Proceedings of Automotive Science and Technology 1(1): 1-6 (2025)

ISASTECH2025

Real-Time Detection of Heavy-Duty Vehicle Lane Violations: A Comparative Analysis of Yolov11 and Yolov12

Samet Sülün[1*], Ravzanur Tok[2], Furkan Çeken[3], Ömer Faruk Boyraz[4]

[1 2 3 4] R&D Center, Anadolu Isuzu Automotive Industry and Trade Inc., Kocaeli, 41435, Türkiye

Abstract

The unnecessary occupation of the left lane on highways and main roads by heavy-duty vehicles (such as trucks and buses) poses a significant traffic safety problem. Due to their low speeds and large sizes, these vehicles disrupt traffic flow, forcing other drivers into sudden braking or lane-changing maneuvers, thereby increasing the risk of accidents. Studies conducted indicate that left-lane violations by commercial vehicles contribute to nearly 15% of traffic accident globally.

In this study, an AI-based prototype system was developed to detect and prevent left-lane violations by heavy-duty vehicles. The system employs a vehicle-mounted camera that analyzes lane markings, road barriersand surrounding vehicles in real time. The captured data were processed using two state-of-the-art object detection models, YOLOv11 and YOLOv12, and their performances were comparatively evaluated.

Experimental results demonstrate that YOLOv12 outperforms YOLOv11 in terms of overall performance, achieving higher values in both mAP@50 (85.6%) and precision (89%). YOLOv12 yielded superior detection results for the car (93.6%), bus (72.6%), and lane violation (96.6%) classes. However, for the truck class, YOLOv11 achieved slightly better accuracy (86.0%) compared to YOLOv12 (80.6%). Training curves further revealed that YOLOv12 stabilized its losses more rapidly and exhibited a more consistent learning process.

In conclusion, the proposed system provides real-time detection of left-lane violations and delivers visual and auditory warnings to drivers, thereby encouraging safer lane usage. The comparative analysis of YOLOv11 and YOLOv12 highlights that YOLOv12 generally offers superior performance, while class-specific variations underline the importance of model selection in traffic safety applications.

Keywords: Heavy-Duty Vehicles, Lane Occupation, Safe Driving, Traffic Congestion, YOLOv11, YOLOv12

http://dx.doi.org/10.29228/pastech.89117

1. Introduction

Increasing urbanization, the expansion of highway networks, and the rapid growth in the number of motor vehicles have turned traffic safety issues into a serious global threat. According to the reports of the World Health Organization (2021), approximately 1.3 million people lose their lives in traffic accidents each year worldwide, and millions more are injured [1]. Lane violations by heavy-duty vehicles (such as trucks and buses) on highways and major roads can lead to sudden braking or lane-change maneuvers by high-speed vehicles, resulting in chain-reaction accidents [2]. Analyses conducted globally indicate that left-lane violations by commercial vehicles contribute to nearly 15% of traffic accidents, and that such violations often trigger dangerous driving behaviors [3]. In recent years, there has been growing interest in the use of artificial intelligence and image-processing–based systems to enhance traffic safety. In such systems, cameras and sensors integrated into vehicles collect environmental data and analyze the vehicle's lane position, the movements of surrounding vehicles, lane markings, and roadside barriers [4]. The collected data are processed by deep learning algorithms that operate with high accuracy and low latency, enabling real-time object detection. At this point, algorithms from the YOLO (You Only Look Once) family stand out in terms of both speed and accuracy, allowing traffic elements to be classified effectively [5].

Smart systems developed in the literature not only regulate traffic flow but also play a critical role in enhancing highway safety and preventing accidents caused by heavy-duty vehicles. Moreover, with the expansion of deep learning–based models, vehicle-to-vehicle (V2V) communication can be supported. In this way, heavy-duty vehicles that violate the left lane could warn not only their own drivers but also surrounding vehicles, enabling safer lane-change maneuvers [6].

This study aims to develop a system designed to detect and reduce left-lane occupation by heavy-duty vehicles in Türkiye's traffic conditions. The proposed system analyzes images captured by vehicle-mounted cameras to determine whether a heavy-duty vehicle is occupying the left lane, and in the case of a violation, it directs the driver to move to the right lane through visual and auditory warnings. The novelty of this study lies in the comparative evaluation of the YOLOv11 algorithm and the most recent version, YOLOv12, within the system. In this way, class-based performance differences between various YOLO versions are revealed, highlighting the importance of selecting the most suitable model for traffic safety applications.

2. Literature

In recent years, numerous studies have been conducted to reduce the impact of heavy-duty vehicles on traffic flow and improve traffic safety. Within this scope, advanced systems utilizing vehicle-mounted sensors and cameras for object detection have been widely explored.

Weng and Meng [7] analyzed the lane-changing behaviors of heavy-duty vehicle drivers and examined the impact of lane violations on accidents, emphasizing that proper lane usage by heavy-duty vehicles can ensure safer traffic flow.

The YOLO algorithm developed by Redmon and Farhadi [5] achieved successful results in detecting traffic objects in real time with high accuracy, highlighting the importance of rapid object recognition for traffic safety. Such deep learning–based object detection systems can instantly identify situations such as lane violations by large vehicles and provide visual and auditory warnings to the driver.

Additionally, in a study conducted by Zhang and Du [8], deep convolutional neural networks were used to enhance the effectiveness of lane tracking for heavy-duty vehicles, demonstrating that these technologies have the potential to reduce accidents by up to 20%.

Moreover, Chen et al. [9] showed that vehicle-to-vehicle (V2V) communication–based solutions play a significant role in mitigating accidents caused by heavy-duty vehicles, emphasizing that such communication protocols should be widely adopted to ensure traffic safety.

3. Methods

3.1. Construction and Characteristics of the Collected Dataset

The dataset used in this study was created by compiling videos recorded under real driving conditions. To increase the diversity and realism of the dataset, recordings obtained from different sources were combined, thereby covering a wide range of driving scenarios.

The recordings conducted with Anadolu ISUZU's test vehicles were planned to include daytime and nighttime driving, sunny and rainy weather conditions, urban traffic, and intercity roads. Additionally, the vehicle was intentionally driven across all lanes to include lane changes, overtaking maneuvers, and varying road dynamics in the dataset. In this way, a rich and comprehensive data pool was obtained that reflects the real driving behavior of heavy-duty vehicles.

Sample images selected from the prepared dataset are presented in Figure 1, offering a visual representation of different environmental conditions and road types. This diversity enables a more accurate evaluation of the real-time lane detection performance of the YOLOv11 and YOLOv12 algorithms.

In conclusion, the constructed dataset covers a wide range of vehicle speeds, varying illumination conditions, and diverse environmental scenarios, providing a suitable infrastructure for testing the models under real-world conditions.

Figure 1. Data Recorded Under Different Conditions (Daytime, Tunnel, Nighttime, Rainy)

3.2. Materials

The deep learning–based system's AI model training was carried out on a computer equipped with an Intel Core i7-vPRO processor, an Nvidia GeForce RTX 2080 GPU, and 32 GB of RAM. After training, the developed models were tested in real time both on the computer and on the embedded device.

The Nvidia Jetson Nano embedded device used in the system (Figure 2) features a 4-core ARM A57 CPU, a 128-CUDA-core NVIDIA Maxwell™ GPU, and 4 GB of 64-bit LPDDR4 memory with a bandwidth of 25.6 GB/s. This device provides high-performance processing by feeding real-time camera input into deep learning models. Additionally, for real-time image processing, the NVIDIA Jetson Nano was paired with a CSI camera equipped with a Sony IMX219 sensor and a 77-degree field of view. This camera uses the MIPI (Mobile Industry Processor Interface) standard and connects directly to the device, ensuring high-speed and low-latency data transmission. Furthermore, its direct connection eliminates the need for an additional interface or conversion board, enabling faster data processing [10, 11, 12].

Figure 2. Nvidia Jetson Nano and CSI Camera

3.3. Methods

After obtaining sufficient data, road lanes and vehicles were annotated using the Roboflow tool, preparing the dataset for model training. Subsequently, data augmentation techniques were applied during the AI training phase of the developed system. These augmentation operations included angle variations, scaling, saturation adjustments, exposure modifications, and hue shifts.

While angle variations and scaling were integrated directly into the model training process, the other parameters introduced random changes in image color saturation, brightness (exposure), and hue, thereby improving the model's robustness and adaptability under varying conditions.

These enhancements contributed to maintaining high model performance regardless of environmental or lighting variations. The parameters and their corresponding values were added to the configuration file, and the parameters used are explained in detail in Table 1.

Table 1. YOLOv11 – YOLOv12 Configuration Parameters and Data Augmentation Settings

Width × Height (pixels)	416 x 416
Batch Size	32
Rotation Angle (degrees)	8
Saturation	1.5
Exposure	1.5
Hue	0.1

YOLOv11 is one of the latest versions in the YOLO series developed by Ultralytics. Key architectural innovations include the C3K2 (Cross Stage Partial with kernel size 2) blocks, the SPPF (Spatial Pyramid Pooling – Fast) module, and the C2PSA (Convolutional block with Parallel Spatial Attention) unit. These components enable the model to achieve improved performance, particularly in detecting small objects and capturing fine-grained spatial details. With a broad range of model variants (from nano to x-large), YOLOv11 offers scalability to achieve both parameter efficiency and high mAP. Additionally, in real-time applications, the largest model in the series (YOLOv11-x) has been reported to achieve an impressive 54.7% mAP on the COCO dataset. This level of performance represents a significant advancement in both accuracy and efficiency compared to earlier models [13].

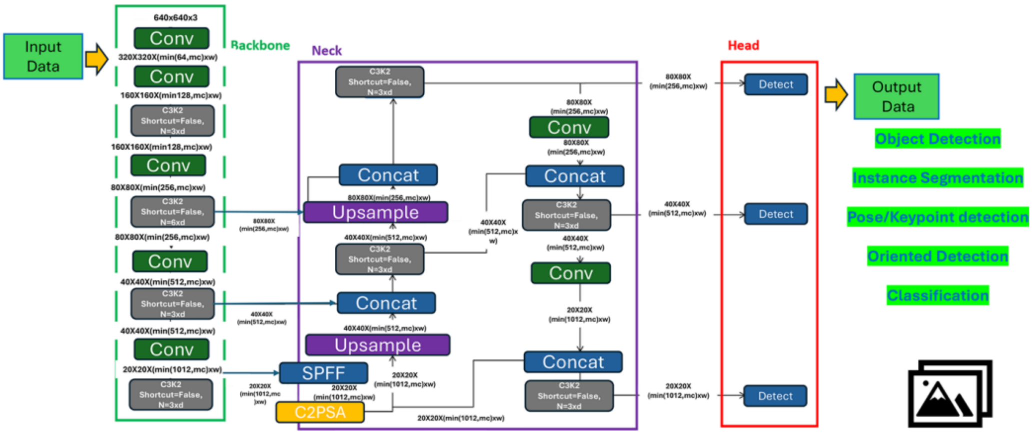


Figure 3. YOLOv11 Architecture [13]

YOLOv12 is an advanced version that builds upon YOLOv11, incorporating attention-based enhancements that further improve both architecture and performance. The model integrates several advanced components, including the Area Attention module, R-ELAN (Residual Efficient Layer Aggregation Networks), 7×7 depthwise separable convolutions, and a FlashAttention-based attention structure alongside CNN-based approaches. Through these additions, the model benefits from the contextual modeling advantages provided by attention mechanisms while maintaining high real-time processing speed.

For example, the YOLOv12-N model achieved a mAP of 40.6% with only 1.64 ms latency on an NVIDIA T4 GPU, demonstrating performance improvements of 1.2% over YOLOv11-N and 2.1% over YOLOv10-N. These advancements indicate that the model elevates both accuracy and efficiency to a new level [13].

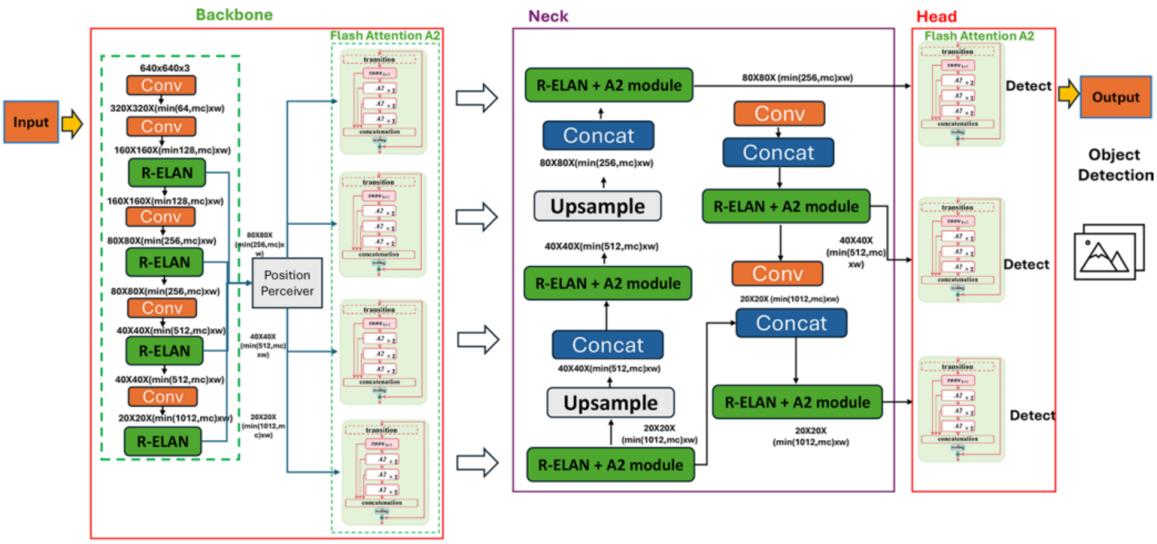


Figure 4. YOLOv12 Architecture [13]

Image and video analysis can be performed using OpenCV, an open-source library that supports deep learning models and includes image-processing algorithms for object detection. In this study, after detecting objects in three separate classes, OpenCV was used to determine lane occupation by the vehicle and to ensure safe lane-change maneuvers.

The camera integrated into the vehicle was positioned to center the front view of the vehicle. The images obtained from the camera were virtually divided into two sections with a vertical line, and specific regions on the right and left sides of this line were monitored to determine whether the vehicle was moving within the correct lanes. If dashed lane markings were detected on both the right and left sides, it was concluded that the vehicle was not in the far-left lane. If no dashed markings were detected on the left side of the line, the vehicle was determined to be in the far-left lane. When the vehicle was in the left lane, the next step was to determine whether it was appropriate to merge into the right lane. If the scenario was suitable, a warning was issued to the driver.

For this part, the image was horizontally divided into upper and lower sections. If vehicles ahead on the right side remained in the upper section of the frame, the area was marked as safe. If they appeared in the lower section, the area was marked as unsafe, as this indicated a reduced following distance.

After the initial analysis, the vehicle was determined to be in the left lane. In the subsequent analysis, if the right lane was marked as a safe zone, it was concluded that the vehicle was occupying the left lane and that the warning mechanisms needed to be activated. In this stage, messages transmitted from the onboard computer were integrated into the vehicle's dashboard software to provide visual alerts to the driver. Additionally, a buzzer module was used to deliver audible warnings at specific frequencies and durations. When the driver moved to the right lane, dashed lane markings would be detected on both the right and left sides, thereby deactivating the warning mechanisms.

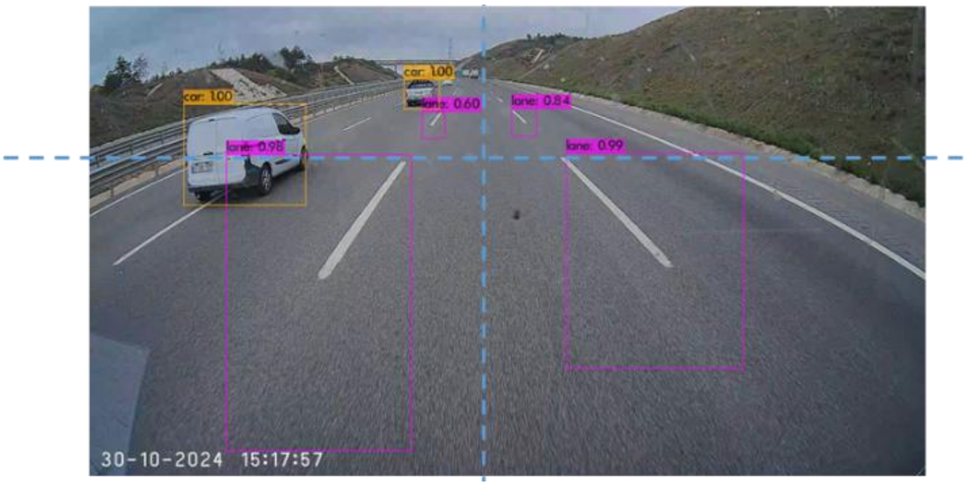
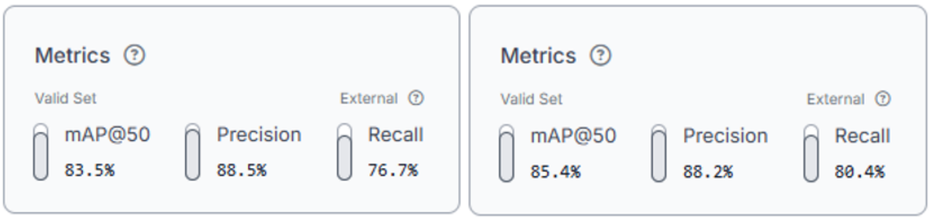


Figure 5. Sample Image Processing Integrated into the Highway Object Detection Model

4. Conclusions

In this study, a system was developed to detect left-lane violations by heavy-duty vehicles, and the performance of the system was evaluated comparatively using the YOLOv11 and YOLOv12 deep learning algorithms. The results show that both models can accurately detect cars, trucks, and lane markings. However, differences become noticeable in class-based comparisons.

The YOLOv12 model achieved higher overall performance compared to YOLOv11, reaching a mAP@50 of 85.6% and a precision of 89%. YOLOv12 demonstrated superior accuracy in the car (93.6%), bus (72.6%), and lane-violation (96.6%) classes. On the other hand, YOLOv11 outperformed YOLOv12 in the truck class, achieving 86.0% accuracy compared to YOLOv12's 80.6%. Examination of the training curves indicates that YOLOv12 stabilizes its losses more rapidly and exhibits a more consistent learning process. Figure 6 presents the metrics obtained from the training of the YOLOv11 and YOLOv12 models.



(a) (b)

Figure 6. (a) YOLOv11 Metric Results (b) YOLOv12 Metric Results

In addition to the model outputs, lane detection for the right and left lanes and safe-area marking were implemented using image-processing libraries. The developed system was supported with a visual alert mechanism that displays icons on the vehicle's dashboard, as well as an audible warning mechanism using a buzzer. In this way, left-lane violations were detected instantly, encouraging drivers to return to the right lane.

In conclusion, the comparative analysis of YOLOv11 and YOLOv12 revealed that YOLOv12 generally provides superior performance, whereas YOLOv11 offers advantages specifically for the truck class. This finding highlights the importance of selecting models based on class-specific requirements in system design and presents an innovative solution to reduce left-lane violations caused by heavy-duty vehicles.

References

[1] Global status report on road safety (2021). https://www.itf-oecd.org/sites/default/files/docs/irtad-road-safety-annual-report-2021.pdf

[2] Weng, J., & Meng, Q. (2011). *Analysis of driver behavior in freeway lane-changing events using vehicle trajectory data. Journal of Transportation Safety & Security,* 3(2), 117–129

[3] Leduc, G. (2009), *Road traffic data: Collection methods and applications. Working Papers on Energy, Transport and Climate Change,* No. 1, JRC, European Commission.

[4] Zhang, L., & Du, B. (2017). *High-level semantics extraction from traffic images using deep convolutional neural networks. Neurocomputing, 239, 64–73.*

[5] Redmon, J., & Farhadi, A. (2018). *YOLOv3: An incremental improvement.* arXiv preprint arXiv:1804.02767.

[6] Bongiorno, C., Wilson, R., & Wright, S. (2022). *Advanced vehicle monitoring for heavy vehicles: Addressing lane violations using AI systems. IEEE Transactions on Intelligent Transportation Systems, 23(9), 14532–14541.*

[7] Weng, J., & Meng, Q. (2011). *Analysis of driver behavior in freeway lane-changing events using vehicle trajectory data. Journal of Transportation Safety & Security,* 3(3), 153-167.

[8] Zhang, K., & Du, X. (2017). *Lane detection based on deep convolutional neural network. IEEE Intelligent Transportation Systems Conference (ITSC),* 1-6.

[9] Chen, L., Englund, C., & Wang, H. (2019). *Cooperative intelligent transport systems: A V2X-based approach for reducing traffic accidents. IEEE Transactions on Intelligent Transportation Systems,* 20(1), 178-186.

[10] Jetson Nano [Online]. Available: *https://www.nvidia.com/tr-tr/autonomous-machines/embedded-systems/jetson-nano/product-development/*

[11] Jetson Nano Brings AI Computing to Everyone, *https://developer.nvidia.com/blog/jetson-nano-ai-computing/*

[12] Kwang-Hyun Baek,Kyusam Lim, Gye Su Kim, Suki Kim (2010). *A multi-lane MIPI CSI receiver for mobile camera applications,* IEEE Transactions on Consumer Electronics

[13] Sapkota, Ranjan, et al. "YOLO advances to its genesis: a decadal and comprehensive review of the You Only Look Once (YOLO) series." *Artificial Intelligence Review* 58.9 (2025): 274.