*Uluslararası İleri Doğa Bilimleri ve Mühendislik Araştırmaları Dergisi*
*Sayı 9, S. 82-89, 12, 2025*

***Araştırma Makalesi***

IJANSER
https://as-proceeding.com/index.php/ijanser
ISSN:2980-0811

*International Journal of Advanced Natural Sciences and Engineering Researches*
*Volume 9, pp. 82-89, 12, 2025*

***Research Article***

# Deep Learning Image Analysis for Pedestrian and Animal Safety in Traffic intersection using YOLOv10 and YOLOv11

Bayan Sheikh Omar[1] and Önder Yakut [2,*]

[1]*Department of Information Systems Engineering, Faculty of Technology, Kocaeli University, Kocaeli, Turkey (ORCID: https://orcid.org/0009-0009-6452-3832)*

[2]*Department of Information Systems Engineering, Faculty of Technology, Kocaeli University, Kocaeli, Turkey (ORCID: https://orcid.org/0000-0003-0265-7252)*

[*]*(onder.yakut@kocaeli.edu.tr) Email of the corresponding author*

*(Received: 22 November 2025, Accepted: 03 December 2025)*

(6th International Conference on Contemporary Academic Research ICCAR 2025, November 25-26, 2025)

**ATIF/REFERENCE:** Omar, B. S. & Yakut, Ö. (2025). Deep Learning Image Analysis for Pedestrian and Animal Safety in Traffic intersection using YOLOv10 and YOLOv11, *International Journal of Advanced Natural Sciences and Engineering Researches*, 9(12), 82-89.

***Abstract*** **–** Ensuring the safety of pedestrians and animals at traffic intersections remains a critical challenge for intelligent transportation systems, particularly in complex urban environments. Traditional object detection methods often fail to deliver reliable real-time performance. However, recent advancements in deep learning, particularly within the YOLO (You Only Look Once) family, have significantly improved detection speed and accuracy. While earlier versions, such as YOLOv5 to YOLOv9, have been extensively studied, the potential of the more recent YOLOv10 and YOLOv11 architectures remains largely unexplored. This study provides a comparative analysis of these models in the context of detecting pedestrians and animals at intersections. The results show that YOLOv11 achieves an impressive mean Average Precision (mAP) of 89%, outperforming previous YOLO versions and current state-of-the-art methods. This positions YOLOv11 as a strong candidate for real-time traffic safety applications. The findings underscore the potential of cutting-edge deep learning models in advancing traffic safety and contribute to the development of smarter, more responsive transport infrastructure.

*Keywords – Deep Learning, Object Detection, Pedestrian and Animal Detection, Traffic Safety, YOLOv10, YOLOv11.*

## I. INTRODUCTION

Ensuring pedestrian and animal safety at traffic intersections is a critical challenge in modern transportation systems. According to the World Health Organization, over 1.3 million people die annually in road traffic accidents, with pedestrians representing a significant proportion of these fatalities. Animal–vehicle collisions also contribute to human injuries and ecological disruption, particularly in regions with frequent wildlife crossings, underscoring the need for intelligent real time detection systems.

Deep learning based object detection has become a key solution for traffic safety. The YOLO family of models is particularly notable for balancing accuracy and efficiency. Earlier versions, including YOLOv5 and YOLOv7, have shown strong performance in pedestrian detection in urban areas and animal detection on highways [1]-[4].

Recent YOLO advancements, particularly YOLOv10 and YOLOv11, feature improved backbone designs, enhanced feature aggregation, and optimized training strategies, leading to higher detection accuracy and faster inference. YOLOv10, for instance, has shown notable gains in precision and recall for real-time pedestrian detection under complex traffic conditions [3] Similarly, YOLOv11 has demonstrated superior performance in vehicle and pedestrian detection tasks, making it a promising candidate for intelligent transportation systems [2]. Despite these advances, the application of YOLOv10 and YOLOv11 to the combined challenge of pedestrian and animal safety at traffic intersections remains underexplored.

This research aims to conduct a comparative analysis of YOLOv10 and YOLOv11 for pedestrian and animal detection at intersections and evaluating performance across multiple metrics including mean Average Precision (mAP), recall and precision.

### 1.1.Background and Related Works

Research on intelligent transportation systems has increasingly adopted deep learning models, particularly YOLO, to address diverse safety challenges.

Early works focused on integrating YOLOv3 into driver assistance systems. For instance, Kankaria et al [7] developed a driver alert system combining traffic sign, light, and pedestrian detection using YOLOv3. Trained on GTSRB, COCO, and Bosch datasets, the model achieved 91.12% accuracy at 30 fps, outperforming SSD, R-FCN, and R-CNN, and demonstrated feasibility for real-time deployment on low-cost hardware.

Although these approaches are building on static traffic elements, later studies expanded the scope toward animal detection and nighttime scenarios.

Rajesh and Manivannan [8] introduced a YOLOv8n-based system for automatic traffic sign and animal detection to mitigate human-animal road conflicts. Using a custom dataset of Indian traffic signs and animal images (monkey, deer), the model achieved 99.3% accuracy and 73.7% mAP, demonstrating strong real-time performance and potential for preventing wildlife-related accidents.

Similarly, Parkavi et al. [9] enhanced nighttime animal detection on highways using an optimized YOLOv5 model combined with CLAHE and Robust Retinex preprocessing. The system achieved 0.923 precision, 0.773 recall, 0.802 mAP50, and 0.567 mAP50-95, outperforming conventional infrared and thermal imaging methods, though challenges remain under adverse weather conditions and just focus on specific types of animals.

Beyond static detection, researchers began to address accident detection and prediction.

In [5], the authors proposed an Automated Accident Detection System (AADS) leveraging a fine-tuned YOLOv10 model (YOLOv10 outfitted with Dw-Conv, C3Ghost, and PSA.) for nighttime surveillance. The system achieved 72.3% mAP, outperforming the baseline YOLOv10 in accident detection tasks, and demonstrated the potential of deep learning-based surveillance to enhance emergency response efficiency.

Parallel efforts focused on traffic monitoring tasks such as speed detection and safe distance estimation.

Syachputra et al. [6], integrated YOLOv11 with DeepSORT to detect speeding vehicles and monitor safe distances in real traffic. Their system achieved 87.3% mAP and 81.5% F1-Score, outperforming prior studies limited to detection or counting, and demonstrated robustness even with limited and low-quality traffic data.

Other studies emphasized collision prevention frameworks and pedestrian safety.

Vijayan et al. [12] proposed a collision detection and prevention framework using YOLOv8 on dash cam data. The system focused on six traffic-related categories (person, car, bike, motorcycle, bus, truck) and integrated night vision technology to detect pedestrians, cyclists, and animals in low-light conditions, thereby enhancing road safety and supporting autonomous driving applications.

Qu et al. [15] developed a YOLOv5-based vision system to detect pedestrian red-light violations. The framework integrates real-time alerts to guide pedestrians back to safe zones, aiming to reduce accidents caused by jaywalking and distracted crossing behaviors.

Li et al. [16] proposed an enhanced YOLOv10 framework for real-time pedestrian detection in autonomous vehicles. By integrating EfficientNet, C2F-DM modules, BiFormer attention, and a multi-

scale fusion detection head, the model achieved efficient multi-scale pedestrian detection even in complex backgrounds, highlighting its suitability for intelligent transportation systems.

Finally, comprehensive reviews and architectural improvements highlighted the evolution of YOLO toward more scalable and accurate frameworks.

Li et al. [13] introduced YOLO11-AMF, an improved YOLO11 framework integrating Mamba-Like Linear Attention, Asymptotic Feature Pyramid Network, and Focaler-IoU loss. The model achieved 96.5% precision, 82.9% recall, 90.0% mAP50, and 66.0% mAP50-95, outperforming baseline YOLO11n, and demonstrated robustness for real-time accident detection under diverse traffic conditions.

Sadik et al. [14] evaluated YOLOv8 and RT-DETR models for real-time vehicle and pedestrian detection in complex urban environments. Their study demonstrated that YOLOv8-Large achieved superior performance, particularly in pedestrian recognition, establishing new benchmarks for intelligent traffic monitoring and autonomous driving systems.

Das et al. [11] conducted a comprehensive review of YOLO models from v7 to v11 for real-time vehicle and pedestrian detection. Their study showed that YOLOv11 achieved 88% mAP and 45 FPS, outperforming earlier YOLO versions and other state-of-the-art methods, highlighting its scalability and robustness for intelligent transportation systems.

## II. MATERIALS AND METHOD

In this study, the latest versions of the YOLO architecture—YOLOv10 and YOLOv11—were applied to the COCO 2017 dataset, which includes 80 object categories. The experiments focused specifically on detecting persons and animals, given their relevance to traffic safety applications.

The experimental workflow contains key stages including model training, validation, and performance evaluation under different conditions. Each model was evaluated using metrics such as mean average precision (mAP), recall, training loss, and confusion matrix analysis.

The overall structure of these processes is shown in figure1, which outlines methodology adopted for YOLO versions.

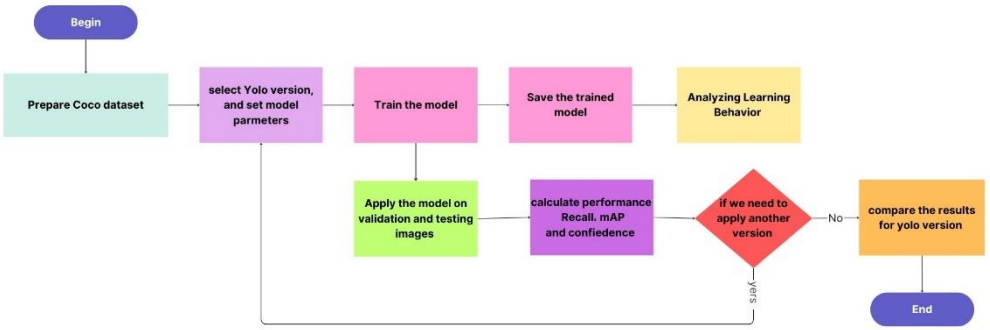


Fig. 1 Workflow process for apply many versions of Yolo on the COCO dataset

This flowchart describes implementing a YOLOv10 or YOLOv11 object detection model, starting with training on the COCO dataset, followed by generating detection outputs, and concluding with performance evaluation using standard metrics.

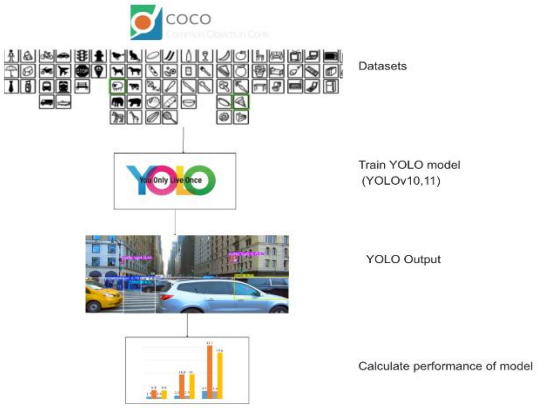


Fig. 2 Workflow of YOLOv10/11 object detection using the COCO dataset

### *A. Dataset*

The COCO (Common Objects in Context) dataset is a widely recognized benchmark for object detection [18], segmentation, and captioning tasks [10]. It contains over 330,000 images. These categories include vehicles (cars, motorcycles, bicycles etc), pedestrians (person class), and animals (cat, dog, elephant, zebra ..etc), which are particularly relevant to this study focusing on road safety. The dataset COCO 2017 is divided into training, validation and testing [17] as shown in table 1.

Table 1. Dataset Split Overview

| **Data** | **COCO 2017** |
|---|---|
| Training | 118k |
| Validation | 5k |
| Testing | 40k |

### *B. Training Parameters*

The training of YOLOv10 and YOLOv11 was conducted in Google Colab using a Tesla T4 GPU (16 GB). Implementation employed Python 3.12.6 with PyTorch 2.0, CUDA 11.3, and cuDNN 8.6.0. For reproducibility, experiments adopted a batch size of 32, 100 epochs, a learning rate of 0.001 with cosine annealing, AdamW optimizer, and input resolution of 640×640 pixels and data augmentation including random scaling, horizontal flipping, color jittering, and MixUp.

## III. RESULTS

In this section, we present the experimental results obtained from evaluating YOLOv10 and YOLOv11 on pedestrian and animal detection tasks. The assessment relies on several key performance indicators, including precision, mean average precision (mAP) and recall and the convergence behaviour of training loss. The primary objective is to compare the two versions in terms of their ability to enhance detection accuracy, computational efficiency, and real-time performance.

### *A. Comparison of Training*

The comparison of training loss reveals distinct convergence behaviors between YOLOv10 and YOLOv11. YOLOv10 shows steady and smooth convergence through its hybrid design, while YOLOv11 achieves faster convergence and consistently lower loss values. This improvement stems from innovations such as Sparse Attention and Dynamic Inference Scaling, enabling greater efficiency and reduced computational overhead. As illustrated in fig 3, both models converge effectively, but YOLOv11 maintains a clear advantage in training speed and overall loss reduction.

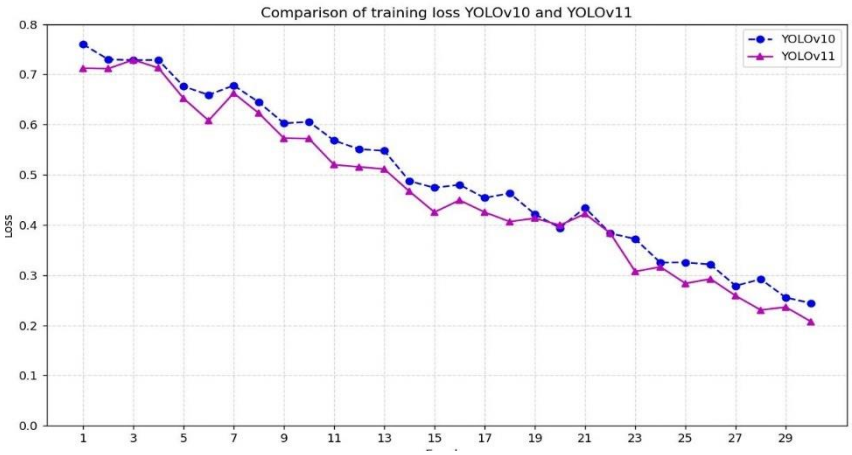


Fig. 3 Comparison of training loss for YOLOv10 and YOLOv11

### *B. Comparison of MAP*

Mean Average Precision (mAP) is a key metric for evaluating object detection models in multi-object scenarios. The comparison between YOLOv10 and YOLOv11 shows that both improve steadily with training, but YOLOv11 consistently achieves higher accuracy. YOLOv10 reached 84% mAP, while YOLOv11 attained 89%, supported by innovations such as Sparse Attention and Dynamic Inference Scaling. These features, combined with advanced augmentation and knowledge distillation, enable YOLOv11 to converge faster, maintain stability, and outperform YOLOv10 in pedestrian and animal detection tasks, as shown in fig 4.

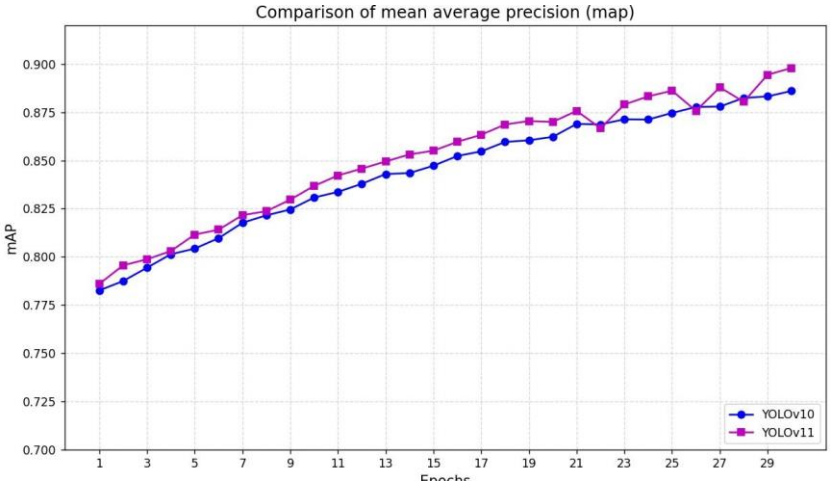


Fig. 4 Comparison of mean average precision (map) for YOLOv10 and YOLOv11

### *C. Comparison of Confusion Matrix*

A confusion matrix is a key tool for evaluating classification models, summarizing predictions through true positives, true negatives, false positives, and false negatives. It provides insights into accuracy, precision, recall, and class-specific errors. Confusion matrices reveal misclassifications among categories with strong diagonal dominance indicating high correct prediction rates and overall model robustness.

Below is the confusion matrix on the validation set for YOLOv10 in fig 5 and for YOLOv11 in fig 6.


Fig. 5 Confusion Matrix on the validation set for YOLOv10

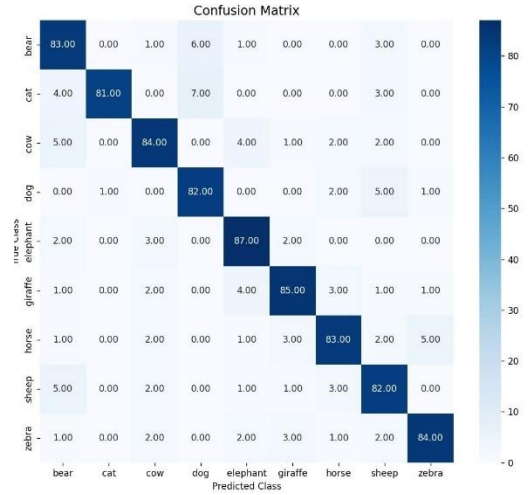


Fig. 6 Confusion Matrix on the validation set for YOLOv11

### D. *Comparison of Performance Metrics (mAP, Precision, Recall)*

As shown in table 2 summarizing the comparative performance of YOLOv10 and YOLOv1, YOLOv11 achieved consistently higher values in all metrics, reflecting its improved detection accuracy and robustness compared to YOLOv10.

Overall, YOLOv11 provides a more reliable detection framework for real-time applications where both speed and accuracy are critical.

Table 2. map and Precision and Recall for YOLO

| Model | mAP | Recall | Precision |
|---|---|---|---|
| YOLOv10 | 84 | 83 | 80 |
| YOLOv11 | 89 | 87 | 85 |

## IV. DISCUSSION

The comparative evaluation of YOLOv10 and YOLOv11 highlights the growing potential of deep learning in enhancing pedestrian and animal safety at traffic intersections. Results demonstrate that YOLOv11 consistently outperforms YOLOv10 in terms of mean Average Precision (mAP) and inference speed, achieving a more favourable balance between accuracy and efficiency. This improvement can be attributed to architectural refinements in YOLOv11, which reduce computational overhead while maintaining robust detection capabilities under diverse urban conditions.

our work leverages the latest versions of YOLO, enabling improved detection accuracy and faster inference in real-world traffic scenarios.

Most earlier studies focused mainly on detecting pedestrians or monitoring vehicles. In our work, we also include animals as an important category. This makes the system more useful for improving safety not only in busy city intersections but also in rural areas where animals often cross the roads.

Despite these advances, several challenges remain. YOLOv10 and YOLOv11 do not work as well in tough conditions like heavy rain, fog, or low light. To overcome these limits, the system could use extra data sources such as thermal cameras or LiDAR to support visual detection.

## V. CONCLUSION

The study compared the performance of YOLOv10 and YOLOv11 in detecting pedestrians and animals using the COCO 2017 dataset. The results showed that YOLOv11 consistently outperformed YOLOv10 in terms of mean Average Precision (mAP), aided by Sparse Attention and Dynamic Inference Scaling, and achieved faster training loss convergence. Confusion matrix analysis confirmed the robustness of YOLOv11, which exhibited fewer misclassifications and stronger diagonal dominance. Although YOLOv10 delivered stable performance, YOLOv11 demonstrated superior generalisation and efficiency, rendering it more suitable for real-time traffic safety applications.

This comparative evaluation highlights the growing potential of deep learning to enhance pedestrian and animal safety at traffic intersections. The results demonstrate that YOLOv11 consistently outperforms YOLOv10 in terms of mean average precision (mAP) and inference speed, achieving a favourable balance between accuracy and efficiency. This improvement can be attributed to architectural refinements in YOLOv11 that reduce computational requirements.

## REFERENCES


[1] J. Redmon, S. Divvala, R. Girshick, and A. Farhadi, *You only look once: Unified, real-time object detection*: arXiv preprint arXiv:1506.02640, 2016.

[2] J. Mo, G. Wu, and R. Li, *An Enhanced YOLOv11-based Algorithm for Vehicle and Pedestrian Detection in Complex Traffic Scenarios*, in Proc. 2025 IEEE 6th International Seminar on Artificial Intelligence, Networking and Information Technology (AINIT), Shenzhen, China, Apr. 11–13, 2025, pp. 1–6, doi: 10.1109/AINIT65432.2025.11035514.

[3] Y. Li, W. Leong, and H. Zhang, *YOLOv10-Based Real-Time Pedestrian Detection for Autonomous Vehicles,* in Proc. 2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, Sept. 3–5, 2024, pp. 1–6, doi: 10.1109/ICSIPA62061.2024.10686546..

[4] S. You, Z. Gu, and K. Zhu, *Pedestrian detection method based on improved YOLOv5,* Systems Science & Control Engineering: An Open Access Journal*, vol. 12, no. 1, pp. 1–13, 2024, doi: 10.1080/21642583.2023.2300836.

[5] M. Karthikeyan, P. R. Kshirsagar, T. K. Tak, K. Lalit, K. Upreti, and R. Jain, *AADS: An Automated Accident Detection and Nighttime Surveillance System Using Fine-Tuned YOLOv10 Deep Learning Techniques*, in Proc. 2025 International Conference on Intelligent Control, Computing and Communications (IC3), Mathura, India, Feb. 13–14, 2025, pp. 1–6, doi: 10.1109/IC363308.2025.10956514.

[6] D. I. A. Syachputra, F. Hamami, and I. Darmawan, *Optimizing Traffic Safety: YOLOv11 and DeepSORT for Speed Detection and Safe Distance Monitoring*, in Proc. 2025 4th International Conference on Electronics Representation and Algorithm (ICERA), Yogyakarta, Indonesia, Jun. 12, 2025, doi: 10.1109/ICERA66156.2025.11087337.

[7] R. V. Kankaria, S. K. Jain, P. Bide, A. Kothari, and H. Agarwal, *Alert System for Drivers based on Traffic Signs, Lights and Pedestrian Detection*, in Proc. 2020 International Conference for Emerging Technology (INCET),Belgaum, India, Jun. 5–7, 2020, pp. 1–6,doi: 10.1109/INCET49848.2020.9154167.

[8] R. Rajesh and P. V. Manivannan, *Automatic traffic sign, animal detection, and recognition using YOLOv8n to avoid human-animal road conflicts*, Engineered Science, vol. 31, pp. 1–12, 2024, doi: 10.30919/espub.esj.1192.

[9] K. Parkavi, A. Ganguly, S. Sharma, A. Banerjee, and K. Kejriwal, *Enhancing Road Safety: Detection of Animals on Highways During Night*, IEEE Access, vol. 13, pp. 36877–36896, Feb. 25, 2025,doi: 10.1109/ACCESS.2025.3545490.

[10] Tsung-Yi Lin, M. Maire, S. Belongie, L. Bourdev, R. Girshick, J. Hays, P. Perona, D. Ramanan, C. L. Zitnick, and P. Dollár, *Microsoft COCO: Common Objects in Context*, arXiv:1405.0312 [cs.CV], Feb. 2015, doi: 10.48550/arXiv.1405.0312.

[11] S. Das, M. I. Ibrahem, and M. M. Fouda, *A Comprehensive Review on Real-Time Vehicle and Pedestrian Detection Using YOLO,*, in Proc. 2025 IEEE 4th International Conference on Computing and Machine Intelligence (ICMI), MI, USA, Apr. 5–6, 2025, pp. 1–7, doi: 10.1109/ICMI65310.2025.11141119.

[12] N. Vijayan, S. Pandey, and P. R. M, *Collision Detection and Prevention for Automobiles using Machine Learning*, in Proc. 2024 International Conference on Trends in Quantum Computing and Emerging Business Technologies (TQCEBT) Pune, India, Mar. 22–23, 2024, pp. 1–6, doi: 10.1109/TQCEBT59414.2024.10545080.

[13] W. Li, L. Huang, and X. Lai, *A Deep Learning Framework for Traffic Accident Detection Based on Improved YOLO11*, Vehicles, vol. 7, no. 3, p. 81, Aug. 2025, doi: 10.3390/vehicles7030081.
[14] M. N. Sadik, T. Hossain, and F. Sayeed, *Real-Time Detection and Analysis of Vehicles and Pedestrians using Deep Learning*, arXiv preprint arXiv:2404.08081, Apr. 2024, doi: 10.48550/arXiv.2404.08081.
[15] Z. Qu, W. Kong, H. Liu and L. Deng, *Pedestrian Red-Light Violation Detection Based on Deep Learning Vision* , in Proc. 2023 IEEE 6th International Conference on Automation, Electronics and Electrical Engineering (AUTEEE) , Shenyang, China, Dec. 15–17, 2023, pp. 1–6, doi: 10.1109/AUTEEE60196.2023.10407497.
[16] Y. Li, W. Leong, and H. Zhang, *YOLOv10-Based Real-Time Pedestrian Detection for Autonomous Vehicles*, in Proc. 2024 IEEE 8th International Conference on Signal and Image Processing Applications (ICSIPA), Kuala Lumpur, Malaysia, Sep. 3–5, 2024, pp. 1–6, doi: 10.1109/ICSIPA62061.2024.10686546.
[17] COCO Consortium, *COCO Dataset – Download*: COCO Dataset, [Online]. Available: https://cocodataset.org/#download. [Accessed: Nov. 22, 2025].
[18] A. Bochkovskiy, C.-Y. Wang, and H.-Y. M. Liao, *YOLOv4: Optimal Speed and Accuracy of Object Detection*: arXiv preprint arXiv:2004.10934, 2020.