

Simulation of Environment Recognition Systems for Autonomous Vehicles in CARLA Simulator [†]

Dávid Gorza ^{1,*}, Gábor Saly ^{1,*}  and Dániel Csikor ² 

¹ Department of Road and Rail Vehicles, Széchenyi István University, H-9026 Győr, Hungary; gorza.david@hallgato.sze.hu

² Vehicle Industry Research Center, Széchenyi István University, H-9026 Győr, Hungary; csikor.daniel@ga.sze.hu

* Correspondence: saly.gabor@sze.hu

[†] Presented at the Sustainable Mobility and Transportation Symposium 2025, Győr, Hungary, 16–18 October 2025.

Abstract

Towards the introduction of autonomous vehicles, studying their functionality is becoming increasingly important. Detecting the environment in a self-driving vehicle is a very complex issue. The combination of different sensors is essential for safe and reliable operation. Detection enables the vehicle to accurately recognize and track surrounding objects, understand changes in the dynamic environment, and adapt to different situations. Improving environmental sensing and object recognition is essential for the widespread deployment of self-driving vehicles. In addition to real-world tests, simulation environments provide an opportunity to investigate the operation of autonomous vehicles. Simulations are cost-effective methods for examining the processing of information from the vehicle environment and identifying the current limitations and problems of these technologies. In the CARLA simulator environment, object detection is reproduced in realistic traffic situations. Based on the results, the detection performance was analyzed using interference matrices, F1 scores, accuracy, and coverage metrics.

Keywords: autonomous vehicle; simulation; object detection



Academic Editors: Szabolcs
Kocsis-Szürke and Boglárka
Eisinger Balassa

Published: 3 November 2025

Citation: Gorza, D.; Saly, G.; Csikor, D. Simulation of Environment Recognition Systems for Autonomous Vehicles in CARLA Simulator. *Eng. Proc.* **2025**, *113*, 30. <https://doi.org/10.3390/engproc2025113030>

Copyright: © 2025 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/>).

1. Introduction

In recent decades, humanity has had to contend with global population growth and urbanization processes. At present, almost 75% of the European population lives in cities, and this ratio is expected to rise to approximately 84% by 2050. These processes may pose new challenges for sustainable urban growth. Smart urban development strategies focus on technological and social solutions aimed at improving the quality of urban life, reducing environmental impact, and increasing transport safety [1]. Intelligent mobility solutions could be key elements in these efforts, as they have the potential to reduce the number of accidents caused by human factors, manage traffic congestion, and contribute to reducing transport-related environmental impacts [2]. Based on research results, the majority of road accidents are attributable to human error. The most prominent sources of danger include speeding, inattention, drunk driving, and failure to use passive safety devices. The Vision Zero principle aims to create a transport system that can tolerate human error without leading to serious or fatal consequences. In this systematic approach, injury prevention is ensured through the use of infrastructure and vehicle technology [3].

Advanced driver assistance systems (ADASs) operate on the basis of data collected by various sensors integrated in vehicles [4]. An investigation of the road safety effects of ADAS technologies in different driving contexts, based on accident statistics from the United Kingdom. The research analyzed different driving environments, including urban, rural, and highway road types, as well as clear, rainy, foggy, snowy, and stormy weather conditions, and daytime and nighttime lighting conditions. Based on the results of the analysis, the widespread use of ADAS across the entire vehicle fleet could reduce the number of road accidents by 23.8% annually [5].

The introduction of autonomous vehicles into the transport system could have an impact on further improving road safety. These vehicles eliminate the risks of human errors and can potentially reduce the number of fatal and serious accidents. Their safe operation is closely related to the reliability of sensors, the accuracy of environmental perception, and the decision-making capabilities of algorithms. The introduction of these vehicles will require an interdisciplinary approach that takes into account the characteristics of all components of the system, including people, machines, and infrastructure [6]. The fundamental task of autonomous vehicles is to continuously detect and interpret their environment. They use sensors and algorithms that recognize road obstacles, other road users, vehicles, pedestrians, traffic signs, traffic lights, and road markings. Based on the collected information, the vehicle is required to make fast and reliable decisions that ensure smooth travel in a dynamic traffic environment [7].

LiDAR sensors provide high-precision, high-resolution 3D mapping, which helps vehicles detect objects and measure distances. They are capable of detecting static and moving objects in detail, even at long distances. However, unfavorable weather conditions such as rain, fog, or snow can significantly reduce LiDAR sensitivity, as the light beams can be absorbed or distorted. Radar sensors are similarly suited to detecting moving objects and accurately measuring distances and velocities. The technology is particularly advantageous in fog, rain, or darkness. However, radar cannot provide as high-resolution information as LiDAR, making it less suitable for detailed object recognition. Cameras provide visual information for recognizing traffic signs, lane markings, signal lights, and pedestrians. The images enable the identification of the color, shape, and texture of objects, which serve as the basis for machine vision algorithms. However, these sensors are sensitive to lighting changes, such as low light conditions, nighttime conditions, or strong sunlight. Sensor fusion is responsible for the coordinated operation of the aforementioned sensors [8].

Image processing is used to improve image quality, extract information, and support automated decision-making in autonomous vehicle systems. During preprocessing, image data is prepared, which includes noise reduction, contrast enhancement, and correction of geometric distortions to improve the visibility of relevant information. Object segmentation involves dividing images into regions to separate individual objects [9].

Dursun et al. focused on the real-time recognition of traffic signs and traffic lights using the YOLOv3 (You Only Look Once) algorithm. A unique dataset consisting of images of traffic sign models under various lighting conditions was created. The system was trained based on these images. The system became capable of providing reliable recognition in real-time environments, which is necessary to support the safe operation of autonomous systems [10]. Priya et al. presented a real-time, integrated system that detected and tracked vehicles, traffic signs, and signals using the YOLOv8 algorithm. YOLOv8 was able to perform detection and classification in a single step during image processing. This ensured high speed and accuracy. The system was complemented by a convolutional neural network-based classification. The method was trained on open-source datasets, and its performance was evaluated using precision, F1 score, and mAP metrics, confirming its effectiveness in various traffic situations [11]. Barrozo and Lazcano presented a simulated

control system for an autonomous vehicle designed for off-road environments, where navigation is based solely on visual information, specifically road surface recognition. All decisions are made based on the ratio of road pixels within the image field. The results showed that the approach could be more effective than edge detection-based strategies, especially in low-texture environments [12]. A distinctive contribution of this study is the combined evaluation of object detection performance across multiple object classes and environmental conditions, such as varying weather and lighting scenarios.

Most previous studies focused either on evaluating the effect of a single environmental condition across multiple object classes [13,14] or on the detection performance of a specific class under varying conditions [15,16]. This research integrates both aspects into a unified experimental setup. This approach enables a more realistic and comprehensive understanding of perception challenges in autonomous driving. The resulting framework offers a valuable baseline for developing and benchmarking detection systems under diverse and complex traffic situations.

2. Materials and Methods

The development of environmental perception systems for autonomous vehicles requires a realistic, flexibly customizable simulation platform that supports sensor integration and the modeling of complex traffic situations. Three well-known tools were examined: CARLA, LGSVL, and AirSim. CARLA and LGSVL offer detailed, real-time sensor modeling and urban environments, while AirSim is primarily designed for drone simulation, with more limited on-road vehicle simulation capabilities. Based on a comparison of functionality, flexibility, and realism, CARLA proved to be the most suitable for testing object tracking systems for autonomous vehicles [17–19]. CARLA (Car Learning to Act) is an open source simulation environment designed specifically for autonomous vehicle research and development. It is built on the Unreal Engine graphics engine, enabling detailed modeling of realistic urban environments and traffic situations. The platform is able to support real-time data generation from camera, LIDAR, radar, and GPS sensors and their joint operation, enabling the testing of sensor fusion algorithms under various environmental conditions such as rain, fog, daytime, or nighttime lighting. CARLA provides full control over vehicle movements, traffic scenarios, pedestrian behavior, and weather conditions. The platform is suitable for creating complex and interactive simulation scenarios. The movement of all dynamic and static actors can be controlled, facilitating the testing of decision-making algorithms for autonomous systems [17].

The simulation data were generated using CARLA version 0.9.14, which allows data from various sensors to be recorded. The maps used, Town02 and Town05, are available on the official CARLA website and provide a realistic and detailed urban structure. To generate traffic, not only autonomous vehicles but also simulation-controlled vehicles and pedestrians, as uncontrolled actors, were created to ensure realistic traffic scenarios. During the tests, the current speed, position, and timestamp were saved for each frame. While annotating, YOLO-format .txt files were automatically generated for the images, containing the object class and normalized bounding box coordinates. The classes examined were car, motorcycle, bicycle, traffic light, traffic sign, and background, for which no label was created, but it was taken into account during the evaluation. During the tests, YOLOv7, a single-stage object detection network, was used, processing the entire input image at once, enabling fast and efficient recognition. The network consists of three main layers within a single forward convolutional architecture. The backbone network extracts visual features such as edges, textures, and shapes from the input image using various convolution and pooling layers. The neck network is a multi-level feature processing layer that facilitates the detection of objects of different sizes. The head network is the final component, which

divides the image into a grid and examines the probability of the object's presence in each cell. Object recognition is not simply the selection of pixels, but a statistical decision about the content of image segments [20,21].

To test the model under challenging conditions, two weather presets were used in CARLA. The critical weather setup included 90% cloudiness, 90% precipitation, 80% precipitation deposits, 50% wind intensity, 90% fog density, and 100% surface wetness, simulating an extreme low-visibility scenario with a low sun angle (-5°). In addition, a nighttime scenario was created with 80% cloudiness, 100% precipitation, and a sun altitude angle of -10° , representing heavy rain in near-darkness. These configurations aimed to evaluate detection performance under high-variance environmental conditions. Figure 1 is taken from the simulation of the latter case.



Figure 1. Screenshot from the object recognition simulation in a rainy city environment at night.

The YOLOv7 model was used from an open source implementation (source: <https://github.com/WongKinYiu/yolov7> (accessed on 14 November 2024)), which is PyTorch-1.11.0 based and easily configurable for unique data. Based on the images and annotations extracted from CARLA, a YOLO-compatible data structure was created, which is summarized in a data.yaml configuration file. The file contains a list of classes and the paths to the training and validation datasets. The dataset containing 779 images was divided into 80–20% for training and validation. Training was performed in a GPU environment over 150 episodes, monitoring the evolution of loss functions, which consist of three parts: localization error (bounding box accuracy), classification error, and object presence error. During optimization, we followed the metrics of mean average precision (mAP@0.5) and mean precision at various IoU thresholds (mAP@0.5:0.95), which show the overall and detailed recognition performance of the model. The limited size of the dataset and the uneven distribution of object classes probably contributed to the instability observed in the accuracy and recall performance discussed below, especially for smaller or less frequent categories such as traffic signs and traffic lights. The simulation was performed using an ASUS ROG Strix G15 G512 notebook (Taipei, Taiwan) equipped with an Intel® Core™ i5-10300H processor (2.5 GHZ, 4 cores) (Santa Clara, CA, USA), NVIDIA® GeForce® GTX 1650 Ti 4GB GDDR6 graphics card, and 16GB DDR4 memory (Santa Clara, CA, USA).

3. Results

This section analyzes the model's performance across different object classes. A detailed examination was conducted in terms of recognition accuracy, F1 score, precision, and recall values. The confusion matrix was used to map the classes that the model recognizes accurately and where frequent errors occur, such as the misidentification of smaller objects as background. The results provide a comprehensive overview of the strengths and weaknesses of the object detection system and information on further development opportunities.

Figure 2 displays a confusion matrix illustrating the model's classification performance across different object classes. The best-performing category is vehicle, which the model correctly identifies in 66% of cases. However, in one-third of cases (33%), it fails to detect

the object and ignores it as background. The motorcycle class performs similarly, with 58% correct classification. The model performs worse in the case of bicycles, correctly recognizing only 41% of objects, while treating the remaining 59% as background. The performance of the traffic light and, in particular, the traffic sign classes declines further. Traffic lights have a correct classification rate of 38%, while traffic signs have a rate of only 8%, indicating that the model treats this class almost entirely as background. The last column of the matrix shows the false positive rate. The vehicle class stands out with a value of 0.56, which means that the model often detects vehicles where there are none.

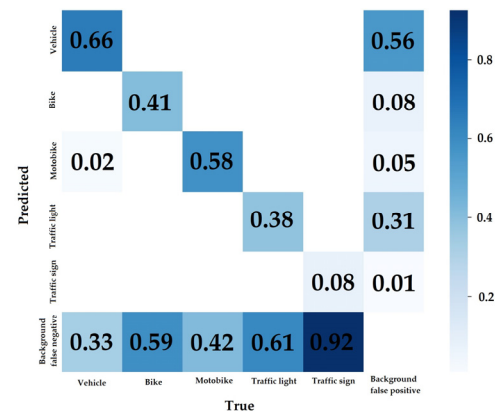


Figure 2. Confusion matrix of the simulation.

The model's performance was more favorable for larger, easily recognizable classes such as vehicles and motorcycles, while it exhibited significant weaknesses in recognizing smaller or visually more difficult objects such as traffic signs and lights. These results suggest that the model's limitations in detecting smaller or less visually distinct objects may be partly due to insufficient training exposure. A larger and more balanced dataset—especially one that includes a greater number of annotated examples of traffic signs and lights under varying conditions—could have provided the model with a more comprehensive basis for learning the characteristic features of these classes.

Figure 3a displays the F1 score of the model as a function of the confidence value, divided into different object classes. The F1 score reflects the balance between accuracy and recall, thus characterizing detection performance. The vehicle class performs best, with an F1 value rising to nearly 0.6 and remaining stable across a wider confidence range. These values suggest that the model is quite confident and accurate in recognizing vehicles. For the bike and motobike classes, the F1 value peaks at around 0.4, but based on the variability of the curves, the model's performance is less consistent in these cases. This may indicate similarities between the classes or image uncertainties. The traffic light class also shows a medium F1 value, but its curve is stepped, indicating fluctuating performance, probably due to more difficult visual recognition. Traffic sign detection is the weakest, with an F1 value rarely exceeding 0.2, indicating low reliability. The combined, wider blue curve shows the overall performance of the model for all classes. The best average F1 score is 0.499 at a confidence threshold of 0.35, as highlighted in the figure. In practice, this value may represent the optimal confidence threshold.

Figure 3b shows the precision value of the model by class as a function of confidence. The motobike class behaves most favorably with high confidence, with a precision value exceeding 0.9, meaning that false alarms are extremely rare. The vehicle class also shows consistently high accuracy, around a maximum of 0.6. The precision value of the traffic sign class increases sharply with greater confidence, but the curve is irregular, which may indicate a small number of elements. The bike and traffic light classes show lower and more fluctuating performance, especially as confidence increases. Based on the curve displaying

the cumulative performance, the model achieves its highest average precision value (0.72) with a confidence of 0.996. High accuracy requires very strict threshold values. The model is accurate in detecting motorcycles and vehicles, while for other classes, accuracy depends more on the confidence setting, which determines the threshold value to minimize false positives.

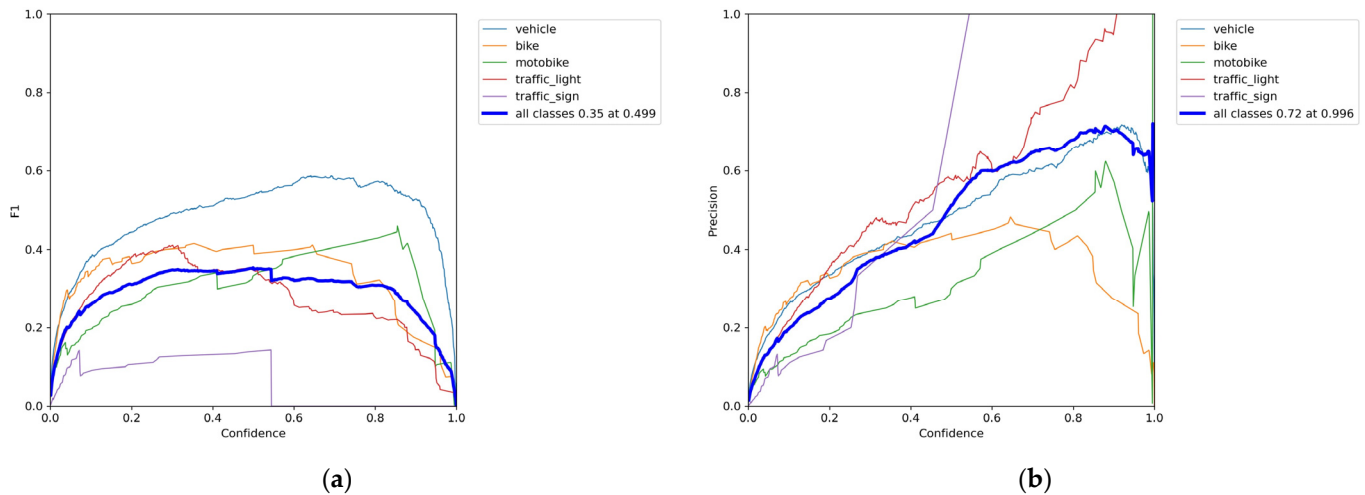


Figure 3. Simulation results showing the relationship between (a) F1 and confidence values and (b) precision and confidence values.

Figure 4a presents the precision–recall curves and average precision (AP) for each model class. The vehicle class performs reasonably well, with an AP value of 0.474, supported by a balanced precision–recall curve. The traffic light (0.288) and motobike (0.276) classes also perform acceptably, although they are characterized by high precision and low recall, meaning that they are less frequent but more accurate. The AP value of the bike class is lower at 0.238, and its curve drops rapidly, indicating that many detections are missing or incorrect. The traffic sign class performs the worst (0.087), as shown by its curve rapidly approaching 0, indicating few correct detections and many false detections. The mAP@0.5 value calculated for the entire model is 0.272, which indicates moderate performance, especially due to the weaker recognition of smaller and visually complex classes.

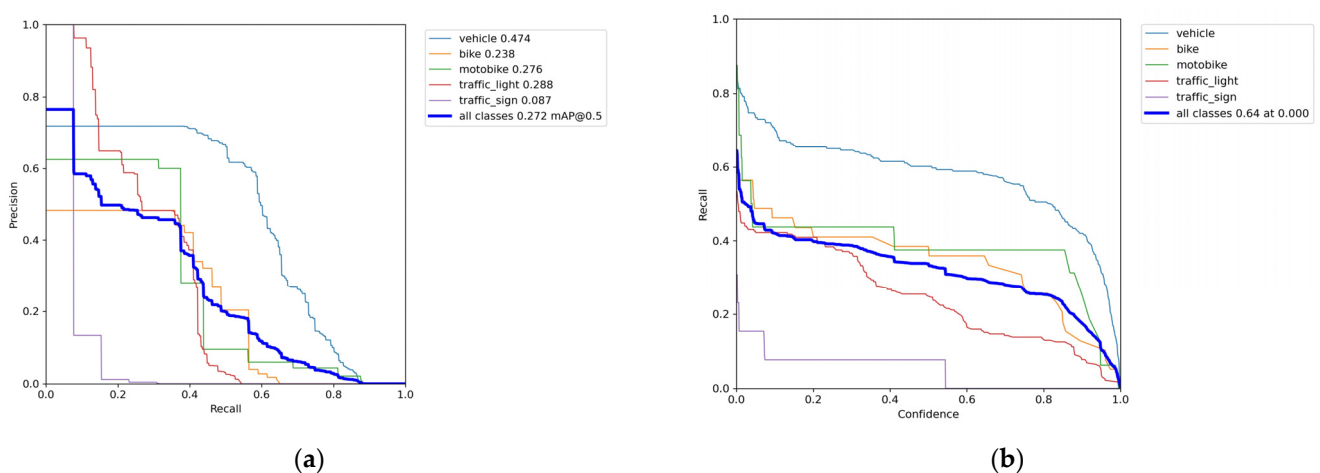


Figure 4. Simulation results showing the relationship between (a) precision and recall values and (b) recall and confidence values.

Figure 4b shows the relationship between recall and confidence by class. The vehicle class shows the most favorable performance, achieving a recall value above 0.8 even with

low confidence, meaning that it recognizes the majority of vehicles. The bike, motobike, and traffic light classes show moderate recall (0.3–0.5), with performance rapidly deteriorating as confidence increases. Traffic signs are the least accurate, with a recall value of just over 0.1. The maximum overall recall is 0.64, which the model achieves at the lowest confidence value (0.0). Based on this, the model is most accurate at recognizing vehicles and often misses smaller objects.

Figure 5 illustrates the development of the model training and validation process along different metrics, by epoch. The top row shows the values for the training data, while the bottom row shows the values for the validation data. In the case of the loss functions (box, objectness, classification), it is clear that all components decrease continuously during training, indicating the convergence of the model and the effectiveness of the learning process. In terms of validation losses, box and classification losses show a similar decreasing trend, but objectness loss begins to rise slightly in later epochs, which could indicate overfitting. The accuracy metrics, precision, recall, mAP@0.5, and mAP@0.5:0.95, stabilize over the learning process. The precision value is between 0.6 and 0.7, the recall is around 0.4, while the mean average precision (mAP@0.5) reaches 0.3. The stricter mean average precision (mAP@0.5:0.95) value ranges between 0.15 and 0.2, indicating moderate object recognition performance. It can be concluded that the model learns effectively, which is reflected in the improvement of the metrics. However, based on the increase in validation objectness loss and moderate mAP values, it might be worth improving performance with further development or data augmentation, especially for object classes that are more difficult to identify.

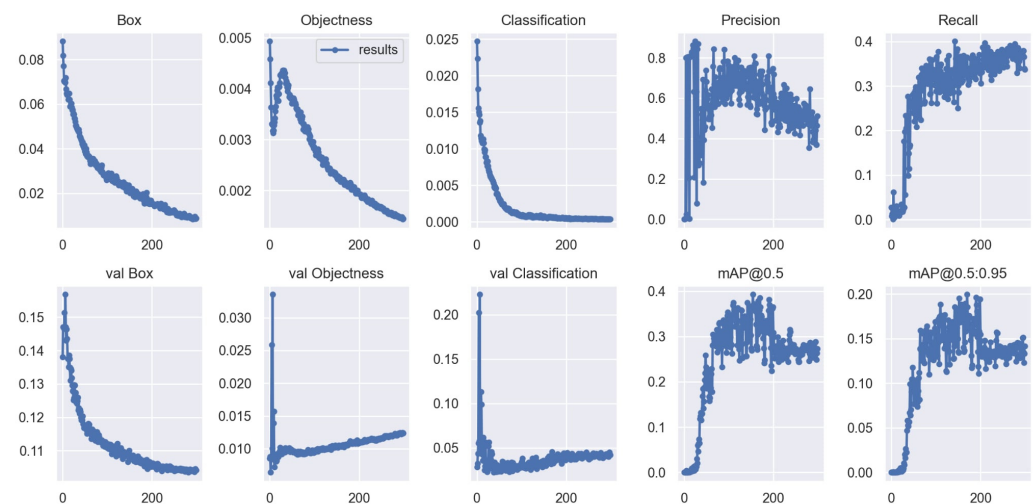


Figure 5. The development of the model training and validation process along different metrics.

A key limitation affecting the model's performance is the relatively small and imbalanced dataset used for training and evaluation, consisting of only 779 annotated images. This constrained the model's ability to generalize across object classes—particularly for small, less frequent categories such as traffic signs and lights—resulting in unstable precision–recall behavior and low F1 scores for these classes. Based on these limitations, this study should be viewed as a starting point for further investigations. The primary goal was to demonstrate the feasibility of a simulation-based testing framework using CARLA and YOLOv7. Future research will build upon this foundation by expanding the dataset, improving class balance, and validating results in real-world environments. The presented methodology offers a reproducible and extensible basis for such comparative studies and for continued development in environment perception for autonomous driving systems.

4. Conclusions

The objective of this research was to investigate an environment recognition system for autonomous vehicles using the CARLA simulation platform and the YOLOv7 object detection model. Various urban traffic scenarios were simulated, and the model was trained using YOLO-format annotations generated from 779 images. Performance evaluation was based on key detection metrics, including F1 score, precision, recall, mAP@0.5, and confusion matrix. The results showed that the model achieved acceptable recognition performance for large and frequently occurring object classes such as vehicles and motorcycles. However, detection performance dropped significantly for smaller or less visually distinct classes, such as traffic signs and traffic lights.

A major limitation of the study was the small and imbalanced dataset, which constrained the model's ability to generalize and led to unstable detection behavior—particularly in the case of underrepresented object categories. The low recall and irregular F1 score patterns across confidence thresholds reflect the insufficient exposure of the model to diverse object instances during training. These shortcomings highlight the need for a larger and more representative dataset that better captures variability in object types, scales, and environments.

The novelty of the present study lies in the combined assessment of multiple object classes under varying environmental and weather conditions, which provides a more realistic simulation of autonomous vehicle perception challenges. This comprehensive evaluation setup establishes a methodological foundation for further research.

Future work will extend this framework by increasing the dataset size, validating the model in real-world settings, and performing comparative analysis with alternative object detection architectures, including other YOLO variants. The flexible and reproducible pipeline demonstrated in this study offers a solid baseline for iterative development and algorithmic benchmarking in simulated and practical environments.

Author Contributions: Conceptualization, G.S. and D.C.; methodology, D.G., G.S. and D.C.; software, D.G.; validation, D.G. and D.C.; formal analysis, G.S. and D.C.; investigation, D.G. and G.S.; resources, G.S. and D.C.; writing—original draft preparation, D.G. and G.S.; writing—review and editing, G.S. and D.C.; visualization, D.G.; supervision, G.S. and D.C.; project administration, D.C.; funding acquisition, D.C. All authors have read and agreed to the published version of the manuscript.

Funding: The research was supported by the European Union within the framework of the National Laboratory for Autonomous Systems (RRF-2.3.1-21-2022-00002).

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: The raw data supporting the conclusions of this article will be made available by the authors upon request.

Conflicts of Interest: The authors declare no conflict of interest.

References

1. Orejon-Sanchez, R.D.; Crespo-Garcia, D.; Andres-Diaz, J.R.; Gago-Calderon, A. Smart Cities' Development in Spain: A Comparison of Technical and Social Indicators with Reference to European Cities. *Sustain. Cities Soc.* **2022**, *81*, 103820. [\[CrossRef\]](#)
2. Elassy, M.; Al-Hattab, M.; Takruri, M.; Badawi, S. Intelligent Transportation Systems for Sustainable Smart Cities. *Transp. Eng.* **2024**, *16*, 100303. [\[CrossRef\]](#)
3. Rizzi, M.; Strandroth, J. Road Safety Analysis. In *The Vision Zero Handbook*; Springer: Cham, Switzerland, 2022; pp. 1–23.
4. Xiao, H.; Ju, C.; Zhao, J. Research on the Development Limitations of ADAS under the Intelligent Trend of New Energy Vehicles. In Proceedings of the 18th International Conference on Computational Intelligence and Security (CIS 2022), Chengdu, China, 23–26 September 2022; IEEE: Piscataway, NJ, USA, 2022; pp. 234–242.

5. Masello, L.; Castignani, G.; Sheehan, B.; Murphy, F.; McDonnell, K. On the Road Safety Benefits of Advanced Driver Assistance Systems in Different Driving Contexts. *Transp. Res. Interdiscip. Perspect.* **2022**, *15*, 100670. [[CrossRef](#)]
6. Lie, A.; Tingvall, C.; Håkansson, M.; Boström, O. Automated Vehicles—How Do They Relate to Vision Zero? In *The Vision Zero Handbook*; Springer: Cham, Switzerland, 2022; pp. 1–16.
7. Morales-Alvarez, W.; Sipele, O.; Léberon, R.; Tadjine, H.H.; Olaverri-Monreal, C. Automated Driving: A Literature Review of the Take Over Request in Conditional Automation. *Electronics* **2020**, *9*, 2087. [[CrossRef](#)]
8. Gu, J.; Lind, A.; Chhetri, T.R.; Bellone, M.; Sell, R. End-to-End Multimodal Sensor Dataset Collection Framework for Autonomous Vehicles. *Sensors* **2023**, *23*, 6783. [[CrossRef](#)] [[PubMed](#)]
9. Mohammed, S.A.; Ralescu, A.L. Insights into Image Understanding: Segmentation Methods for Object Recognition and Scene Classification. *Algorithms* **2024**, *17*, 189. [[CrossRef](#)]
10. Dursun, C.; Erdei, T.I.; Husi, G. Artificial Intelligence Applications in Autonomous Vehicles: Training Algorithm for Traffic Signs Recognition. *IOP Conf. Ser. Mater. Sci. Eng.* **2020**, *898*, 012027. [[CrossRef](#)]
11. Priya, S.; Kumar, S.S.; Lavanya, P.; Sadik, S.; Kumar, A.K. Real-Time Image Segmentation and Object Tracking for Autonomous Vehicles. In Proceedings of the 3rd International Conference on Advances in Computing, Communication and Applied Informatics (ACCAI 2024), Chennai, India, 3–5 January 2024; IEEE: Piscataway, NJ, USA, 2024.
12. Barrozo, J.I.; Lazcano, V. Simulation of an Autonomous Vehicle Control System Based on Image Processing. In Proceedings of the 5th International Conference on Frontiers of Signal Processing (ICFSP), Marseille, France, 18–20 September 2019; IEEE: Piscataway, NJ, USA, 2019; pp. 88–94.
13. Al-Haija, Q.A.; Gharaibeh, M.; Odeh, A. Detection in Adverse Weather Conditions for Autonomous Vehicles via Deep Learning. *AI* **2022**, *3*, 303–317. [[CrossRef](#)]
14. Ali, E.; Khan, M.N.; Ahmed, M.M. Real-Time Snowy Weather Detection Based on Machine Vision and Vehicle Kinematics: A Non-Parametric Data Fusion Analysis Protocol. *J. Saf. Res.* **2022**, *83*, 163–180. [[CrossRef](#)] [[PubMed](#)]
15. Ghintab, S.S.; Hassan, M.Y. CNN-Based Visual Localization for Autonomous Vehicles under Different Weather Conditions. *Eng. Technol. J.* **2023**, *41*, 375–386. [[CrossRef](#)]
16. Niranjana, D.R.; Vinaykarthik, B.C.; Mohana. Performance Analysis of SSD and Faster RCNN Multi-Class Object Detection Model for Autonomous Driving Vehicle Research Using CARLA Simulator. In Proceedings of the 4th International Conference on Electrical, Computer and Communication Technologies (ICECCT 2021), Erode, India, 15–17 September 2021; IEEE: Piscataway, NJ, USA, 2021.
17. Dosovitskiy, A.; Ros, G.; Codevilla, F.; Lopez, A.; Koltun, V. CARLA: An open urban driving simulator. In Proceedings of the 1st Annual Conference on Robot Learning, Mountain View, CA, USA, 13–15 November 2017; JMLR, Inc.: Cambridge, MA, USA, 2017; Volume 78, pp. 1–16.
18. Rong, G.; Hyun Shin, B.; Tabatabaee, H.; Lu, Q.; Lemke, S.; Možeiko, S.; Boise, E.; Uhm, G.; Gerow, M.; Mehta, S.; et al. LGSVL Simulator: A High Fidelity Simulator for Autonomous Driving. In Proceedings of the IEEE 23rd International Conference on Intelligent Transportation Systems (ITSC), Rhodes, Greece, 20–23 September 2020; pp. 1–6.
19. Shah, S.; Dey, D.; Lovett, C.; Kapoor, A. AirSim: High-Fidelity Visual and Physical Simulation for Autonomous Vehicles. In *Springer Proceedings in Advanced Robotics, Proceedings of the 11th International Conference on Field and Service Robotics (FSR), Zurich, Switzerland, 12–15 September 2017*; Springer: Cham, Switzerland, 2018; Volume 5, pp. 621–635.
20. Wang, C.-Y.; Bochkovskiy, A.; Liao, H.-Y.M. YOLOv7: Trainable Bag-of-Freebies Sets New State-of-the-Art for Real-Time Object Detectors. In Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR), Vancouver, BC, Canada, 18–24 June 2023; pp. 7464–7475.
21. Aulia, S.; Suksmono, A.B.; Mengko, T.R.; Alisjahbana, B. A Novel Digitized Microscopic Images of ZN-Stained Sputum Smear and Its Classification Based on IUATLD Grades. *IEEE Access* **2024**, *12*, 51364–51380. [[CrossRef](#)]

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.