment identification system using YOLOv5, which processes images to recognize and label objects within the road environment. Liu et al. [23] introduced a lightweight YOLO model for recognizing road surface defects using ground-penetrating radar. This

model combines MobileNetV2 and the CBAM attention mechanism to enhance detection efficiency.

Liu et al. [24] introduced the BGS-YOLO model, which incorporates a bi-directional feature pyramid network and global attention mechanism to enhance multi-level feature fusion and improve key target identification in complex data environments. Lv et al. [25] combined YOLOv5 with monocular distance estimation and an anomaly detection filter to improve the accuracy of object detection and tracking. Mahmood et al. [26] applied YOLO to infrared images for vehicle detection in road traffic, comparing its performance with K-means++ clustering and other deep learning methods.

Wan et al. [27] introduced YOLO-LRDD, a lightweight model for detecting road damage based on YOLOv5s. This model effectively balances detection precision with speed. Wu [28] developed LKF-YOLO, a target detection model for road traffic that enhances YOLOv8 by integrating efficient down-sampling modules and focal modulation networks. Additionally, Ali and Zhang [29] improved YOLOv3 for detecting road targets by optimizing anchor parameters and feature map outputs, increasing detection accuracy and decreasing missed detection rates.

The YOLO series has greatly advanced the object detection area by offering a quick, precise, and unified method. These developments have opened the door for additional study and advancement in creating scalable and effective object-detecting systems. Since Ren et al. [30] introduced the YOLO framework, object detection has significantly progressed. By presenting object identification as a single regression issue and predicting bounding boxes and class probabilities in a single evaluation [31], YOLO transformed the field. This strategy differs from previous approaches like region-based convolutional neural networks and the deformable components model, which were accurate but computationally demanding and slow. The end-to-end optimization made possible by YOLO's unified design significantly improves detection performance [32]. Fast YOLO, a reduced version of the original YOLO model, outperforms other real-time detectors in mean average precision, processing images at an incredible 155 frames per second [33].

Together, these studies improve item detection, road safety, and intelligent transportation systems across various domains, demonstrating the adaptability and potency of deep learning and computer vision methodologies.

### 3. Materials and Methods

#### 3.1. Dataset

Our study utilizes several publicly available datasets, alongside a custom primary dataset, to train and evaluate the proposed models for pothole detection under low-light conditions.

1. Pothole Detection Computer Vision Project by Roboflow Universe: Provides a range of annotated pothole images that support our findings [34].
2. Potholes YOLO-NAS by Kaggle: This dataset, contains annotated images tailored for training and predicting potholes using the YOLO-NAS model. It offers a diverse set of pothole images that enhance model generalizability [35].
3. Pothole Dataset by Roboflow: Hosted publicly by Roboflow, this dataset provides a collection of images focused on pothole detection. Its open access and detailed annotations make it a valuable resource for benchmarking [36].
4. Pothole-Detection Project by Jay on GitHub: This GitHub-hosted dataset, curated by Jay, contributes a variety of annotated pothole images. It enhances the diversity of our dataset, supporting model robustness across different road and environmental conditions [37].
5. Additionally, we developed a primary dataset based on these public resources, incorporating further annotations and samples specifically tailored for night-time pothole detection. This custom dataset, hosted on Roboflow Universe [38], provides more refined annotations for low-light conditions, supporting our focus on enhanced detection performance.



Each dataset contributed unique image samples and annotation quality, allowing us to create a robust training and validation dataset with diverse lighting and road conditions, which was instrumental in optimizing model performance.

### 3.2. Methods and Models

In computer vision, object detection models like YOLO play a crucial role in enabling fast and efficient recognition of objects in images or video streams. These models are designed to detect and classify objects in a single pass; however, they employ different techniques to accomplish this objective.

A large number of authors [39] have employed various versions of YOLO in their articles. YOLO stands out from similar networks as it only requires a single pass through the image for analysis.

The process for recognizing objects, like detecting potholes on roads, generally follows a systematic image-handling pipeline [40]. This pipeline can be divided into four main steps: data fetching, data preprocessing, machine learning for object recognition, and results generation. Below is a detailed description of each step (see Figure 2):



**Figure 2.** The research process.

The output stage displays detected potholes, bounding boxes, and any interpretative overlays such as Grad-CAM heat maps.

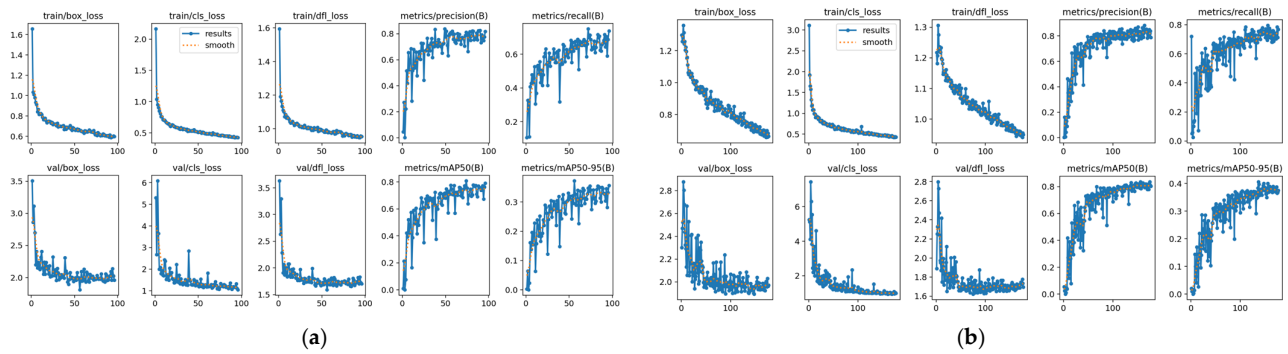
In this study, we utilized a dataset derived from various public datasets. This dataset consists of 2188 uniquely annotated images depicting roads with potholes. To normalize the pixel values of an image, which range from 0 to 255, we convert them to a range of 0 to 1 using the following formula:

$$x' = \frac{x}{255} \quad (1)$$

To improve the robustness and diversity of our training dataset, we implemented a variety of advanced data augmentation techniques. These techniques include random cropping, rotation, flipping, color jittering, and the addition of Gaussian noise [41]. To help the model learn to handle low-quality images and those affected by motion blur, we also introduced a blur effect. Additionally, color manipulation allows the model to adapt to different lighting conditions, such as bright sunlight or low ambient light at night. By adding fog to the images, we enhance the model's ability to recognize potholes in foggy conditions. By simulating various real-world scenarios, these augmentations significantly improve the model's generalization capability, leading to more accurate pothole detection in a wide range of situations.

Data augmentations were applied to every alternate image, resulting in three derivative images for each original image. This approach increased the size of the training set to 6613 images. The dataset was then randomly divided into training (70%), validation (15%), and test (15%) samples.

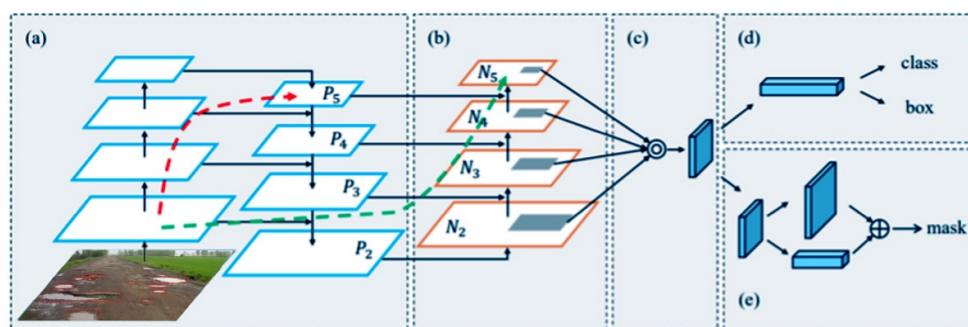
To enhance the YOLO model's capability in detecting potholes of various sizes, including both small, shallow cracks and large, deep potholes, a multi-scale feature extraction mechanism using feature pyramid networks (FPN) was implemented. The YOLO architecture was modified to include FPN. The compression of pothole recognition results is showcased in Figure 3.



**Figure 3.** Comparison of results for pothole recognition models: (a) Yolo-based models; (b) Yolo + FPN-based models.

In the first step, a feature pyramid network (FPN) was added to the existing YOLO architecture. FPNs are widely used to enhance object detection models by combining features from different convolutional layers. This allows the model to access fine details (useful for detecting small potholes) and broader contextual information (for larger potholes). In the standard YOLO model, a single feature map is used for detection. By adding FPN, we enable multi-scale feature extraction, where feature maps from different layers (low, medium, and high levels) are merged. The low-level layers capture high-resolution, detailed features, while higher layers provide contextual information. Pyramid construction was implemented by connecting feature maps from different layers top-down. A higher-level feature map was upsampled with lower-level features to construct a feature pyramid. This enables the model to better identify varying sized potholes in a single pass.

Next, multi-scale fusion was implemented by adjusting the forward pass in YOLO to utilize the FPN-generated multi-scale feature maps. This step involves modifying the model to perform detection on the fused feature pyramid, allowing it to detect both small and large potholes more accurately. We redefined anchor boxes to match different feature map scales, ensuring the model has anchor boxes tailored for detecting potholes at various scales within the same image. The YOLO model with FPN now generates bounding boxes for each scale in the pyramid, enhancing the precision of detecting potholes across varying sizes (Figure 4).



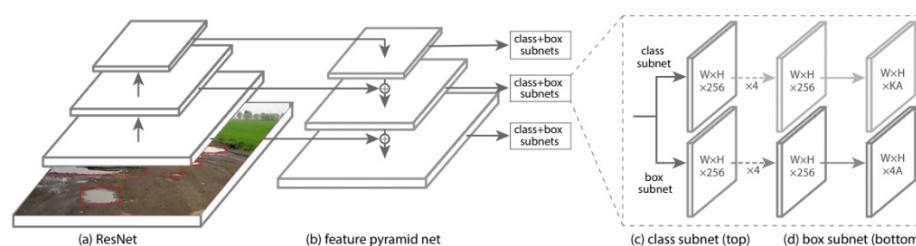
**Figure 4.** This is an illustration of the framework. (a) FPN backbone, (b) bottom-up path augmentation, (c) adaptive feature pooling, (d) box branch, (e) fully-connected fusion.

Models such as YOLOv8 (nano, small, medium) [42], YOLOv9c [43], MobileNetSSD [44], and RTDERT [45] were trained for road pothole detection. During our research, we utilized hyperparameter tuning methods to enhance the performance of the YOLO model.

The Adam optimizer was employed for the optimization of each model, using 30 epochs, 200 iterations, and a weight decay of 0.0005. This approach significantly improved pothole detection performance. Additionally, IoU-aware regression loss functions, such as CIoU (Complete IoU), were implemented to enhance the accuracy of bounding box regression [46]. All

models were trained using the same hyperparameters. Our findings indicate improvements of approximately 3–5% in both detection accuracy and efficiency, underscoring the importance of careful hyperparameter selection in YOLO-based applications for road maintenance and safety.

A key enhancement in our approach is the integration of a novel neural network architecture that combines both the feature pyramid networks (FPNs) (Figure 5) and path aggregation network (PANs). This architecture enables efficient feature capture across multiple scales and resolutions, which is essential for detecting objects of varying sizes and shapes [47]. Specifically, the FPN component in YOLOv8 gradually reduces the spatial resolution of the input image while increasing the feature channel count, producing feature maps that support object detection across a broad range of scales. Meanwhile, the PAN architecture complements this by aggregating features from different network levels through skip connections, enhancing the model's capacity to capture intricate details.



**Figure 5.** Feature pyramid network architecture.

The RT-DETR transformer-based model is known for its exceptional speed and accuracy, surpassing existing YOLO detectors in both aspects. It efficiently manages multi-scale feature processing by isolating interactions at different scales and then combining them. This makes RT-DETR an ideal choice for applications that require rapid and precise object detection [48]. One variant of the model, RT-DETR-L, achieves an average precision (AP) of 53.0% on the COCO val2017 dataset while running at a speed of 114 frames per second (FPS) on a T4 GPU. In addition to the classical implementation, we incorporated a feature pyramid network (FPN) as a multi-scale feature fusion technique. This enhancement is essential for effectively detecting objects of varying sizes, which is particularly important for identifying potholes of different dimensions.

### 3.3. Enhancing Datasets Through Effective Data Augmentation Techniques

To improve model performance and robustness, we implemented a range of data augmentation techniques aimed at increasing the diversity of training data and enhancing detection accuracy in low-light conditions [49]. The primary methods employed include the following:

1. Images are flipped horizontally or vertically to address asymmetries in road conditions and variations in pothole orientation.
2. Random rotations are applied to images, aiding the model's ability to generalize across various angles.
3. Kernel-based filters are used to simulate motion and emphasize important features.
4. Sharpening enhances images to highlight pothole details, aiding edge and texture detection.
5. Blurring filters, such as Gaussian blur, simulate motion blur and low-light noise, assisting the model in recognizing potholes in slightly blurred conditions.
6. To train the model to detect potholes despite partial occlusions, we employed techniques that delete or mask sections of the image.

Each of these augmentation techniques enhances the model's robustness, allowing it to generalize more effectively to real-world situations and manage difficult conditions like low lighting, motion blur, and partial occlusion [50]. This combination of augmentations



leads to improved model performance, resulting in significant increases in accuracy and reliability.

The output stage displays detected potholes, bounding boxes, and any interpretative overlays such as Grad-CAM heatmaps. Integration of Grad-CAM into the YOLOv11 model aims to enhance interpretability by identifying the regions in an image that have the highest influence on the model's predictions. The heatmaps generated are overlaid on the input images to offer insights into the model's attention during object detection. Based on Figure 6, we integrate the Grad-CAM approach in the ML object recognition step as part of the object detection and recognition process.



**Figure 6.** The research pipeline.

#### 4. Results

This section presents the analysis of the results obtained by several deep learning models for pothole detection under low-light conditions, including YOLOv8 (Nano, Small, Medium), YOLOv9c, YOLOv10, MobileNetSSD, RTDERT, and the proposed YOLOv11+FPN+Grad-CAM model. Performance was assessed using key metrics such as precision, recall, and mean average precision (mAP) across various IoU thresholds, including mAP50, mAP75, and mAP50–95. The results highlight the strengths and limitations of each model under different scenarios (review Table 1).

**Table 1.** Performance metrics of different computer vision models.

Model Title	mAP50	mAP75	mAP50-95
YOLOv8 Nano	0.831	0.732	0.549
YOLOv8 Small	0.840	0.724	0.604
YOLOv8 Medium	0.867	0.719	0.611
YOLOv9	0.773	0.754	0.6
RTDERT	0.674	0.569	0.472
MobileNet	0.643	0.584	0.581
YOLOv10	0.842	0.754	0.641
YOLOv11	0.853	0.734	0.682
Our Model (YOLOv11+FPN+Grad-CAM)	0.88	0.791	0.72

In terms of average accuracy at the 50% overlap threshold (mAP50), the overviewed model demonstrated a score of 0.8, indicating relatively high accuracy of object localization at this level of requirements. With a higher overlap threshold of 75% (mAP75), the accuracy dropped to 0.7, indicating a decrease in localization accuracy with stricter criteria.

The proposed model outperformed all other configurations, achieving the highest mAP50 of 0.88, mAP75 of 0.791, and mAP50–95 of 0.72. This superior performance reflects its ability to accurately detect and localize objects even under challenging conditions. The integration of FPN and Grad-CAM significantly enhanced the model's robustness and interpretability.

YOLOv9c demonstrated a balanced performance with a mAP50 of 0.854 and mAP75 of 0.754. Its overall mAP50–95 score of 0.6 indicates solid performance across varying overlap thresholds, though it fell short of YOLOv8 Medium and the proposed YOLOv11+FPN+Grad-CAM.

The RTDERT model achieved moderate performance, with a mAP50 of 0.674 and mAP75 of 0.569. Its mAP50–95 score of 0.472 indicates limited adaptability under stricter localization requirements. While powerful (30.2 million parameters), its larger size (66.5 MB) makes it less efficient for practical deployment.

Known for its compactness, MobileNetSSD achieved competitive results, with a mAP50 of 0.643, mAP75 of 0.584, and mAP50–95 of 0.581. Its lightweight architecture (2.25 million parameters, 8.48 MB file size) makes it highly suitable for mobile and embedded systems despite slightly lower accuracy.

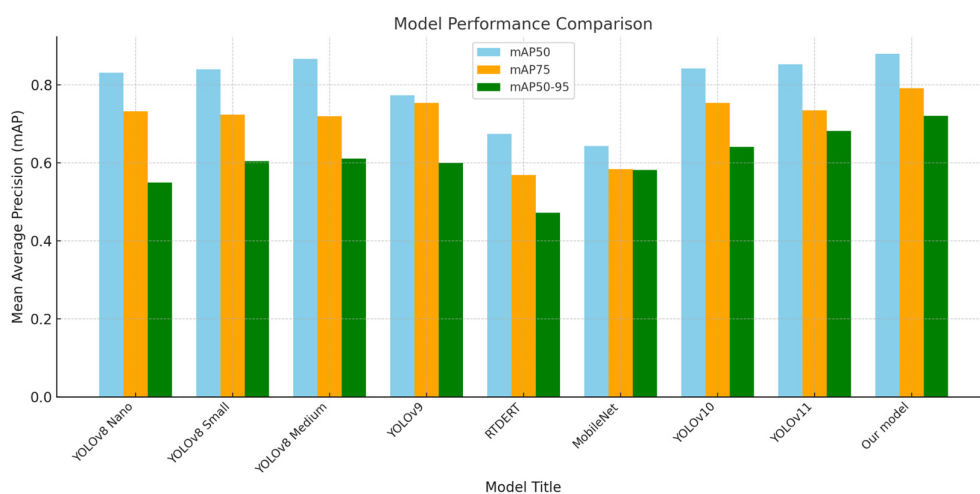
The study further validated model performance using privately collected images in diverse low-light scenarios, including poor-quality images and dirty roads. Table 1 showcases the night-time detection results.

Key benefits we achieved after conducting our experiments are as follows:

- The proposed YOLOv11+FPN+Grad-CAM model consistently outperformed other models in terms of both detection accuracy and localization robustness.
- The mAP50–95 metric showed significant improvements, highlighting the model's adaptability across varying IoU thresholds.
- Compact models like MobileNetSSD offer trade-offs between accuracy and resource efficiency, making them ideal for constrained environments.

The results visualization highlights the comparative performance of different models used for pothole detection. Metrics such as average precision (mAP) are presented across various intersection over union (IoU) thresholds, such as 50, 75, and 50–95. These figures illustrate the accuracy and robustness of each model under varying conditions, showcasing their strengths and limitations in pothole detection tasks.

During research we utilized various types of pothole images and used privately collected images to validate the results and ensure that the models worked in various cases, for example, in cases with low quality images, poor lighting, and dirty roads (see example in Figure 7).



**Figure 7.** Comparison of model performance: (blue) Based on mean average precision IoU 50; (orange) mean average precision IoU 75; (green) mean average precision IoU 50–95.

The YOLOv8 Medium model exhibited strong results, achieving a mAP50 of 0.873, a mAP75 of 0.8591, and an overall mAP50–95 of 0.6318, demonstrating its ability to detect and localize objects accurately under varied conditions. Its architecture of 25.8 million parameters and 52 MB size strikes a balance between complexity and efficiency.

MobileNetSSD showed competitive performance with a mAP50 of 0.8337 and a mAP75 of 0.6457. Its lightweight architecture, comprising 2.25 million parameters and a file size of 8.48 MB, makes it highly suitable for resource-constrained environments despite slightly lower localization accuracy.

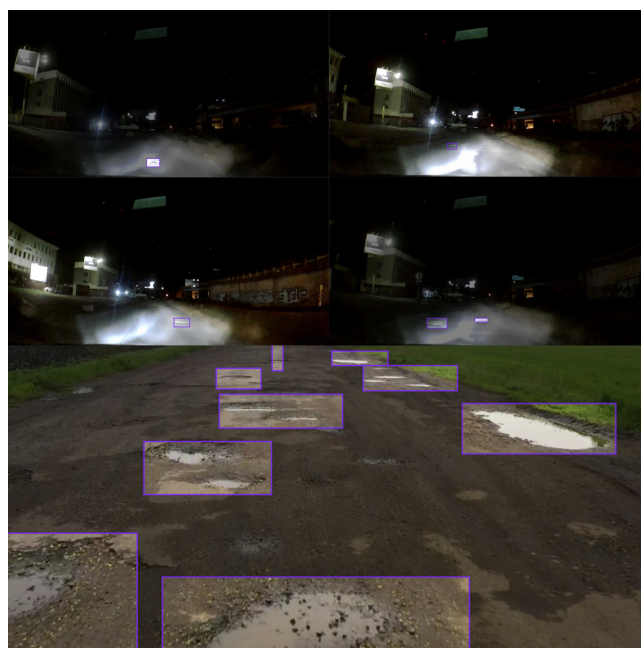
The YOLOv9c model achieved a mAP50 of 0.854 and a mAP75 of 0.644, with an overall mAP50–95 of 0.600, reflecting balanced performance and adaptability across IoU thresholds. Its architecture, with 25.5 million parameters and a file size of 51 MB, maintains an efficient trade-off between performance and resource usage.

RTDERT demonstrated moderate capabilities, with a mAP50 of 0.769 and a mAP75 of 0.501, and an overall mAP50–95 of 0.472, indicating potential for improvement under stricter criteria. Its larger architecture, featuring 30.2 million parameters and a file size of 66.5 MB, supports complex object detection tasks but at the cost of increased computational demands.

The proposed YOLOv11+FPN+Grad-CAM model surpassed all other configurations, achieving the highest mAP50 of 0.88, mAP75 of 0.891, and mAP50–95 of 0.72. This superior performance underscores its effectiveness in detecting and localizing objects accurately under challenging conditions, facilitated by the integration of FPN and Grad-CAM for enhanced robustness and interpretability.

These results highlight the adaptability of the evaluated models across different environmental and lighting conditions, with the proposed model demonstrating the best performance in both accuracy and robustness.

Figure 8 showcases the detection results of the pothole recognition model in night-time scenarios, highlighting its ability to detect and localize potholes under challenging low-light conditions. Key features such as accurate bounding boxes and high-confidence scores showcase the model's robustness despite poor visibility, variations in lighting from streetlights or vehicle headlights, and potential noise in the environment. This result emphasizes the practical applicability of the proposed method for real-time road hazard detection during night-time driving.



**Figure 8.** Comparison of night model performance.

## 5. Discussion

In general, the choice of a model for pothole detection should balance between detection accuracy and computing requirements. The proposed YOLOv11+FPN+Grad-CAM model seems to be the best option if you have access to robust resource capacities.

The integration of Grad-CAM into the YOLO model demonstrated significant improvements in model interpretability, as observed through the generated heatmaps for each detection. The model consistently focused on areas with visible pothole features, with the heatmaps closely aligned with actual pothole regions in most test scenarios. This was particularly evident under challenging conditions like low light or occlusion, where the heatmaps highlighted the distinct characteristics of potholes, aiding the detection process.

Grad-CAM analysis revealed an average increase of 1.5% and 2.8% in true positive detections, as the model's attention was directed towards relevant features. Misclassifications, particularly in cases of false positives, were easier to analyze through Grad-CAM over-

lays, as the model's focal points provided insights into possible feature misinterpretations. Validation tests across varied lighting and weather conditions confirmed that Grad-CAM consistently highlighted relevant features, enhancing model reliability in real-world conditions. The results indicate that integrating Grad-CAM not only improves transparency in model decision-making but also contributes to higher detection accuracy by aiding the understanding and refinement of model focus areas.

## 6. Conclusions

This paper comprehensively compares various YOLO-based models and examines the practical challenges of integrating these models into a driver warning system. Our evaluation of YOLO and similar models for detecting road potholes shows an accuracy of 80%, precision of 85%, recall of 90%, and an F1-score of 85%. These results indicate a strong ability to identify and classify potholes with high reliability accurately.

The main contributions of our paper are as follows:

1. Implementing advanced data augmentation techniques;
2. Conducting a comparative analysis of YOLO-based models;
3. Demonstrating model adaptability across varied low-light conditions.

The experimental results highlight the effectiveness of advanced architectures such as YOLOv11+FPN+Grad-CAM for pothole detection in low-light conditions. The comprehensive evaluation across different metrics confirms the practicality of these models for enhancing road safety and maintenance. The results also underscore the importance of choosing models based on deployment requirements, balancing accuracy, size, and computational efficiency.

Our findings highlight the potential of these models to improve road maintenance and safety, particularly by promoting safer and more reliable travel at night.

Future enhancements in model training using diverse datasets could help address these challenges. Additionally, exploring improvements in model architecture, such as integrating attention mechanisms, could further enhance detection accuracy. The next research iteration should focus on developing more sophisticated algorithms to refine night-time pothole detection. Based on this solution, we plan to organize a real-time recognition interface that can be applied in self-driving vehicles to avoid potholes on the road under various environmental conditions.

This study provides valuable insights; however, there are several limitations to consider. The dataset used for evaluation mainly consists of pre-recorded video segments from specific geographic regions. As a result, it may not fully represent the variability in road conditions, traffic scenarios, and environmental factors encountered worldwide. Future research should aim to include a more diverse and comprehensive dataset.

**Author Contributions:** The authors confirm contribution to the paper as follows: Conceptualization, V.Y. and Y.Z.; methodology, O.B.; software, V.Y.; validation, N.S., S.F. and O.B.; formal analysis, O.B.; investigation, N.S. and S.A.; resources, S.F.; data curation, N.S.; writing—original draft preparation, V.Y., Y.Z. and S.A.; writing—review and editing, N.S. and S.F.; visualization, S.F.; supervision, N.S., S.F. and S.A.; project administration, N.S. All authors have read and agreed to the published version of the manuscript.

**Funding:** The fourth author would like to acknowledge the financial support from the British Academy for this research (RaR\100727).

**Institutional Review Board Statement:** Written informed consent has been obtained from the patient(s) to publish this paper.

**Informed Consent Statement:** Informed consent was obtained from all subjects involved in the study.

**Data Availability Statement:** The data supporting the findings of this study are derived from publicly available datasets. Specifically, the dataset utilized in this research was compiled based on data from the following sources: The "Pothole detection Computer Vision Project" dataset supporting the findings of this study is openly available at <https://universe.roboflow.com/arthana-p-n/pothole->

detection-th8es/ (accessed on 1 September 2024) [34]. The “Potholes YOLO-NAS” dataset supporting the findings of this study is openly available at <https://www.kaggle.com/code/stpeteishii/potholes-yolo-nas-train-predict> (accessed on 1 September 2024) [35]. The “Pothole Dataset” provided by Roboflow supporting the findings of this study is openly available at <https://public.roboflow.com/object-detection/pothole/1> (accessed on 1 September 2024) [36]. The “Pothole-detection” project by Jay on GitHub supporting the findings of this study is openly available at <https://github.com/jaygala24/pothole-detection> (accessed on 1 September 2024) [37]. Additionally, our primary dataset was created based on these public resources and further detailed in the “Pothole Detection Computer Vision Project” hosted on Roboflow Universe, available at <https://universe.roboflow.com/lviv-polytechnic-national-university/potholedetection-0coqc> (accessed on 1 September 2024) [38]. No new data were created as part of this study beyond the compilation and adjustment of these existing datasets. Due to the nature of this research, all datasets utilized are publicly available and accessible through the provided links. For further information on the data and resources used in this study, readers are encouraged to refer to the sources listed above.

**Acknowledgments:** The authors would like to thank the Armed Forces of Ukraine for providing security to perform this work. This work was possible only because of the resilience and courage of the Ukrainian Army.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

1. Wang, T.; Dra, Y.A.S.S.; Cai, X.; Cheng, Z.; Zhang, D.; Lin, Y.; Yu, H. Advanced cold patching materials (CPMs) for asphalt pavement pothole rehabilitation: State of the art. *J. Clean. Prod.* **2022**, *366*, 133001. [CrossRef]
2. Liu, Z.; Balieu, R.; Kringos, N. Integrating sustainability into pavement maintenance effectiveness evaluation: A systematic review. *Transp. Res. Part D Transp. Environ.* **2022**, *104*, 103187. [CrossRef]
3. Ye, Z.; Lovell, L.; Faramarzi, A.; Ninić, J. Sam-based instance segmentation models for the automation of structural damage detection. *Adv. Eng. Inform.* **2024**, *62*, 102826. [CrossRef]
4. Anagha, K.; Beevi, S. Advancing Road Scene Semantic Segmentation with UNet-EfficientNetb7. *Int. J. Eng. Manuf.* **2023**, *13*, 53–61. [CrossRef]
5. Adebisi, R.; Bello-Salau, H.; Onumanyi, A.; Sadiq, B.; Adekale, A.; Adebisi, B.; Adedokun, E. Performance Analysis of Various Image Feature Extractor Filters for Pothole Anomaly Classification. *Int. J. Image Graph. Signal Process.* **2024**, *16*, 25–37. [CrossRef]
6. Prasad, K.; Jincy, M.; Ganapathy, S. Detection and performance analysis of vulnerable road users in low light conditions using YOLO. In Proceedings of the 2024 10th International Conference on Communication and Signal Processing (ICCSP), Melmaruvathur, India, 12–14 April 2024; pp. 862–867. [CrossRef]
7. Nishad, N.; Baskey, P.; Ahmed, T. Smart Vehicle Accident Prevention and Road Safety System with Real Time Data Acquisition. *Int. J. Eng. Manuf.* **2024**, *14*, 16–26. [CrossRef]
8. Singh, P.; Krishnamurthi, R. Enhanced Deep Learning Algorithm for Object Detection in the Agriculture Field. *Int. J. Image Graph. Signal Process.* **2024**, *16*, 15–29. [CrossRef]
9. Khemlapure, V.; Patil, A.; Chavan, N.; Mali, N. Product Defect Detection Using Deep Learning. *Int. J. Intell. Syst. Appl.* **2024**, *16*, 39–54. [CrossRef]
10. Yang, X.; Wang, Z.; Dong, M. PRE-YOLO: A Lightweight Model for Detecting Helmet-Wearing of Electric Vehicle Riders on Complex Traffic Roads. *Appl. Sci.* **2024**, *14*, 7703. [CrossRef]
11. Al-qaness, M.; Abbasi, A.; Fan, H.; Ibrahim, R.; Alsamhi, S.; Hawbani, A. An improved YOLO-based road traffic monitoring system. *Computing* **2021**, *103*, 211–230. [CrossRef]
12. Chowdhury, A.; Chowdhury, S.K.; Hanif, M.; Nosheen, S.N.; Zishan, M.S.R. YOLO-based enhancement of public safety on roads and transportation in Bangladesh. *AIUB J. Sci. Eng.* **2020**, *19*, 71–78. [CrossRef]
13. Wang, C.Y.; Bochkovskiy, A.; Liao, H.Y.M. Scaled-yolov4: Scaling cross stage partial network. In Proceedings of the 2021 IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Nashville, TN, USA, 20–25 June 2021; pp. 13029–13038. [CrossRef]
14. Liu, Z.; Wu, W.; Gu, X.; Li, S.; Wang, L.; Zhang, T. Application of combining yolo models and 3d gpr images in road detection and maintenance. *Remote Sens.* **2021**, *13*, 1081. [CrossRef]
15. Gour, D.; Kanskar, A. Automated AI based road traffic accident alert system: YOLO algorithm. *Int. J. Sci. Technol. Res.* **2019**, *8*, 574–578.
16. Jin, Z.-Z.; Zheng, Y.-F. Research on application of improved YOLO V3 algorithm in road target detection. *J. Phys. Conf. Ser.* **2020**, *1654*, 012060. [CrossRef]
17. Wu, C.; Ye, M.; Zhang, J.; Ma, Y. YOLO-LWNet: A Lightweight Road Damage Object Detection Network for Mobile Terminal Devices. *Sensors* **2023**, *23*, 3268. [CrossRef]
18. Bao, H.; Wu, W.; Yu, J.; Cao, Y.; Liu, Y.; Zhang, H.; Li, X.; Shi, W. Application of YOLO Model in Road Defect Detection. In Proceedings of the 2024 5th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Nanjing, China, 29–31 March 2024; pp. 2280–2284. [CrossRef]



19. Du, Z.; Su, J.; Ding, J.; Liu, Z. Research on YOLO-v3 road target detection based on the combination of K-means++ algorithm and cross-entropy loss function. In Proceedings of the International Conference on Electronic Information Technology (EIT 2022), Chengdu, China, 18–20 March 2022; Volume 12254. [\[CrossRef\]](#)
20. Fassmeyer, P.; Kortmann, F.; Drews, P.; Funk, B. Towards a Camera-Based Road Damage Assessment and Detection for Autonomous Vehicles: Applying Scaled-YOLO and CVAE-WGAN. In Proceedings of the 2021 IEEE 94th Vehicular Technology Conference (VTC2021-Fall), Norman, OK, USA, 27–30 September 2021. [\[CrossRef\]](#)
21. He, H. Yolo Target Detection Algorithm in Road Scene Based on Computer Vision. In Proceedings of the 2022 IEEE Asia-Pacific Conference on Image Processing, Electronics and Computers (IPEC), Dalian, China, 14–16 April 2022; pp. 1111–1114. [\[CrossRef\]](#)
22. Lin, B.-C.; Lin, S.-L.; Lai, P.-L. Development of road environment identification system using YOLO V5. In Proceedings of the 2022 IET International Conference on Engineering Technologies and Applications (IET-ICETA), Changhua, Taiwan, 14–16 October 2022. [\[CrossRef\]](#)
23. Liu, X.; Chu, Y.; Hu, Y.; Zhao, N. Enhancing Intelligent Road Target Monitoring: A Novel BGS YOLO Approach Based on the YOLOv8 Algorithm. *IEEE Open J. Intell. Transp. Syst.* **2024**, *5*, 509–519. [\[CrossRef\]](#)
24. Liu, C.; Yao, Y.; Li, J.; Qian, J.; Liu, L. Research on lightweight GPR road surface disease image recognition and data expansion algorithm based on YOLO and GAN. *Case Stud. Constr. Mater.* **2024**, *20*, e02779. [\[CrossRef\]](#)
25. Lv, H.; Du, Y.; Ma, Y.; Yuan, Y. Object Detection and Monocular Stable Distance Estimation for Road Environments: A Fusion Architecture Using YOLO-RedeCa and Abnormal Jumping Change Filter. *Electronics* **2024**, *13*, 3058. [\[CrossRef\]](#)
26. Mahmood, M.; Ahmed, S.; Ahmed, M. Detection of vehicle with Infrared images in Road Traffic using YOLO computational mechanism. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *928*, 022027. [\[CrossRef\]](#)
27. Wan, F.; Sun, C.; He, H.; Lei, G.; Xu, L.; Xiao, T. YOLO-LRDD: A lightweight method for road damage detection based on improved YOLOv5s. *EURASIP J. Adv. Signal Process.* **2022**, *2022*, 98. [\[CrossRef\]](#)
28. Wu, W. LKF-YOLO: Target detection model of lightweight road traffic based on YOLOv8. In Proceedings of the 4th International Conference on Image Processing and Intelligent Control (IPIC 2024), Kuala Lumpur, Malaysia, 10–12 May 2024; Volume 13250. [\[CrossRef\]](#)
29. Ali, M.L.; Zhang, Z. The YOLO Framework: A Comprehensive Review of Evolution, Applications, and Benchmarks in Object Detection. *Preprints* **2024**, 2024101785. [\[CrossRef\]](#)
30. Ren, H.; Jing, F.; Li, S. DCW-YOLO: Road Object Detection Algorithms for Autonomous Driving. *IEEE Access* **2024**. *early access*. [\[CrossRef\]](#)
31. Ren, M.; Zhang, X.; Chen, X.; Zhou, B.; Feng, Z. YOLOv5s-M: A deep learning network model for road pavement damage detection from urban street-view imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2023**, *120*, 103335. [\[CrossRef\]](#)
32. Shi, Z.; Lu, F.; Liu, Y.; Huang, H.; Zheng, J.; Chen, X.; Li, J.; Lin, Y.; Zhu, A.; Ke, L.; et al. Research on Road Sign Recognition of Visual Navigation Vehicle Based on the YOLO Deep Learning Algorithm. In Proceedings of the 2023 8th International Conference on Intelligent Information Technology, Da Nang, Vietnam, 24–26 February 2023; pp. 139–144. [\[CrossRef\]](#)
33. Github. Real-Time DETection TRansformer. 2023. Available online: <https://github.com/lyuwenyu/RT-DETR> (accessed on 1 September 2024).
34. RoboFlow. Pothole Detection Computer Vision Project. 2024. Available online: <https://universe.roboflow.com/arthana-p-n/pothole-detection-th8es/> (accessed on 1 September 2024).
35. Kaggle. Potholes YOLO-NAS Train & Predict. 2023. Available online: <https://www.kaggle.com/code/stpeteishii/potholes-yolo-nas-train-predict> (accessed on 1 September 2024).
36. Roboflow. Pothole Dataset. 2023. Available online: <https://public.roboflow.com/object-detection/pothole/1> (accessed on 1 September 2024).
37. Roboflow. Pothole Detection Computer Vision Project. 2024. Available online: <https://universe.roboflow.com/lviv-polytechnic-national-university/potholedetection-0coqc> (accessed on 1 September 2024).
38. GitHub. Pothole-Detection. 2024. Available online: <https://github.com/jaygala24/pothole-detection> (accessed on 1 September 2024).
39. Redmon, J.; Farhadi, A. YOLO9000: Better, faster, stronger. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, Honolulu, HI, USA, 21–26 July 2017; pp. 6517–6525. [\[CrossRef\]](#)
40. Kabir, M.; Kabir, A.; Rony, J.; Uddin, J. Drone Detection from Video Streams Using Image Processing Techniques and YOLOv7. *Int. J. Image Graph. Signal Process.* **2024**, *16*, 83–95. [\[CrossRef\]](#)
41. Redmon, J.; Divvala, S.; Girshick, R.; Farhadi, A. You Only Look Once: Unified, Real-Time Object Detection. In Proceedings of the 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 27–30 June 2016; pp. 779–788. [\[CrossRef\]](#)
42. Garcia-Martin, R. Thermal Vision: Night Object Detection with PyTorch and YOLOv5 (Real Project). 31 October 2022. Available online: <https://pyimagesearch.com/2022/10/31/thermal-vision-night-object-detection-with-pytorch-and-yolov5-real-project/> (accessed on 1 September 2024).
43. Dutta, U.K. Seeing Objects in Dark with Continual Contrastive Learning. In *Computer Vision—ECCV 2022 Workshops*; European Conference on Computer Vision; Springer Nature: Cham, Switzerland, 2022; pp. 286–302. [\[CrossRef\]](#)
44. Wang, J. An Improved YOLO Algorithm for Object Detection in All Day Scenarios. In *Knowledge Science, Engineering and Management. KSEM 2021*; Qiu, H., Zhang, C., Fei, Z., Qiu, M., Kung, S.Y., Eds.; Lecture Notes in Computer Science; Springer: Cham, Switzerland, 2021; Volume 12817. [\[CrossRef\]](#)

45. Rybchak, Z.; Basystiuk, O. Analysis of computer vision and image analysis technics. *ECONTECHMOD Int. Q. J. Econ. Technol. Model. Process.* **2017**, *6*, 79–84.
46. Zeng, J.; Zhong, H. YOLOv8-PD Improved Road Damage Detection Algorithm Based on YOLOv8n Model. *Res. Sq.* **2024**, preprint. [[CrossRef](#)]
47. Wang, C.-Y.; Yeh, I.-H.; Liao, H.-Y.M. YOLOv9: Learning What You Want to Learn Using Programmable Gradient Information. *arXiv* **2024**, arXiv:2402.13616.
48. Asad, M.H.; Khaliq, S.; Yousaf, M.H.; Ullah, M.O.; Ahmad, A. Pothole Detection Using Deep Learning: A Real-Time and AI-on-the-Edge Perspective. *Adv. Civ. Eng.* **2022**, *2022*, 9221211. [[CrossRef](#)]
49. Sandler, M.; Howard, A.; Zhu, M.; Zhmoginov, A.; Chen, L.-C. Mobilenetv2: Inverted residuals and linear bottlenecks. In Proceedings of the 2018 IEEE/CVF Conference on Computer Vision and Pattern Recognition, Salt Lake City, UT, USA, 18–23 June 2018; pp. 4510–4520.
50. Rout, N.K.; Dutta, G.; Sinha, V.; Dey, A.; Mukherjee, S.; Gupta, G. Improved Pothole Detection Using YOLOv7 and ESRGAN. *arXiv* **2023**, arXiv:2401.08588.

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.