Improving YOLO11 Architecture for Reckless Driving Detection on the Road

Sutikno, Aris Sugiharto, Retno Kusumaningrum
Department of Informatics, Diponegoro University, Semarang, Indonesia

***Abstract*—Reckless driving behavior on the road can increase the risk of traffic accidents for drivers and other road users. Currently, supervision remains weak, particularly in direct supervision, due to the limited number of officers. This study developed an automated system to detect reckless drivers based on their road trajectories. This system comprised three subsystems: car detection, car tracking, and driving trajectory detection. In the driving trajectory detection subsystem, we proposed an improved YOLO11n-cls method developed from YOLO11n-cls by adding convolution and C3k2 blocks. The test results showed that the proposed model achieved an accuracy increase of 4.4% over YOLO11n-cls. The proposed model achieved an accuracy of 0.935 and an inference time of 0.5 ms for car trajectory classification. In addition, the proposed model achieved higher accuracy than all YOLO11 models (YOLO11n-cls, YOLO11s-cls, YOLO11m-cls, YOLO11l-cls, and YOLO11x-cls) and all YOLO12 models (YOLO12n-cls, YOLO12s-cls, YOLO12m-cls, YOLO12l-cls, and YOLO12x-cls). Therefore, the proposed model is better suited to support traffic law enforcement, especially the real-time detection of reckless drivers on highways.**

Keywords—Reckless driving detection; improved YOLO11n-cls; added convolution blocks; added C3k2 blocks

I. Introduction

Deaths from traffic accidents remain a serious global problem. One contributing factor is drivers who disobey traffic regulations, such as dangerous or reckless driving. These violations not only endanger the driver but also put other road users at risk, thereby increasing the likelihood of accidents. Furthermore, many violations still occur. Therefore, it is crucial to strengthen supervision and enforcement of the rules. Currently, supervision is still carried out directly on the road due to limited personnel. This way has the disadvantage of potentially undetected violations. One solution to this problem is to utilize roadside surveillance cameras in conjunction with automated detection systems to identify drivers who drive dangerously or recklessly, thereby preventing traffic accidents. This system requires high accuracy and speed to be implemented in real time.

Research on detecting reckless driving behavior on the highway began about a decade ago. Based on the location of the sensor installation, this research was divided into two categories, namely sensors installed inside the car and surveillance camera sensors installed above the highway. Sensors installed inside the car monitor driver behavior directly. Meanwhile, surveillance camera sensors installed above the highway analyze driver trajectory patterns and speeds. Some methods used for sensor-based car systems included AdaBoost [1], Backpropagation [2], Support Vector Machine (SVM) [3], Graph Convolutional Long Short-term Memory networks (GConvLSTM) [4], YOLOv5 [5], YOLOv8 [6], and ResNet18+SCConv [7]. Meanwhile, some methods using surveillance cameras installed above the highway were SVM [8], Boosting Artificial Intelligence (BAI) [9], Lightweight Graph Convolution (LGC) [10], Convolutional Neural Network (CNN) [11][12], YOLOv2 [13], and YOLOv3 [14].

Cameras installed above the road are better suited for law enforcement, especially for detecting reckless driving. YOLOv2 and YOLOv3 are superior to other methods in single-frame multi-object detection, but they have weaknesses in detecting small objects and require substantial computational resources [13][14]. The models make them challenging to implement in real-time. These models produce low accuracy on small-looking car objects.

YOLO excels because it is fast, accurate, efficient, and easy to implement. It is also suitable for real-time object detection across various applications, including detecting parking space availability [15], video summarization [16], anomaly detection from a video surveillance camera [17], and student behavior detection [18]. Another investigation aimed at detecting pedestrians [19], vehicles [20], and road objects [21]. YOLOv2 and YOLOv3 have been used for reckless driver detection, but these models produce relatively low accuracy and long computation time compared to YOLO11.

This study proposed a system for detecting reckless car drivers on the road based on their trajectories, suitable for real-time conditions. This system comprised three subsystems: car detection, car tracking, and driving trajectory detection. We utilize the modified YOLO11 in the car detection subsystem, the ByteTrack algorithm in the car tracking subsystem, and the improved YOLO11n-cls model for driving trajectory detection. YOLO11n-cls still produces substantial errors in classifying car trajectories on the highway, especially in the smooth lane-change class. This study develops the improved YOLO11n-cls model, which is based on YOLO11n-cls. This model added convolutional layers and C3k2 blocks to expand the model's representation capacity, resulting in a YOLO11n-cls architecture that was more powerful at classifying car driver trajectories on the road. The development of this model aims to improve the accuracy of car trajectory classification on the road while maintaining a consistent speed.

II. Materials and Methods

A. Dataset

The data used in this study are vehicle trajectory images from a 70-minute video recording on a highway. The video has a resolution of 960 × 540 pixels and a frame rate of 30 frames per second (fps). The recording was made with a camera mounted above the highway. The data was recorded on a highway in Semarang City, Indonesia. Fig. 1 is an example of a frame from this video.

Fig. 1. Example of a video frame.

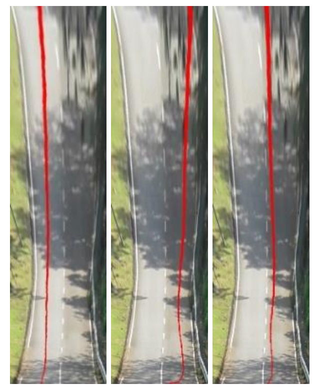
The trajectory data consisted of 153 images across three classes: 87 for straight driving, 30 for smooth lane changes, and 36 for reckless driving. This dataset was split into two parts: 70% for training and 30% for testing. The straight driving class represents the trajectory of a car that does not change lanes, and the smooth lane change class represents the trajectory of a car that changes lanes once. Meanwhile, the reckless driving class represents a car that changes lanes more than once. An example of this dataset is shown in Fig. 2.

B. Proposed System

The method proposed in this study was a reckless road traffic detection system based on the driver's car trajectory. This system consisted of three subsystems: car detection, car tracking, and driving trajectory detection, as illustrated in Fig. 3. The car detection subsystem determined the car's position on the road. The method used was modified YOLO11m, which has been shown to achieve higher mAP than YOLOv8x, YOLOv8m, and all YOLO11 models [22]. The next subsystem was car tracking, which determined each car's trajectory on the road. The method used in this tracking was the ByteTrack algorithm. This algorithm has been demonstrated to outperform other tracking methods, including SORT, DeepSORT, and DeepMOT [23].

The final subsystem, driving trajectory detection, is illustrated in Fig. 4. It comprises several processes: car trajectory reconstruction, Region of Interest (ROI) frame determination, and trajectory classification. The trajectory reconstruction process was used to redraw the car's trajectory on a base frame. This base frame was obtained from one of the video recording frames. Next, the ROI was determined to identify the key areas of the trajectory, namely the road. The ROI was determined by determining 4 points on the highway boundary. The ROI area was left fixed while the others were colored black. The next step was to crop the ROI, then continue the perspective transformation (warping) to produce only the ROI area of the car's trajectory on the highway.

The final process involved classifying the trajectory of the car in this image. This study developed a YOLO11n-cls model for classifying three car-trajectory classes: straight driving, smooth lane change, and reckless driving. The proposed method is named improved YOLO11n-cls, as shown in Fig. 5. The YOLO11n-cls model was essentially a YOLO-based classification model that used a lightweight backbone with convolutional blocks and C3k2 modules for feature extraction, along with a C2PSA (Cross-Stage Partial Squeeze-and-Attention) module to capture spatial and contextual dependencies at a high feature level [24]. This structure is designed to be computationally efficient, while still producing rich feature representations, making it suitable for various image classification tasks, including vehicle movement trajectories.



(a)

(b)

(c)

Fig. 2. Example of datasets: (a) Straight driving class, (b) Smooth lane change class, (c) Reckless driving class.

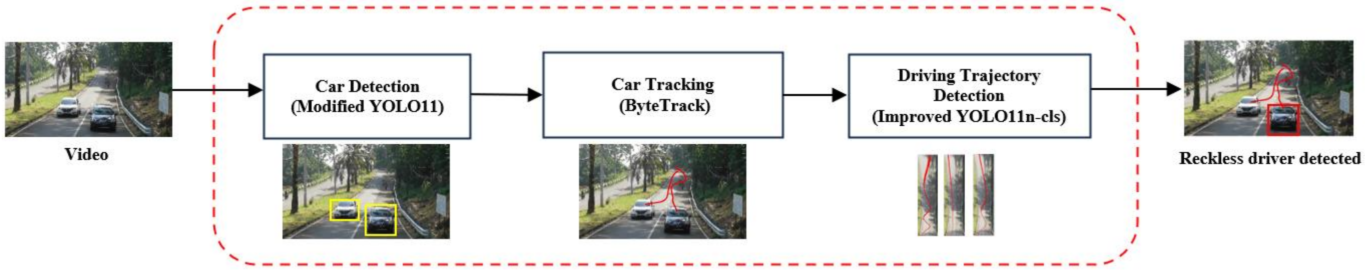


Fig. 3. The proposed system for reckless road traffic detection.

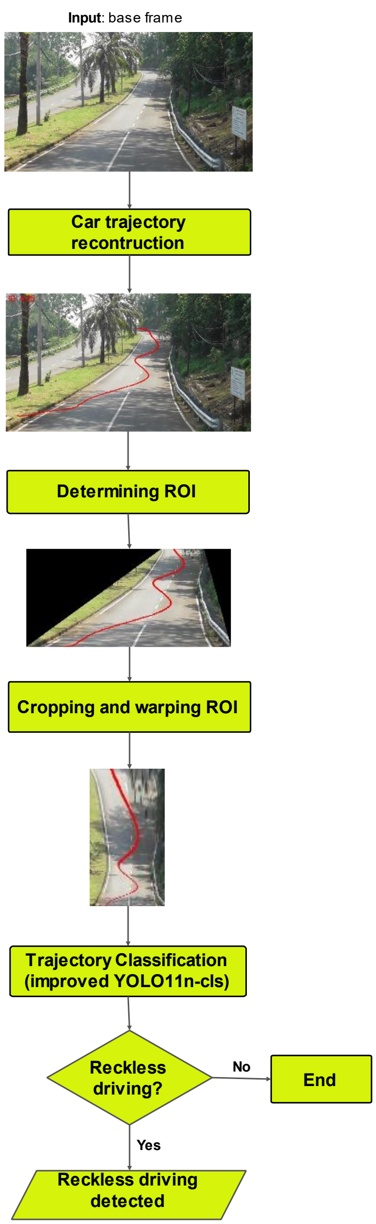


Fig. 4. Driving trajectory detection subsystem.

In the improved YOLO11n-cls model, we modified the backbone to enhance its feature extraction, particularly for complex car trajectory patterns. This modification was achieved by adding a convolutional layer with 2,028 output channels, followed by a C3k2 block at the high-level feature level. This addition aims to expand the model's representation capacity to capture variations in trajectory patterns, increase feature depth, and improve sensitivity to subtle visual differences across trajectory classes. The C3k2 block produces more efficient computation compared to C2f in the previous version of the YOLO model. Furthermore, the C2PSA module was expanded to process higher-resolution features, thereby improving the network's spatial and contextual attention.

The primary objective of this method development is to enhance classification accuracy while maintaining a relatively fast inference time, thereby enabling the model to be applied efficiently in real-time systems. The improved YOLO11n-cls model is expected to deliver more accurate, responsive classification, potentially supporting intelligent transportation systems and real-time driver behavior analysis.

C. Performance Evaluation

The classification model was evaluated using accuracy (Acc), precision, recall, and F1-score as defined by Eq. (1), Eq. (2), Eq. (3), and Eq. (4), respectively [25]. The variable T_{pos} (true positive) represents the number of samples that are correctly classified as belonging to a specific class. In contrast, T_{neg} (true negative) indicates the number of samples that are correctly classified as not belonging to that class. F_{pos} (false positive) refers to the number of samples that are incorrectly classified as belonging to a class when they actually belong to another class, and F_{neg} (false negative) denotes the number of samples that belong to a specific class but are incorrectly classified as another class.

$$Accuration = \frac{T_{pos}+T_{neg}}{T_{pos}+T_{neg}+F_{pos}+F_{neg}} \tag{1}$$

$$Precision = \frac{T_{pos}}{T_{pos}+F_{pos}} \tag{2}$$

$$Recall = \frac{T_{pos}}{T_{pos}+F_{neg}} \tag{3}$$

$$F1-score = 2 \times \frac{Precision \times Recall}{Precision + Recall} \tag{4}$$

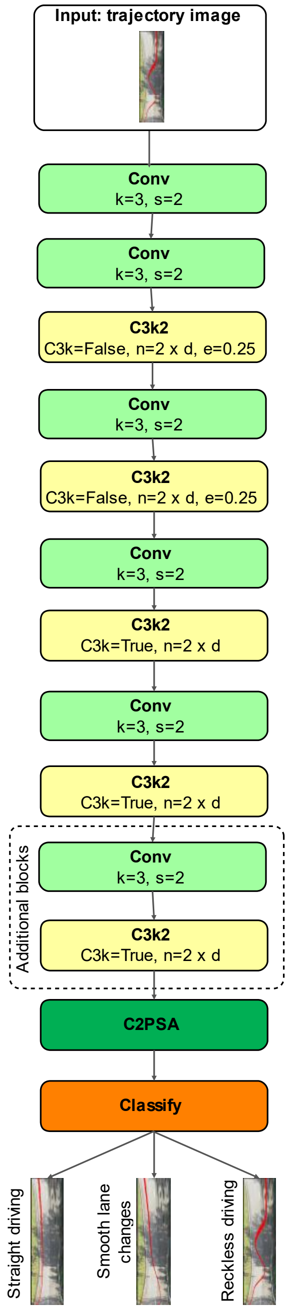


Fig. 5. Improved YOLO11n-cls architecture (proposed model).

III. Experiments and Results

Testing was conducted using several scenarios. First, we tested the proposed method on a dataset. Next, we performed data augmentation to augment the training data. The proposed method was then compared with several other models. The results of the first test using the proposed method are shown in Table I. The precision and recall values for the reckless driving class were quite good. In fact, the recall value was perfect, but the precision and recall for the smooth lane change class could not be classified. Meanwhile, the accuracy, inference time, and training time from this test resulted in 0.891, 0.5 ms, and 0.062 hours, respectively.

The training dataset for the first test was still small—only 107 images—so we augmented the data. The technique used was horizontal scaling. Each vehicle trajectory was scaled proportionally to its starting point, using a scale factor of 0.4-1.5 and a scale increment of 0.1. This technique maintains the original trajectory shape while introducing left-right shifts, making the model more robust to differences in vehicle positions within the lane. The resulting augmented data amounted to 326, so the training data became 433. The test results, incorporating the augmented data, are presented in Table II. The precision and recall values for the reckless driving class were 0.8750 and 1, respectively. The F1-score of the smooth lane change class experienced a significant increase of 0.6667. The confusion matrix of this classification is shown in Table III. The proposed model performs reasonably well at predicting reckless and straight driving, but it still struggles to predict smooth lane changes. Further improvements are needed. Furthermore, additional trajectory datasets are needed to improve accuracy across all classes. Frequent lane changes and high speeds are characteristic of reckless drivers. The proposed method can be combined with speed exceeding threshold detection.

The comparison of accuracy, inference time, and training time between before and after the addition of augmented data is shown in Table IV. From this table, it can be seen that with the addition of augmented data, accuracy increased by 4.4%, while inference time remained the same. Meanwhile, the training time increased threefold. Therefore, it can be concluded that the proposed method, improved YOLO11n-cls, is more suitable for car trajectory classification than YOLO11n-cls.

TABLE I. Test Results of the Proposed Method on the Dataset Without Augmentation

Class	Precision	Recall	F1-score
Reckless driving	0.9375	1.0000	0.9684
Smooth lane change	0.0000	0.0000	0.0000
Straight driving	0.9630	0.8966	0.9285

TABLE II. Test Results of the Proposed Method on the Dataset with Augmentation

Class	Precision	Recall	F1-score
Reckless Driving	0.8750	1.0000	0.9333
Smooth Lane Change	0.6667	0.6667	0.6667
Straight Driving	1.0000	0.9310	0.9646

TABLE III. Confusion Matrix of Trajectory Classification

		Actual Value		
		Reckless driving	***Smooth lane change***	***Straight driving***
Predicted	***Reckless driving***	14	0	0
	Smooth lane change	1	2	0
	Straight driving	1	1	27

TABLE IV. COMPARISON RESULTS WITH AND WITHOUT AUGMENTATION

Augmentasi	Acc	Inference time	Training time
No	0.891	**0.5**	**0.062**
Yes	**0.935**	**0.5**	0.172

Finally, the proposed method was compared with YOLO11 models and YOLO12 models, as shown in Table V. From this table, it is evident that the proposed model's accuracy surpasses that of the other models. At the same time, the inference time is ranked second among all. The proposed method achieves an accuracy of 0.935 and an inference time of 0.5 ms. Therefore, it can be concluded that the proposed method is more suitable for real-time car trajectory classification than all YOLO11 and YOLO12 models. The proposed model improves the average precision of reckless driving detection using vehicle trajectories. Compared to YOLO11, the proposed model achieved higher accuracy due to its enhanced representation capacity and ability to process higher-resolution features.

TABLE V. COMPARISON RESULTS OF THE PROPOSED METHOD WITH ALL YOLO11 AND YOLOV12 MODELS

Model	Acc	Inference time (ms)	Training time (hours)
YOLO11n-cls	0.913	**0.4**	0.170
YOLO11s-cls	0.870	0.5	**0.051**
YOLO11m-cls	0.891	0.5	0.110
YOLO11l-cls	0.891	0.9	0.078
YOLO11x-cls	0.913	0.7	0.100
YOLO12n-cls	0.870	0.8	0.064
YOLO12s-cls	0.848	1.0	0.067
YOLO12m-cls	0.870	0.6	0.076
YOLO12l-cls	0.891	2.0	0.126
YOLO12x-cls	0.870	3.0	0.136
Proposed	**0.935**	0.5	0.172

IV. CONCLUSION

This study developed a system to detect dangerous driving behavior based on vehicle trajectories. The developed system consisted of three subsystems: car detection using a modified YOLO11, car tracking using ByteTrack, and driving trajectory detection using an improved YOLO11n-cls. Improved YOLO11n-cls is a model developed from the original YOLO11n-cls by adding convolution and C3k2 blocks before the C2PSA block. The test results show that the proposed model can improve accuracy by 4.4% compared to YOLO11n-cls. The proposed model achieves an accuracy of 0.935 and an inference time of 0.5 ms. Additionally, the resulting model outperforms all YOLO11-cls and YOLO12-cls models. These results can improve the average precision of reckless driver detection based on their trajectory. However, the proposed model still performs poorly in predicting smooth lane changes. Therefore, the proposed model is better suited to support traffic law enforcement, especially the real-time detection of reckless drivers on highways. Further research can expand the dataset with more diverse traffic conditions and refine the classification model to capture more complex driving patterns.

ACKNOWLEDGMENT

The authors thank the Ministry of Higher Education, Science, and Technology of Indonesia and Diponegoro University for funding this study under contract number: 359-094/UN7.D2.1/PP/VI/2025.

REFERENCES

[1] Y. Qiao, K. Zeng, L. Xu, and X. Yin, "A smartphone-based driver fatigue detection using fusion of multiple real-time facial features," in 2016 13th IEEE Annual Consumer Communications & Networking Conference (CCNC), 2016, pp. 230–235, https://doi.org/10.1109/CCNC.2016.7444761.

[2] L. Xu, S. Li, K. Bian, T. Zhao, and W. Yan, "Sober-Drive: A smartphone-assisted drowsy driving detection system," in 2014 International Conference on Computing, Networking and Communications (ICNC), 2014, pp. 398–402, https://doi.org/10.1109/ICCNC.2014.6785367.

[3] A. Dasgupta, D. Rahman, and A. Routray, "A smartphone-based drowsiness detection and warning system for automotive drivers," IEEE Transactions on Intelligent Transportation Systems, vol. 20, no. 11, pp. 4045–4054, 2019, https://doi.org/10.1109/TITS.2018.2879609.

[4] P. Khosravinia, T. Perumal, and J. Zarrin, "Enhancing road safety through accurate detection of hazardous driving behaviors with graph convolutional recurrent networks," IEEE Access, vol. 11, pp. 52983 – 52995, 2023, https://doi.org/10.1109/ACCESS.2023.3280473.

[5] X. G. Zhang and L. Liao, "Identification methods for driver's dangerous driving behavior in low visibility environments," Advances in Transportation Studies, vol. 2, no. Special Issue, pp. 129 – 142, 2024, https://doi.org/10.53136/979122181412511.

[6] F. M. Al Ali, M. J. Alnuaimi, S. A. Alawadhi, and S. Abdallah, "Abnormal driver behavior detection using deep learning," in 2024 7th International Conference on Signal Processing and Information Security (ICSPIS), 2024, pp. 1–4, https://doi.org/10.1109/ICSPIS63676.2024.10812606.

[7] X. Feng, T. Peng, T. Duan, H. Hou, Q. Chen, and Z. Zhang, "DdbiNet: a dangerous driving behavior identification network based on Resnet18," in 2024 5th International Conference on Information Science, Parallel and Distributed Systems, ISPDS 2024, 2024, pp. 565 – 570, https://doi.org/10.1109/ISPDS62779.2024.10667560.

[8] Q. Xue, K. Wang, J. J. Lu, and Y. Liu, "Rapid driving style recognition in car-following using machine learning and vehicle trajectory data," Journal of Advanced Transportation, vol. 2019, no. 1, p. 9085238, 2019, https://doi.org/https://doi.org/10.1155/2019/9085238.

[9] J. Lu, K. Wang, and Y.-M. Jiang, "Real-time identification method of abnormal road driving behavior based on vehicle driving trajectory; [Real-Time Identification Method of Bad Driving Behavior on Road based on Vehicle Driving Trajectory]," Jiaotong Yunshu Gongcheng Xuebao/Journal of Traffic and Transportation Engineering, vol. 20, no. 6, pp. 227 – 235, 2020, https://doi.org/10.19818/j.cnki.1671-1637.2020.06.020.

[10] R. J. Wang, "Identification method for abnormal driving behavior of vehicles based on lightweight graph convolution," Advances in Transportation Studies, vol. 3, no. Special Issue, pp. 97 – 108, 2024, https://doi.org/10.53136/97912218165708.

[11] V. Bhosale, J. Shah, P. Doshi, R. Mangrulkar, and I. Williams, "Enhancing road safety: reckless driver detection via OpenCV in simulated environments," in Intelligent Computing and Big Data Analytics, 2024, pp. 207–222.

[12] X. W. Yang, "Classification recognition for traffic dangerous driving behavior," Advances in Transportation Studies, vol. 1, no. Special Issue, pp. 115 – 124, 2024, https://doi.org/10.53136/979122181230511.

[13] T. Heo, W. Nam, J. Paek, and J. Ko, "Autonomous reckless driving detection using deep learning on embedded GPUs," in 2020 IEEE 17th International Conference on Mobile Ad Hoc and Sensor Systems (MASS), 2020, pp. 464–472, https://doi.org/10.1109/MASS50613.2020.00063.

[14] A. Chauhan, S. Kumar, L. Mahto, and D. N. Jagadish, "Detection of reckless driving using deep learning," in Proceedings - 19th IEEE

International Conference on Machine Learning and Applications, ICMLA 2020, 2020, pp. 853 – 858, https://doi.org/10.1109/ICMLA51294.2020.00139.

[15] M. F. Asy'ari, C. Fatichah, and N. Suciati, "Parking space availability detections from two overlapping cameras using YOLOv5 and image stitching methods," International Journal of Intelligent Engineering and Systems, vol. 16, no. 4, pp. 276–288, 2023, https://doi.org/10.22266/ijies2023.0831.23.

[16] S. K. Jarallah and S. A. Mahmood, "Query-based video summarization system based on lightweight deep learning model," International Journal of Intelligent Engineering and Systems, vol. 15, no. 6, pp. 247–262, 2022, https://doi.org/10.22266/ijies2022.1231.24.

[17] I. A. Alnajjar, L. Almazaydeh, A. A. Odeh, A. A. Salameh, K. Alqarni, and A. A. B. Atta, "Anomaly detection based on hierarchical federated learning with edge-enabled object detection for surveillance systems in industry 4.0 scenario," International Journal of Intelligent Engineering and Systems, vol. 17, no. 4, pp. 649–665, 2024, https://doi.org/10.22266/IJIES2024.0831.49.

[18] M. A. E. Alkhalisy and S. H. Abid, "Students' behavior detection based on improved YOLOv5 algorithm combining with CBAM attention mechanism," International Journal of Intelligent Engineering and Systems, vol. 16, no. 5, pp. 473–487, 2023, https://doi.org/10.22266/ijies2023.1031.41.

[19] K. C. Saranya and A. Thangavelu, "Vulnerable road user detection using YOLOv3," International Journal of Advanced Computer Science and Applications, vol. 10, no. 12, pp. 576–582, 2019, https://doi.org/10.14569/ijacsa.2019.0101275.

[20] O. Bourja, H. Derrouz, H. A. Abdelali, A. Maach, R. O. H. Thami, and F. Bourzeix, "Real Time vehicle detection, tracking, and inter-vehicle distance estimation based on stereovision and deep learning using Yolov3," International Journal of Advanced Computer Science and Applications, vol. 12, no. 8, pp. 915–923, 2021, https://doi.org/10.14569/IJACSA.2021.01208101.

[21] G. Al-refai and M. Al-refai, "Road object detection using Yolov3 and Kitti dataset," International Journal of Advanced Computer Science and Applications, vol. 11, no. 8, pp. 48–53, 2020, https://doi.org/10.14569/IJACSA.2020.0110807.

[22] Sutikno, A. Sugiharto, and R. Kusumaningrum, "Modified YOLO11 for improving car detection performance on the highway," Bulletin of the Polish Academy of Sciences Technical Sciences (submitted). 2025.

[23] Z. Dong et al., "Fruit detection and yield estimation in Camellia oleifera based on improved YOLOv8 and ByteTrack algorithm," Smart Agricultural Technology, vol. 12, no. September, p. 101435, 2025, https://doi.org/10.1016/j.atech.2025.101435.

[24] R. Khanam and M. Hussain, "YOLOv11: an overview of the key architectural enhancements," 2024. [Online]. Available: http://arxiv.org/abs/2410.17725.

[25] R. Padilla, S. L. Netto, and E. A. B. Da Silva, "A survey on performance metrics for object-detection algorithms," in International Conference on Systems, Signals, and Image Processing, 2020, pp. 237–242, https://doi.org/10.1109/IWSSIP48289.2020.9145130.