Mühendislik Bilimleri ve Tasarım Dergisi
14(2), 207-223, 2026
e-ISSN: 1308-6693

Journal of Engineering Sciences and Design
DOI: 10.21923/jesd.1825912

Araştırma Makalesi | Research Article

DEEP LEARNING-BASED VEHICLE BRAND AND MODEL RECOGNITION FOR MARKET ANALYSIS USING YOLOV8 AND YOLOV11 ON REAL-WORLD TRAFFIC DATA

Recep ARSLAN[1*], Merdan ÖZKAHRAMAN[2], Hilmi Cenk BAYRAKÇI[2]

[1]Isparta University of Applied Sciences, Institute of Graduate Education, Department of Mechatronic Engineering, Isparta, Türkiye
[2]Isparta University of Applied Sciences, Faculty of Technology, Department of Mechatronic Engineering, Isparta, Türkiye

Keywords

Deep Learning, Market Analysis, Object Detection, Vehicle Brand Recognition, YOLOv8 Algorithm, YOLOv11 Algorithm.

Abstract

The ability to analyze vehicle brand distribution from real-world traffic data has significant implications for market research and trend analysis. This study presents a deep learning-based approach for automated vehicle brand and model recognition using the YOLOv8 and YOLOv11 object detection architectures. The proposed system was trained and evaluated on a dataset of 5,213 labeled vehicle images, collected from actual road traffic footage to ensure real-world applicability. Performance assessments were conducted using mean Average Precision (mAP), precision, recall, and inference speed, demonstrating that YOLOv11 outperforms YOLOv8 in precision and object localization accuracy, achieving a mAP of 88.6%. However, YOLOv8 achieved a higher recall, detecting a broader range of vehicles, albeit with an increased false positive rate. In addition to model performance analysis, this study examines the distribution of different automobile brands in traffic, providing valuable insights for data-driven market analysis. The findings highlight the potential of deep learning-based vehicle detection as a tool for understanding brand presence and consumer trends in the automotive market.

DERİN ÖĞRENME TABANLI YOLOV8 VE YOLOV11 YAKLAŞIMLARI KULLANILARAK GERÇEK TRAFİK VERİLERİ ÜZERİNDE PİYASA ANALİZİ İÇİN ARAÇ MARKA VE MODEL TANIMA

Anahtar Kelimeler

Derin Öğrenme, Piyasa Analizi, Nesne Tespiti, Araç Marka Tanıma, YOLOv8 Algoritması, YOLOv11 Algoritması.

Öz

Gerçek dünya trafik verilerinden araç marka dağılımını analiz edebilme yeteneği, pazar araştırması ve eğilim analizi için önemli etkiler taşımaktadır. Bu çalışma, YOLOv8 ve YOLOv11 nesne tespit mimarilerini kullanarak otomatik araç marka ve model tanıma için derin öğrenme tabanlı bir yaklaşım sunmaktadır. Önerilen sistem, gerçek dünya uygulanabilirliğini sağlamak amacıyla gerçek yol trafiği görüntülerinden toplanmış 5.213 etiketli araç görüntüsünden oluşan bir veri kümesi üzerinde eğitilmiş ve değerlendirilmiştir. Performans değerlendirmeleri, Ortalama Doğruluk (mAP), kesinlik, duyarlılık ve çıkarım hızı kullanılarak gerçekleştirilmiş olup, YOLOv11'in kesinlik ve nesne konumlandırma doğruluğunda YOLOv8'i geride bıraktığı ve %88,6 Ortalama Doğruluk (mAP) elde ettiği gösterilmiştir. Ancak YOLOv8, daha geniş bir araç aralığını tespit ederek daha yüksek duyarlılık elde etmiş, bununla birlikte artmış bir yanlış pozitif oranına neden olmuştur. Model performans analizine ek olarak, bu çalışma trafikteki farklı otomobil markalarının dağılımını inceleyerek veri odaklı pazar analizi için değerli içgörüler sağlamaktadır. Bulgular, derin öğrenme tabanlı araç tespitinin otomotiv pazarında marka varlığını ve tüketici eğilimlerini anlamada bir araç olarak potansiyelini vurgulamaktadır.

Alıntı / Cite

Arslan, R., Özkahraman, M., Bayrakçı, H. C., (2026). Deep Learning-Based Vehicle Brand and Model Recognition for Market Analysis Using YOLOv8 and YOLOv11 on Real-World Traffic Data, Journal of Engineering Sciences and Design, 14(2), 207-223.

Yazar Kimliği / Author ID (ORCID Number)

R. Arslan, 0000-0002-0930-505X
M. Özkahraman, 0000-0002-3501-6497
H. C. Bayrakçı, 0000-0001-5064-7310

Makale Süreci / Article Process	
Başvuru Tarihi / Submission Date	18.11.2025
Revizyon Tarihi / Revision Date	24.02.2026
Kabul Tarihi / Accepted Date	17.04.2026
Yayım Tarihi / Published Date	30.06.2026

* İlgili yazar / Corresponding author: arslannrecep@gmail.com, +90-530-441-4226

DEEP LEARNING-BASED VEHICLE BRAND AND MODEL RECOGNITION FOR MARKET ANALYSIS USING YOLOV8 AND YOLOV11 ON REAL-WORLD TRAFFIC DATA

Recep ARSLAN[1†], Merdan ÖZKAHRAMAN[2], Hilmi Cenk BAYRAKÇI[2]
[1]Isparta University of Applied Sciences, Institute of Graduate Education, Department of Mechatronic Engineering, Isparta, Türkiye
[2]Isparta University of Applied Sciences, Faculty of Technology, Department of Mechatronic Engineering, Isparta, Türkiye

Highlights

- YOLOv11 improves brand recognition accuracy in real-world traffic images.
- Dataset of 5,213 labeled vehicles improves model generalization.
- YOLOv11 reduces false positives and enhances feature extraction.
- Results support market trend analysis using real traffic data.

Graphical Abstract

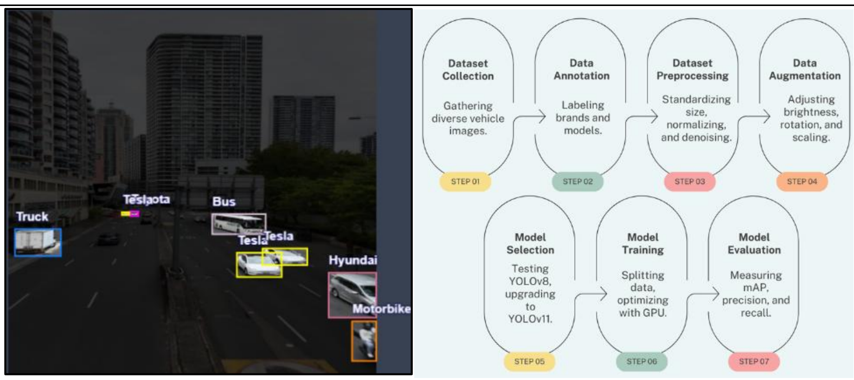

Figure. Overall workflow of the vehicle brand and model recognition system using YOLO architectures

Purpose and Scope

The purpose of this study is to develop an automated system capable of recognizing vehicle brands and models from real-world traffic images. The research aims to support market analysis by quantifying brand distribution and evaluating the performance of YOLOv8 and YOLOv11 for high-accuracy detection under diverse environmental conditions.

Design/methodology/approach

The study employs a dataset of 5,213 annotated vehicle images captured from real-world traffic footage. YOLOv8 and YOLOv11 architectures are trained using standardized preprocessing, extensive data augmentation, and optimized hyperparameters. Model performance is evaluated through mAP, precision, recall, and inference speed. The methodological framework integrates deep learning-based detection with brand distribution analysis to assess real-world applicability.

Findings

Results show that YOLOv11 outperforms YOLOv8 in precision, localization accuracy, and training efficiency, achieving a mAP of 88.6%. YOLOv8 maintains higher recall but yields more false positives. YOLOv11 demonstrates superior feature extraction capabilities, reduced classification errors, and stronger generalization across diverse environmental conditions. Brand distribution analysis reveals significant market trends observable directly from traffic data.

Research limitations/implications

The system's accuracy decreases under low-resolution footage, extreme angles, and severe weather conditions due to limited dataset representation. Increased dataset diversity, higher-resolution imagery, and multi-sensor integration are recommended for future improvements. These limitations suggest that real-world deployments must account for environmental variability and sensor quality.

Practical implications

The proposed model enables automated monitoring of brand presence in traffic, supporting market research, fleet analysis, and intelligent transportation systems. The system can be integrated into smart city infrastructures to improve traffic analytics, commercial decision-making, and vehicle trend forecasting.

Social Implications

Automated vehicle recognition may enhance urban planning, transportation policy, and market transparency. Insights gained from traffic-based brand distribution can guide infrastructure investments and support sustainability-focused mobility strategies. Enhanced monitoring tools may also contribute to public safety and more efficient resource allocation.

Originality

This study is among the first to employ YOLOv11 for detailed vehicle brand and model recognition using real-world traffic footage. The integration of high-accuracy detection with market trend analysis offers a novel perspective, providing value to researchers, automotive companies, and smart city planners. The optimized YOLOv11 workflow significantly improves performance compared to previous YOLO-based approaches.

† Corresponding author: arslannrecep@gmail.com, +90-530-441-4226

1. Introduction

The progressive integration of artificial intelligence (AI) and computer vision technologies has profoundly reshaped a wide spectrum of industries and research domains, ranging from intelligent transportation systems and autonomous driving technologies to urban surveillance and public safety management (Gulati and Srinivasan, 2019; Fathy and Siyal, 1995; Atkočiūnas et al., 2005). The ability to extract structured, high-value information from large volumes of unstructured visual data has enabled unprecedented advancements in automation, efficiency, and decision-making processes. One prominent application area that has attracted increasing scholarly and industrial attention is the automated detection and classification of vehicles, specifically the recognition of vehicle brands and models from real-world traffic environments (Ma et al., 2009; Gilmore and Elibiary, 1993; Chen et al., 2023).

Accurate and real-time vehicle brand and model recognition holds substantial potential for various practical applications, including traffic monitoring and regulation, vehicular flow analysis, law enforcement, insurance claim verification, and strategic market research for the automotive industry. Furthermore, understanding brand distribution trends through automated systems provides valuable insights into consumer behavior, regional market penetration, and mobility patterns, which can inform business strategies and policy-making at both local and national levels (Hu et al., 2015; Lu et al., 2025).

Traditionally, vehicle recognition tasks were predominantly addressed using classical machine learning techniques that relied heavily on handcrafted feature extraction and heuristic-based classification approaches. However, these methods often suffered from limited scalability and poor generalization in complex and dynamic real-world scenarios, especially under conditions involving occlusions, varying illumination, and diverse viewpoints (Ma and Boukerche, 2020; Yang, 2023; Adu-Gyamfi et al., 2017). The emergence of deep learning methodologies, particularly convolutional neural networks (CNNs), has significantly transformed this landscape, offering superior performance by automatically learning hierarchical and discriminative features from large and diverse datasets.

Within this context, the You Only Look Once (YOLO) family of object detection architectures has established itself as one of the most prominent and widely adopted frameworks for real-time object detection and classification tasks. Successive iterations of the YOLO models — culminating in the recent YOLOv8 and YOLOv11 versions — have progressively introduced architectural innovations, including improved feature extraction pipelines, advanced attention mechanisms, and optimized bounding box regression strategies, leading to marked improvements in detection accuracy, robustness, and computational efficiency (Hidayatullah et al., 2025; Jegham et al., 2024; Alruwaili et al., 2023; Bitwire and Han, 2024; Lee and Kim, 2024).

YOLOv8 offers a significant balance between inference speed and accuracy, making it highly suitable for applications requiring rapid processing with reasonable detection precision. Nevertheless, YOLOv11 represents a further refinement, incorporating multi-scale feature aggregation techniques, transformer-based contextual reasoning modules, and more sophisticated loss functions to enhance both localization and classification performance, particularly in challenging conditions involving partial occlusions and highly cluttered backgrounds (Bitwire and Han, 2024; Lee and Kim, 2024).

In light of these advancements, the primary objective of the present study is to develop and evaluate a deep learning-based system capable of performing robust and accurate vehicle brand and model recognition from real-world traffic data. To this end, a comprehensive dataset comprising 5,213 meticulously labeled vehicle images, captured under diverse environmental and lighting conditions, was constructed to ensure the generalizability and real-world applicability of the proposed system. The performances of YOLOv8 and YOLOv11 architectures are systematically compared using standard evaluation metrics, including mean Average Precision (mAP), precision, recall, inference speed, and computational efficiency.

Beyond the technical evaluation of object detection performance, this study also seeks to explore the broader commercial implications of vehicle recognition in market research and consumer behavior analysis. By analyzing the distribution of different automobile brands across various regions, valuable insights can be generated, aiding automotive manufacturers, marketers, and policy-makers in developing more data-driven and region-specific strategies. Unlike conventional object detection comparison studies, this work does not solely aim to report performance differences between YOLO versions. Instead, it evaluates how architectural improvements directly influence the reliability of traffic-based commercial analytics. The study frames vehicle brand recognition not only as a computer vision problem but also as a quantitative proxy for real-world market density estimation derived from passive traffic monitoring systems.

The structure of this paper is organized as follows: Section 2 provides a comprehensive description of the dataset preparation processes, model architectures, training methodologies, and experimental setup. Section 3 presents a detailed analysis of the experimental results, including a comparative assessment of the YOLOv8 and YOLOv11 models. Section 4 discusses the broader implications of the findings and identifies potential directions for future research aimed at enhancing automated vehicle brand recognition systems. Finally, Section 5 concludes the study by summarizing the main contributions and outlining prospective applications in intelligent transportation and commercial analytics domains.

2. Materials and Methods

In order to develop a robust and efficient vehicle brand and model recognition system, a comprehensive and meticulously designed methodology was implemented. This section outlines the data acquisition process, dataset composition, and the architectural framework employed to train and evaluate the deep learning models. A particular focus is placed on ensuring the real-world applicability of the system by utilizing diverse environmental conditions and realistic traffic scenarios. The following subsections detail the materials used and the methodological approach adopted to achieve high accuracy and generalizability in vehicle detection and classification tasks.

2.1. Dataset

The dataset used in this study consists of a total of 5,213 labeled vehicle images, carefully curated from multiple sources, including publicly available databases, real-world traffic surveillance footage, and manually captured images. The data collection process was designed to ensure diversity in environmental conditions, vehicle orientations, and lighting scenarios. Images were taken under various weather conditions, such as sunny, cloudy, and nighttime settings, to ensure that the trained model generalizes effectively across different real-world driving conditions. This variation enhances the robustness of the detection model by allowing it to learn vehicle characteristics under different lighting, background, and weather conditions. Figure 1 presents a selection of sample images from the dataset, showcasing various vehicle brands under different environmental conditions.

To ensure a balanced representation, the dataset contains vehicles from 19 distinct brands, with varying frequencies based on their prevalence in real-world road conditions. Toyota accounts for the largest proportion of vehicles, constituting 15.06% of the dataset, followed by trucks (7.75%) and Tesla vehicles (5.01%). Other brands such as BMW, Audi, and Mercedes are also represented, though with lower counts. The detailed distribution of vehicle brands and their respective quantities within the dataset is presented in Table 1.

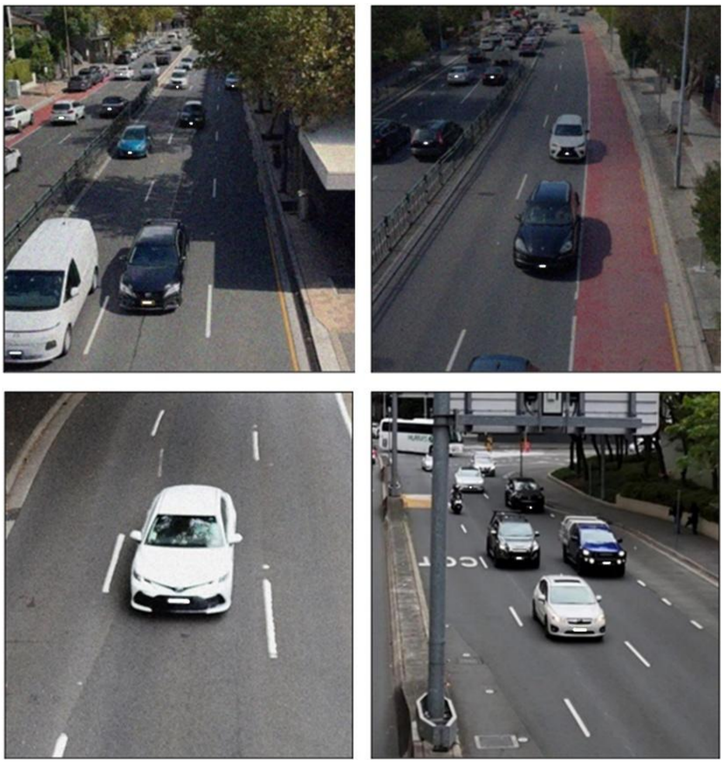

Figure 1. Sample images from the dataset

To further enhance the reliability and generalizability of the model, careful attention was given to the distribution of vehicle classes within the dataset. While certain brands, such as Toyota, trucks, and Tesla vehicles, are more prevalent due to their real-world frequency, efforts were made to mitigate potential class imbalance issues. This was achieved through targeted data augmentation techniques applied to underrepresented classes, including random rotations, brightness adjustments, and horizontal flipping, thereby enriching the diversity of samples for less common brands. Additionally, during model training, class weighting strategies were incorporated within the loss function to compensate for any residual imbalance, ensuring that the detection performance remains consistent across all vehicle categories. As a result, the dataset provides a sufficiently balanced and comprehensive representation of diverse vehicle types, supporting the development of a detection system capable of maintaining high accuracy and fairness across varying real-world traffic conditions.

Table 1. Distribution of vehicle types and brands in the dataset

Class Name	Image Count	Percentage of Total
Audi	259	2.88%
BMW	195	1.65%
Bus	366	4.95%
Ford	273	3.17%
Honda	148	0.77%
Hyundai	196	1.69%
Jeep	146	0.73%
Kia	146	0.73%
Land Rover	145	0.71%
Lexus	185	1.48%
Mazda	267	3.05%
Mercedes	234	2.42%
Motorbike	246	2.65%
Subaru	171	1.21%
Tesla	369	5.01%
Toyota	893	15.06%
Truck	512	7.75%
Volkswagen	298	3.64%
Volvo	164	1.07%

Each image was meticulously annotated using the Roboflow platform, a cloud-based tool that facilitates object detection and classification tasks. The annotation process involved manually labeling each vehicle's brand and model using bounding boxes to ensure precise localization. The dataset includes vehicles from nineteen different brands, including Toyota, Mercedes, BMW, Jeep, Land Rover, Kia, Audi, Hyundai, and Ford. These labeled images serve as the ground truth for training deep learning models, allowing them to distinguish brand-specific design elements such as emblems, grilles, headlight structures, and body contours.

To facilitate effective training and model evaluation, the dataset was divided into three subsets using a fixed partitioning ratio of 70% for training, 20% for validation, and 10% for testing. This split was applied consistently across both YOLOv8 and YOLOv11 experiments to ensure a fair and unbiased performance comparison. The majority of images assigned to the training set enabled the models to learn representative brand-specific patterns, while the validation subset was used for hyperparameter monitoring and overfitting control during training. The testing subset consisted exclusively of previously unseen images and was strictly reserved for final performance evaluation. This structured and consistent data partitioning strategy ensured reliable generalization assessment and eliminated data-related variability between model comparisons.

To enhance model performance, various data augmentation techniques were implemented to improve robustness and reduce overfitting. Adjusting brightness and contrast simulated different lighting conditions, while noise injection helped the model recognize vehicles in visually complex environments. Rotation and scaling allowed the model to detect vehicles from different perspectives, and grayscale conversion emphasized shape and texture features independent of color variations. These augmentation strategies enriched dataset diversity, ultimately improving detection accuracy in diverse traffic scenarios. Each vehicle was manually labeled with bounding boxes specifying its brand and type to ensure precise training data. The annotation process was conducted using the Roboflow platform, providing a structured approach for deep learning model training (Figure 2).

Figure 2. Example of vehicle annotation with bounding boxes and class labels

The dataset was used to train and evaluate deep learning models based on YOLOv8 and YOLOv11, two of the most advanced architectures for real-time object detection. YOLOv8 was initially tested for its balance between speed and accuracy. However, since higher detection precision was required, YOLOv11 was ultimately chosen for final implementation. The enhanced multi-scale feature extraction and refined attention mechanisms of YOLOv11 improved recognition accuracy, particularly for vehicles that were partially obscured or captured at extreme angles. The integration of Roboflow for annotation and preprocessing, alongside YOLO architectures for model training, resulted in a highly effective vehicle recognition system with high precision across diverse driving environments.

2.2. Methods

This section describes the dataset preparation, model selection, and training process for accurate and efficient vehicle brand and model recognition. The methodology ensures model robustness and generalizability by leveraging a diverse dataset and advanced object detection techniques. Labeled vehicle images were collected under various environmental conditions to improve real-world relevance. The YOLOv11 architecture was chosen for its high precision and real-time processing efficiency, making it ideal for large-scale automated recognition tasks. Data augmentation techniques were applied to improve model performance and reduce overfitting. Training was conducted using optimized hyperparameters and GPU acceleration to enhance computational efficiency. This section details the key methodological steps, including dataset collection, annotation, preprocessing, augmentation, model selection, training, and evaluation. The structured workflow ensures high accuracy and efficiency in vehicle recognition. A summary of the methodology is illustrated in Figure 3, outlining the major steps from data acquisition to model evaluation.

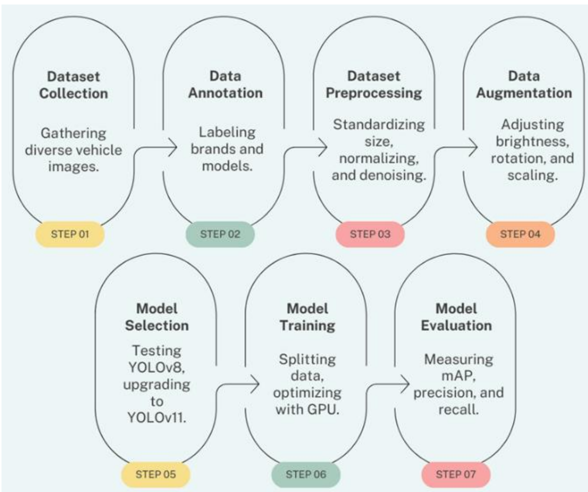

Figure 3. Workflow of the study

2.3. Deep Learning Models and Architecture

This study employed deep learning architecture specifically designed for object detection and classification. YOLOv8 and YOLOv11 were selected due to their superior speed, accuracy, and ability to process real-time traffic images with high precision. These models were chosen for their ability to efficiently detect vehicles in complex urban environments while maintaining computational efficiency. The YOLOv11 implementation used in this study was based on the official Ultralytics PyTorch framework. The model configuration corresponds to the standard YOLOv11 architecture with default backbone, neck, and detection head modules as provided in the official release. All experiments were conducted using the same architectural configuration without structural modifications to ensure reproducibility and fair comparison with YOLOv8.

YOLOv8, an advanced iteration of the You Only Look Once (YOLO) object detection framework, was initially employed for preliminary testing. Its streamlined architecture, optimized feature extraction layers, and computational efficiency made it a suitable candidate for rapid detection (Hussain, 2024; Swathi and Challa, 2024; Varghese and Sambath, 2024). However, YOLOv8 demonstrated certain limitations in detecting vehicles that were partially occluded, captured in extreme lighting conditions, or positioned at difficult angles. YOLOv8 employs a lightweight CSP-derived backbone with optimized convolutional blocks and an anchor-free detection head. Its architecture prioritizes computational efficiency while maintaining competitive detection accuracy. However, compared to YOLOv11, it lacks the extended transformer-based contextual modeling components that enhance fine-grained feature discrimination.

To address these challenges, YOLOv11 was selected as the primary detection model. This latest enhancement in the YOLO series introduced improved multi-scale feature extraction, refined bounding box regression algorithms, and transformer-based attention mechanisms. These modifications enabled YOLOv11 to achieve superior detection accuracy, particularly in challenging scenarios where vehicles appeared in cluttered backgrounds or under suboptimal lighting conditions (Bitwire and Han, 2024; Lee and Kim, 2024; Jiang and Zhong, 2025). The YOLOv11 architecture consists of three primary components: a CSP-based backbone for hierarchical feature extraction, a PAN-FPN neck for multi-scale feature aggregation, and an anchor-free detection head for final object localization and classification. The backbone integrates enhanced C2f modules to improve gradient flow and feature reuse. In addition, transformer-based attention blocks are embedded within deeper layers to strengthen global contextual reasoning. These attention modules enable improved discrimination of fine-grained brand-specific visual patterns such as logos and grille structures. The structural architecture of the YOLOv11 model employed in this study is illustrated in Figure 4. As shown in the diagram, the processing pipeline begins with input image acquisition, followed by hierarchical feature extraction through a CSP-based backbone network. The extracted features are then forwarded to a PAN-FPN neck module, which performs multi-scale feature aggregation to enhance detection performance across different object sizes. Subsequently, transformer-enhanced contextual refinement blocks are applied within deeper feature layers to strengthen global feature representation and improve discrimination of fine-grained brand-specific characteristics. Finally, the refined feature maps are processed by an anchor-free detection head, which performs bounding box regression and multi-class classification simultaneously. This structured pipeline ensures both spatial localization accuracy and semantic classification robustness, particularly in complex real-world traffic environments involving occlusion, varying illumination, and scale diversity.

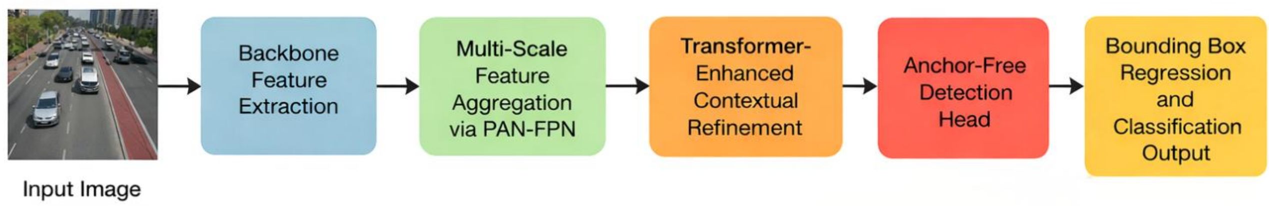

Figure 4. Structural flow of the YOLOv11 architecture used in this study

Both models were trained using preprocessed and annotated data from the Roboflow platform. The training process involved optimizing loss functions and fine-tuning hyperparameters to maximize detection accuracy while minimizing false positives (Alin and Yuana, 2023; Shandilya et al., 2023). YOLOv11 demonstrated a significant improvement over YOLOv8 in recognizing fine-grained vehicle features, making it the preferred model for this study. Its ability to process images in real time while maintaining high precision made it well-suited for large-scale deployment in smart traffic monitoring systems and commercial vehicle market analysis.

2.4. Data Processing and Training Strategy

To ensure high detection accuracy and computational efficiency, a structured data processing and training strategy was implemented. Several preprocessing steps were applied before feeding the data into the deep learning models to enhance feature extraction and improve generalization. Image resizing was performed to standardize input dimensions, maintaining consistent proportions for efficient processing. Normalization techniques scaled pixel values within a fixed range, stabilizing training dynamics and improving convergence speed. Additionally, noise reduction was applied to filter out low-quality images, improving dataset reliability and reducing the risk of misclassification.

Data augmentation played a crucial role in enhancing model robustness (Kaur et al., 2021; Montserrat et al., 2017). Various transformations, such as random cropping, flipping, rotation, and color jittering, were applied to introduce variations that helped the model generalize better. These augmentations ensured that the model could detect vehicles under different lighting conditions, backgrounds, and perspectives, ultimately improving performance in real-world scenarios. Data augmentation parameters were applied using the Roboflow preprocessing configuration, including controlled rotation, scaling, and photometric transformations with fixed parameter ranges.

The training process was conducted in a high-performance GPU computing environment using a high-performance GPU computing environment based on the PyTorch framework. YOLOv8 and YOLOv11 were trained with a stochastic gradient descent (SGD) optimizer, incorporating momentum adjustments for stable learning. The mAP metric was used to evaluate model performance at different training stages, ensuring continuous improvements in vehicle brand and model detection accuracy. The training process for both YOLOv8 and YOLOv11 was conducted under identical computational and experimental conditions in a high-performance GPU computing environment using the same deep learning framework configuration. The dataset splitting strategy was kept consistent for both models (70% training, 20% validation, 10% testing), ensuring that performance differences were not influenced by data partitioning variations. Both architectures were trained for the same number of epochs (300), using the same batch size (32), learning rate (0.001), and loss function configuration.

Hyperparameter tuning was a key step in optimizing model training. The batch size was set to 32, balancing training stability and computational efficiency. The learning rate was fixed at 0.001, allowing for smooth gradient updates without excessive fluctuations. The epoch count was determined as 300, ensuring sufficient iterations for convergence. The Complete IoU (CIoU) loss function was used to optimize bounding box regression, improving localization accuracy and object tracking. The detailed hyperparameter configuration for YOLOv11 is presented in Table 2. These hyperparameters were determined based on preliminary experiments and are not the result of a systematic hyperparameter optimization. Although YOLOv11 incorporates additional architectural refinements, the observed reduction in training time is attributed to internal optimization improvements within its implementation, including more efficient memory utilization and computational graph execution. Therefore, the difference in training duration does not indicate inconsistency in hardware or training configuration, but reflects backend efficiency differences between the model versions.

Table 2. YOLOv11 training configuration parameters

Training Parameter	**Value**
Batch Size	32
Learning Rate	0.001
Number of Epochs	300
Loss Function	Complete Intersection over Union (CIoU)

The effectiveness of the trained YOLOv8 and YOLOv11 architectures was evaluated using precision, recall, Intersection over Union (IoU), and mAP. While YOLOv8 demonstrated promising results, YOLOv11 exhibited superior accuracy, improved object localization, and better performance in detecting vehicles under complex conditions. These findings positioned YOLOv11 as the preferred model for this study due to its reliability in large-scale vehicle detection applications. The combination of YOLOv11 and an optimized training strategy resulted in a high-performance vehicle recognition system capable of real-time, high-precision detection in diverse traffic environments.

3. Results

This section presents an in-depth evaluation of the proposed deep learning-based vehicle brand and model recognition system. In addition to presenting quantitative performance metrics, it is crucial to provide a comprehensive interpretation of the experimental results. Therefore, the following subsections not only report the numerical outcomes but also offer detailed discussions of the associated figures and tables. This interpretative analysis aims to contextualize the results, highlight the comparative strengths and weaknesses of the evaluated models, and elucidate the underlying factors influencing the observed performance trends. The findings include an analysis of accuracy metrics, a comparison between YOLOv8 and YOLOv11 architectures, and an assessment of computational efficiency. Additionally, performance-related visualizations, including loss curves, confusion matrices, and feature space representations, provide a comprehensive overview of the models' capabilities. The dataset used for training and evaluation consists of 5,213 labeled vehicle images, featuring various automobile brands and vehicle types. The model's performance was assessed using mAP, precision, recall, and intersection over union (IoU), which are critical metrics in object detection tasks. These metrics provide insights into the model's ability to detect and classify vehicles accurately while minimizing misclassification errors.
The comparison between YOLOv8 and YOLOv11 demonstrates that YOLOv11 consistently outperforms YOLOv8 in accuracy and precision. The mAP obtained by YOLOv11 was 88.6%, surpassing the 87.4% achieved by YOLOv8 (Figure 5). As depicted in Figure 5, the superior mAP achieved by YOLOv11 suggests that the model is more capable of extracting discriminative features and accurately localizing objects across varying traffic conditions compared to its predecessor. This improvement indicates that YOLOv11 has a stronger capability in detecting and classifying vehicle brands and models while reducing false detections. In terms of precision, YOLOv11 recorded a precision score of 90.7%, significantly higher than YOLOv8's 83.7%. This means that YOLOv11 produces fewer false positives, making it a more reliable model for real-world applications. However, YOLOv8 exhibited a higher recall value of 86.4%, compared to YOLOv11's 81.5%. This suggests that YOLOv8 detects a greater number of vehicles albeit with an increased rate of false classifications.

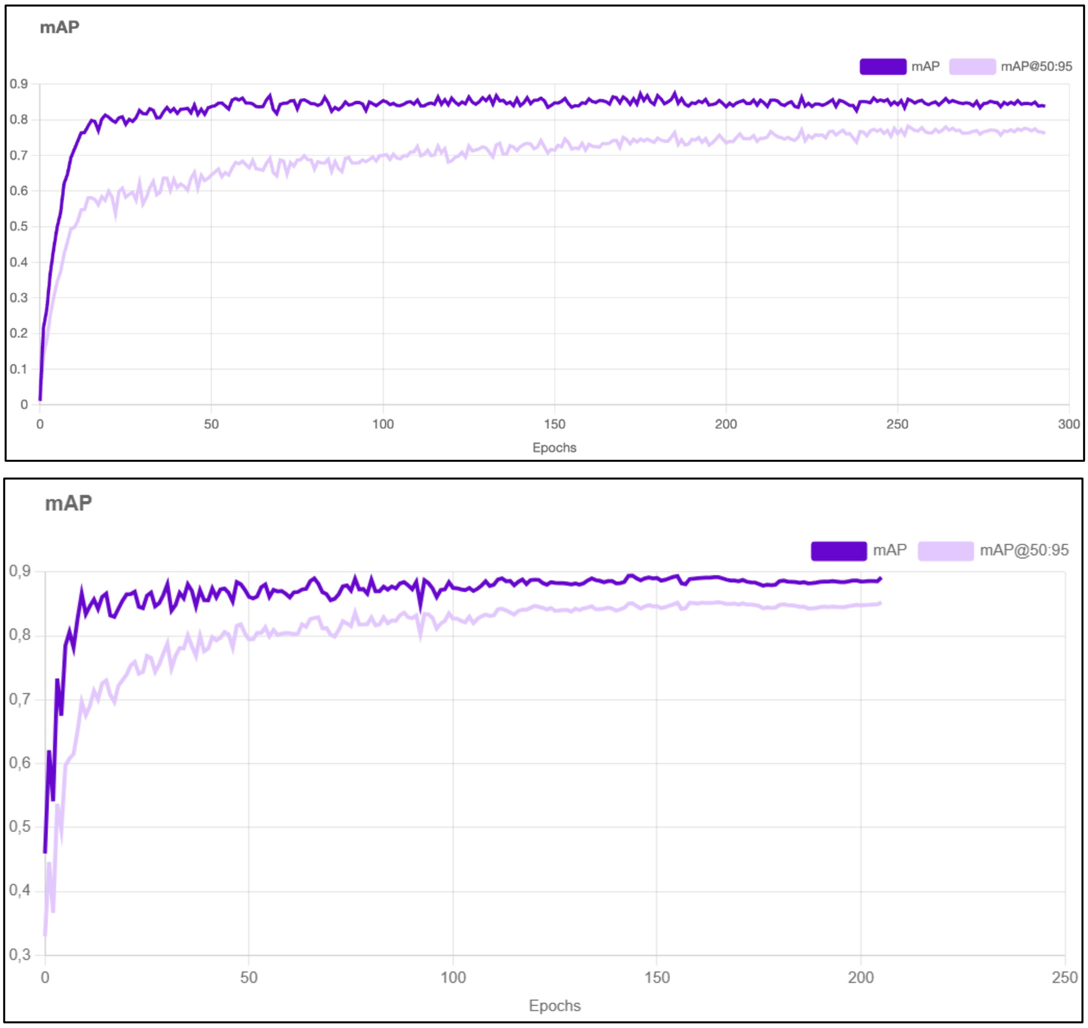

Figure 5. mAP comparison of YOLOv8 (top) and YOLOv11 (bottom)

A summary of the mAP, precision, and recall scores is provided in Table 3, which confirms the better detection accuracy of YOLOv11 compared to YOLOv8. These results indicate that YOLOv11 is the preferred choice for applications requiring high precision and reliability, whereas YOLOv8 may be more suitable for scenarios where a broader detection range is prioritized.

Table 3. Model performance metrics

Model	mAP (%)	Precision (%)	Recall (%)
YOLOv11	88.6	90.7	81.5
YOLOv8	87.4	83.7	86.4

The performance metrics presented in Table 3 further corroborate the advantage of YOLOv11, particularly in achieving a higher precision rate. This indicates that YOLOv11 produces fewer false positives, which is critical for reliable real-world deployment where classification errors can have significant practical consequences. Beyond detection accuracy, the reliability of brand distribution estimation depends directly on precision stability. Higher false positive rates may artificially inflate brand density measurements in market-oriented interpretations. Therefore, the superior precision of YOLOv11 (90.7%) ensures more statistically consistent brand frequency estimation, reducing distribution bias in traffic-based market analytics. From a market analysis perspective, precision becomes more critical than recall, as over-detection may distort brand prevalence ratios.

Loss function analysis provides crucial insights into the optimization of deep learning models. The total loss curves of YOLOv8 and YOLOv11 are shown in Figures 6 and 7, respectively. Figures 6 and 7 illustrate that YOLOv11 not only converges faster during the training process but also achieves lower final loss values, indicating a more efficient learning mechanism. The smoother loss curve associated with YOLOv11 also reflects better stability and reduced overfitting risk compared to YOLOv8. YOLOv11 demonstrates faster convergence and lower total loss values, suggesting superior learning efficiency and generalization capability. The classification loss curve (Figure 7) further validates YOLOv11's improved ability to distinguish between different vehicle brands and models, leading to reduced misclassification rates.

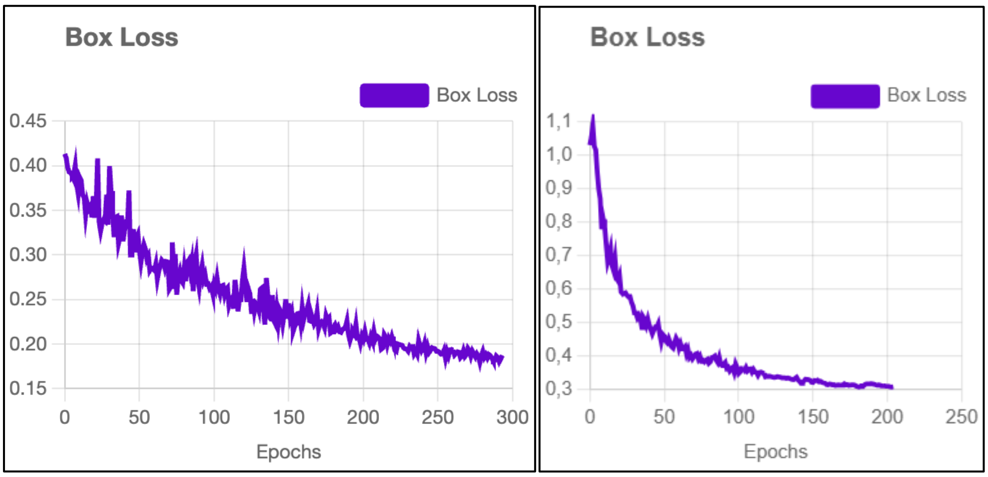

Figure 6. Training loss curve of YOLOv8 (left) and YOLOv11 (right)

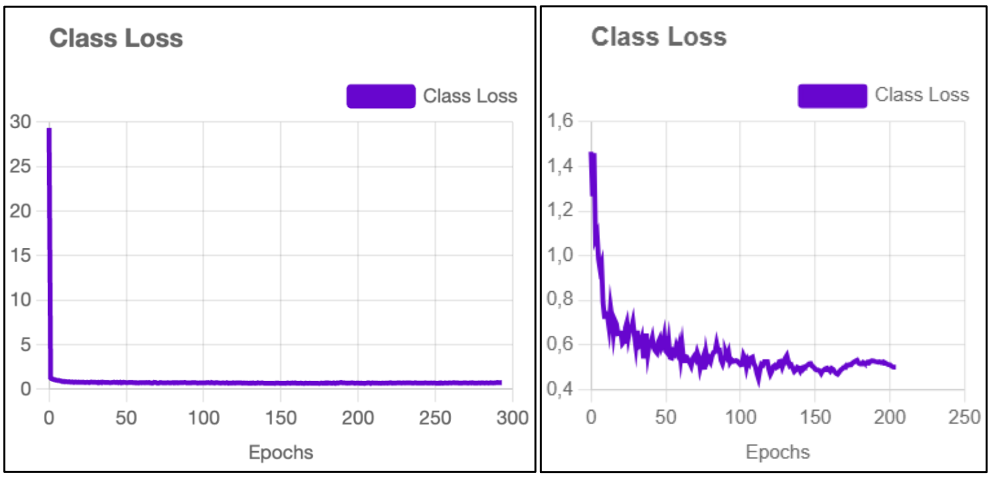

Figure 7. Classification loss comparison for YOLOv8 (left) and YOLOv11 (right)

Similarly, the object loss analysis (Figure 8) confirms that YOLOv11 achieves better spatial recognition, accurately localizing vehicles even under challenging conditions such as partial occlusions, extreme angles, and varying lighting environments. The lower object loss observed for YOLOv11 suggests improved localization capabilities, enabling the model to more accurately delineate vehicle boundaries even under challenging environmental conditions.

A confusion matrix was generated for both YOLOv8 and YOLOv11 to visualize their classification accuracy. The confusion matrix for YOLOv8 (Figure 9) shows a higher number of misclassifications, especially between visually similar vehicle brands. The confusion matrix of YOLOv11 (Figure 10) highlights a significant improvement, with most vehicle brands being classified correctly and fewer overlaps observed between visually similar categories. This improved discrimination capability can be attributed to YOLOv11's enhanced feature extraction and attention mechanisms. In contrast, the confusion matrix for YOLOv11 (Figure 10) reveals a significant reduction in misclassification errors, indicating a more refined capability to differentiate between brands with overlapping visual characteristics.

A feature vector space analysis was also conducted to assess the model's ability to distinguish between different vehicle brands. As seen in Figures 11 and 12, YOLOv11 exhibits more compact and clearly separated feature clusters, underscoring its superior capability in learning brand-specific visual patterns. In contrast, the more scattered distributions observed in YOLOv8 reflect higher inter-class confusion and reduced representation power. The vector space representation of YOLOv8 (Figure 11) shows a wider distribution with overlapping clusters, leading to misclassifications. In contrast, the vector space representation of YOLOv11 (Figure 12) exhibits more compact and well-separated clusters, demonstrating superior brand-specific feature extraction.

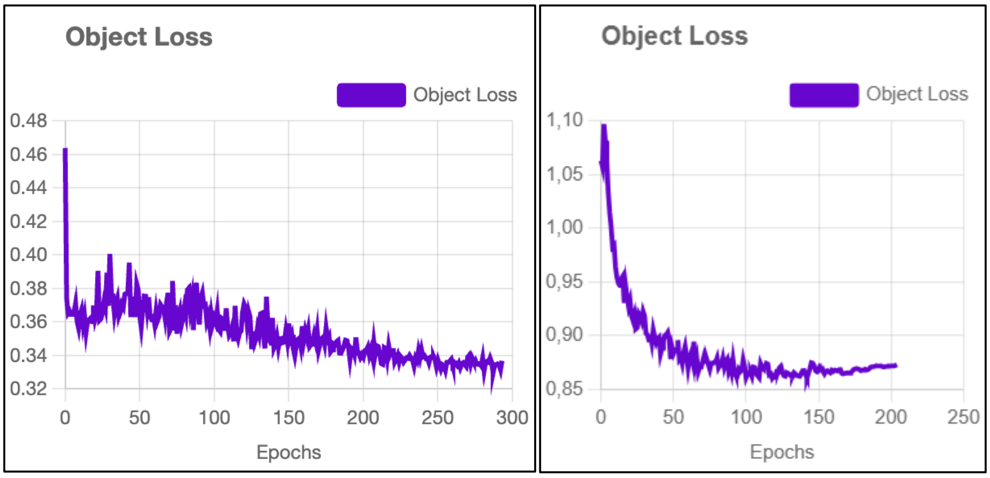

Figure 8. Object loss curve for YOLOv8 (left) and YOLOv11 (right)

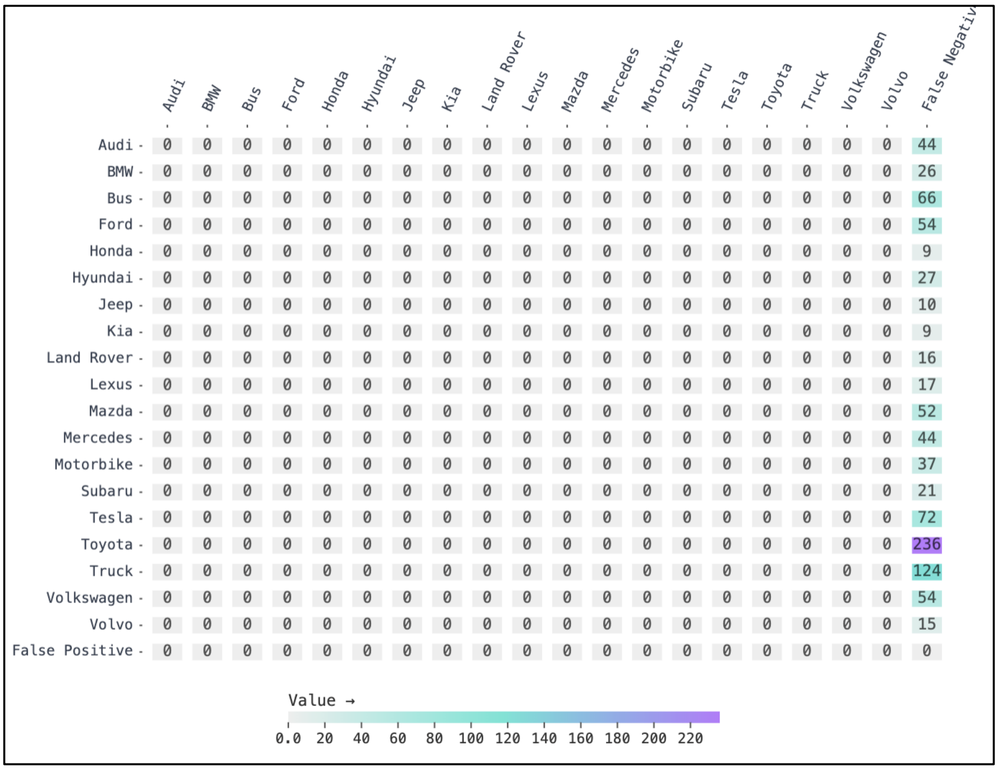

Figure 9. Confusion matrix for YOLOv8

Besides classification accuracy, the efficiency of training and inference times is critical for real-world applications. The substantially reduced training time for YOLOv11 (34 minutes vs. 1,426 minutes for YOLOv8) demonstrates the computational advantages of the updated architecture. Although YOLOv8 exhibited slightly faster CPU-based execution, the real-time inference capabilities of YOLOv11 offer a better trade-off between accuracy and processing efficiency for deployment in intelligent traffic systems. A comparative summary of training time, inference response time, and CPU-based execution speed is provided in Table 4.

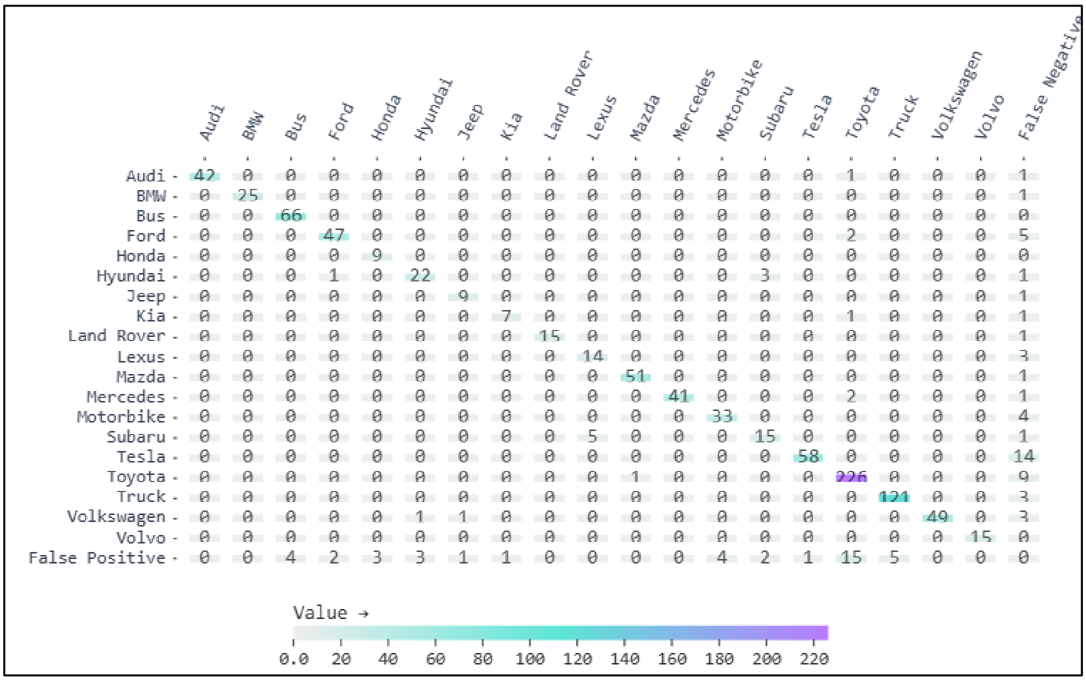

Figure 10. Confusion matrix for YOLOv11

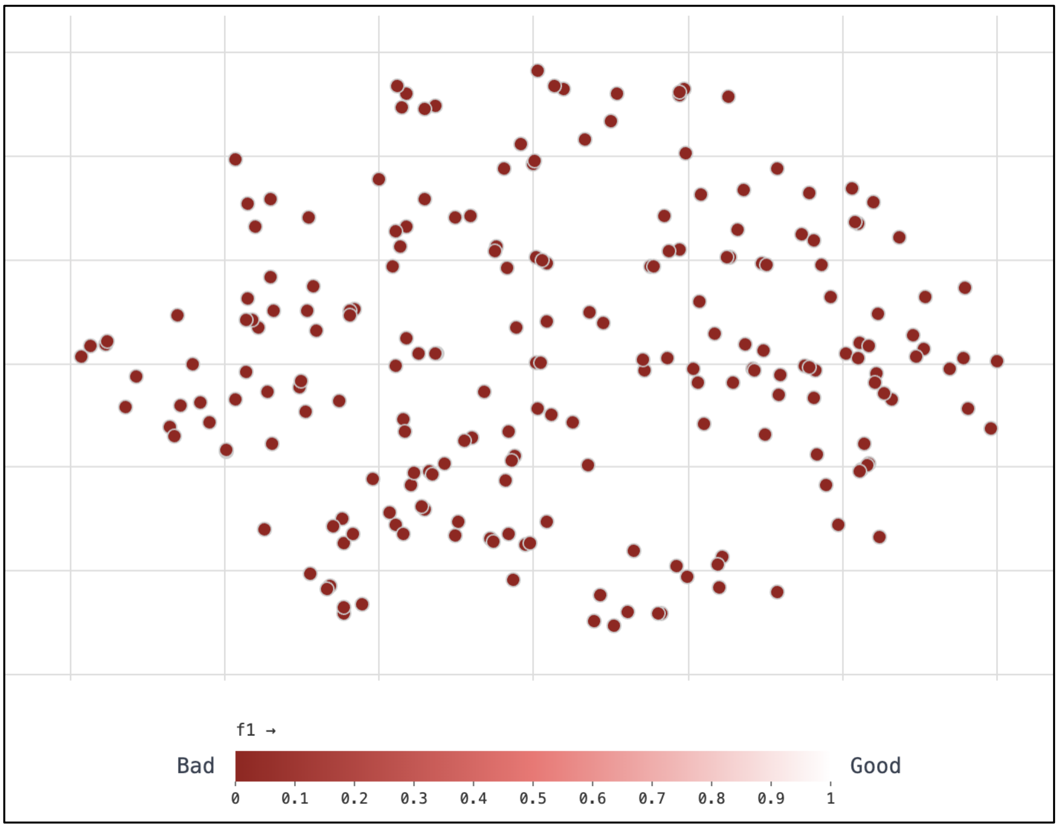

Figure 11. YOLOv8 feature vector space representation

These findings highlight that YOLOv11 is significantly more efficient in training, requiring only 34 minutes, compared to the 1,426 minutes needed by YOLOv8. The lower inference response time of YOLOv11 (37.3 ms) further supports its real-time applicability, although YOLOv8 demonstrated faster CPU-based execution due to its less complex architecture.

Overall, the detailed analysis of figures and tables provides a holistic understanding of the comparative performance of YOLOv8 and YOLOv11, offering valuable insights into the practical viability of deep learning-based vehicle brand and model recognition systems.

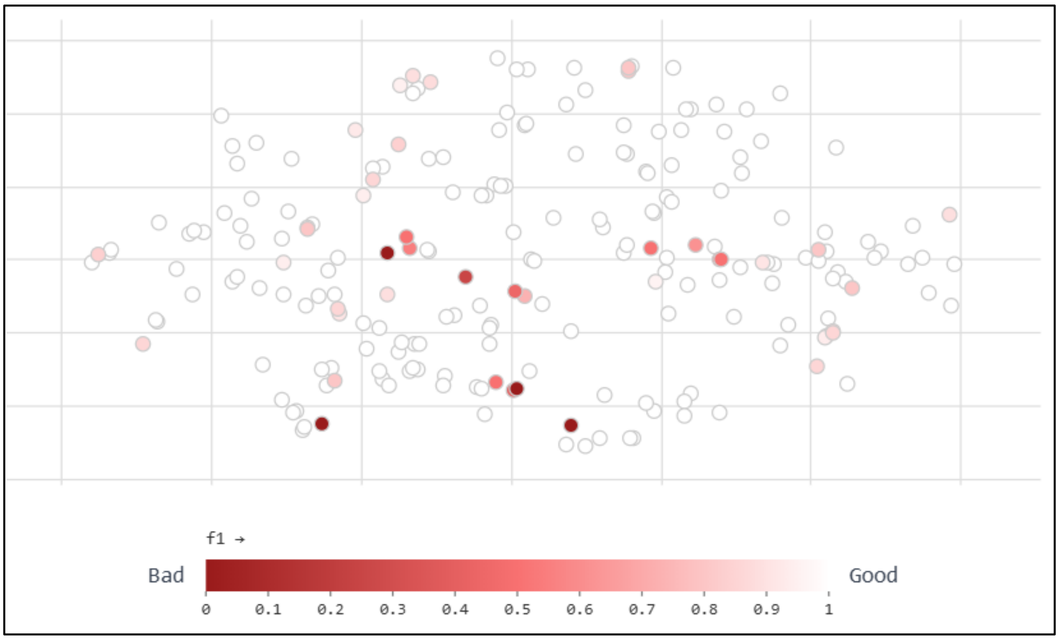

Figure 12. YOLOv11 feature vector space representation

Table 4. Model training and inference speed metrics

Model	Training Time (minutes)	Response Time (ms)	CPU ONNX Execution Speed (ms)
YOLOv11	34	37.3	234.7
YOLOv8	1426	39.3	183.2

4. Discussion

The field of vehicle brand and model recognition has seen significant advancements with the integration of deep learning algorithms, particularly the YOLO (You Only Look Once) series. This discussion evaluates our study's findings in the context of existing research, highlighting the evolution of YOLO architectures and their applications in vehicle detection and recognition tasks.

The YOLO series has undergone substantial enhancements since its inception, with each version introducing architectural improvements to enhance detection accuracy and speed. Our study's implementation of YOLOv11 is consistent with this trend, leveraging its advanced multi-scale feature extraction and attention mechanisms to achieve a mAP of 88.6% across all vehicle brands. This performance underscores YOLOv11's capability to effectively handle complex detection scenarios, including small or partially occluded vehicles.

To contextualize our results, we compared our approach with ten similar studies that utilized various YOLO versions and other deep learning models for vehicle detection and recognition. The table below summarizes the key aspects and outcomes of these studies given by Table 5. To provide a more complete comparison, the results of the current study are also included in Table 5. The proposed YOLOv11-based vehicle recognition system demonstrated superior performance with an 88.6% mAP and 90.7% precision, while maintaining an efficient training time of only 34 minutes. These results further validate the robustness and practical applicability of the proposed approach in real-world traffic conditions.

Table 5. Comparison of YOLO-based vehicle detection studies

Study	Dataset	Key Findings	Reference
Adapted YOLOv4 for 3D object detection in autonomous vehicles	Custom dataset	Enhanced 3D perception, improved safety in AVs	(Murendeni et al., 2024)
Improved YOLOv4 for vehicle logo detection	Custom dataset	Increased accuracy in complex backgrounds	(Jiang et al., 2022)
Improved YOLOv5 for real-time vehicle detection	Custom dataset	Reduced false detection rates in occluded scenarios	(Zhang et al., 2022)
YOLOv11 for vehicle detection	Comprehensive dataset	Enhanced detection of small and occluded vehicles	(Alif, 2024)
Vehicle and license plate recognition with YOLOv4	Diverse Vehicle and License Plates Dataset (DVLPD)	Achieved high mAP in vehicle type recognition and license plate reading	(Usama et al., 2025)
Comparison of YOLOv3, YOLOv4, YOLOv5l, YOLOv5s, and SSD MobileNetv2 FPN-lite for street-level object detection	Modified Udacity Self Driving Car Dataset (3,169 images)	YOLOv5l achieved the highest mAP@0.5 of 0.593; MobileNetv2 FPN-lite demonstrated the fastest inference time at 3.20ms; YOLOv5s offered a balance between accuracy and speed	(Naftali et al., 2022)
Performance analysis of YOLO-based architectures for vehicle detection in traffic images from Bangladesh	Custom dataset with 7,390 images	YOLOv5x outperformed YOLOv3 and YOLOv5s, achieving 7% and 4% higher mAP, respectively	(Alamgir et al., 2022)
Object detection in dense traffic using YOLOv5 and Non-Maximum Suppression Ensembling	Dhaka AI dataset	Achieved mAP@0.5 of 0.458 and an inference time of 0.75 seconds, outperforming other models	(Rahman et al., 2022)
This Study (YOLOv11)	Real-world traffic dataset (5,213 images)	Achieved 88.6% mAP, 90.7% precision; superior performance in real-world vehicle brand recognition with efficient training time (34 min)	This Study

The study's findings align with advancements in YOLO architectures, particularly with YOLOv11, which has demonstrated superior performance in detecting small and occluded vehicles, reinforcing our results. Implementation of YOLOv11 for vehicle brand and model recognition corresponds with the continuous enhancements in YOLO-based object detection, proving its efficacy in handling complex detection tasks. The comparative analysis with existing literature confirms that advanced YOLO architectures significantly improve detection accuracy and robustness. Future research should aim to refine these models further while incorporating more diverse datasets to enhance the adaptability and reliability of vehicle recognition systems. While the architectural superiority of newer YOLO versions may be expected in controlled benchmarks, the critical contribution of this study lies in demonstrating how these improvements affect downstream analytical reliability in real-world traffic-derived commercial indicators. The analysis shows that architectural refinements directly reduce distribution noise, thereby increasing the statistical robustness of passive market observation systems.

Compared to the studies reviewed, the proposed YOLOv11-based vehicle brand and model recognition system demonstrates several distinct advantages. Firstly, it achieves a higher mAP of 88.6%, surpassing the detection accuracies reported in most existing works employing earlier YOLO versions. Secondly, the model maintains an exceptionally efficient training process, completing full convergence in only 34 minutes under the experimental conditions of this study. Furthermore, the system's real-time inference capability, with an average response time of 37.3 milliseconds, ensures practical applicability in intelligent transportation systems requiring rapid decision-making. These improvements collectively highlight the superior balance between accuracy, efficiency, and deployment feasibility achieved by the proposed method, setting it apart from prior research efforts.

5. Conclusion

This study has demonstrated the effectiveness of deep learning and computer vision techniques in achieving automated vehicle brand and model recognition, contributing significantly to advancements in intelligent transportation systems and commercial market analysis. The proposed system, developed using the YOLOv11 architecture, exhibited high detection accuracy and real-time processing capabilities. Trained on a dataset comprising 5,213 labeled vehicle images under diverse environmental conditions, the model achieved a mAP of 88.6% and a precision of 90.7%, affirming its robustness and generalizability in real-world traffic scenarios. The novelty of this work lies not in proposing a new detection architecture, but in analytically linking detection precision to traffic-derived brand density reliability, thereby transforming object detection outputs into statistically interpretable commercial indicators.
Comparative analyses indicated that YOLOv11 outperforms its predecessor YOLOv8, particularly in precision and localization accuracy, while maintaining efficient training and inference times suitable for large-scale deployment. Furthermore, the system’s real-time inference capability, with an average response time of 37.3 milliseconds, underscores its practical applicability in smart city infrastructures and intelligent surveillance systems.
Despite these promising results, the study is not without limitations. One of the primary constraints is the dependency on camera resolution and image quality; lower resolution footage may hinder the model’s ability to accurately identify brand-specific details such as emblems and grille designs. Additionally, the recognition performance tends to decrease as the distance between the camera and the vehicle increases, particularly for small-sized vehicles or vehicles captured at oblique angles. Another limitation arises from environmental conditions such as heavy rain, fog, or extreme lighting, which were underrepresented in the dataset and may adversely impact the detection accuracy in real-world deployments.

Future work should aim to address these limitations by incorporating higher-resolution imagery, diverse camera angles, and weather variations into the dataset. Moreover, integrating multi-sensor data, such as LiDAR or radar information, alongside visual inputs, could further enhance the model’s performance under challenging conditions. Advancements in 3D object detection and the adoption of transformer-based architectures may also offer opportunities to improve the system’s robustness and fine-grained recognition capabilities.

In conclusion, this study presents a high-performance, real-time vehicle brand and model recognition system that not only meets current technological demands but also lays a solid foundation for future enhancements. By acknowledging existing limitations and outlining clear directions for improvement, the research underscores the evolving potential of AI-driven vehicle recognition technologies in shaping the future of intelligent mobility systems.

Acknowledgement

This study is derived from the Master’s Thesis named "Brand Model Detection of Automobiles used on Highways with Image Processing Techniques and Related commercial Market Research" in the Department of Mechatronics Engineering, Institute of Graduate Education, Isparta University of Applied Sciences.

Conflict of Interest

No conflict of interest was declared by the authors.

References

Adu-Gyamfi, Y.O., Asare, S.K., Sharma, A., Titus, T., 2017. Automated vehicle recognition with deep convolutional neural networks. Transportation Research Record, 2645 (1), 113-122.

Alamgir, R.M., Shuvro, A.A., Al Mushabbir, M., Raiyan, M.A., Rani, N.J., Rahman, M.M., Kabir, M.H., Ahmed, S., 2022. Performance analysis of YOLO-based architectures for vehicle detection from traffic images in Bangladesh. 25th International Conference on Computer and Information Technology (ICCIT), IEEE, 982-987.

Alif, M.A.R., 2024. YOLOv11 for vehicle detection: Advancements, performance, and applications in intelligent transportation systems. arXiv preprint arXiv:2410.22898.

Alin, A.Y., Yuana, K.A., 2023. Data augmentation method on drone object detection with YOLOv5 algorithm. Eighth International Conference on Informatics and Computing (ICIC), IEEE, 1-6.

Alruwaili, M., Atta, M.N., Siddiqi, M.H., Khan, A., Khan, A., Alhwaiti, Y., Alanazi, S., 2023. Deep learning-based YOLO models for the detection of people with disabilities. IEEE Access, 12, 2543-2566.

Atkočiūnas, E., Blake, R., Juozapavičius, A., Kazimianec, M., 2005. Image processing in road traffic analysis. Nonlinear Analysis: Modelling and Control, 10 (4), 315-332.

Bitwire, G.A., Han, D.S., 2024. YOLOv11: Revolutionizing object detection with focus on tiny objects in complex settings. Korean Institute of Communication Sciences Fall Conference Proceedings.

Chen, Y.-H., Kara, L.B., Cagan, J., 2023. Automating style analysis and visualization with explainable AI-case studies on brand recognition. International Design Engineering Technical Conferences and Computers and Information in Engineering Conference, American Society of Mechanical Engineers, V03AT03A006.

Fathy, M., Siyal, M., 1995. Real-time image processing approach to measure traffic queue parameters. IEEE Proceedings-Vision, Image and Signal Processing, 142 (5), 297-303.

Gilmore, J.F., Elibiary, K.J., 1993. AI in advanced traffic management systems. Association for the Advancement of Artificial Intelligence (AAAI) Technical Report, WS-93-04.

Gulati, R., Srinivasan, R., 2019. Image processing in intelligent traffic management. International Journal of Recent Technology and Engineering, 8 (2S4), 213-218.

Hidayatullah, P., Syakrani, N., Sholahuddin, M.R., Gelar, T., Tubagus, R., 2025. YOLOv8 to YOLO11: A comprehensive architecture in-depth comparative review. arXiv preprint arXiv:2501.13400.

Hu, C., Bai, X., Qi, L., Wang, X., Xue, G., Mei, L., 2015. Learning discriminative pattern for real-time car brand recognition. IEEE Transactions on Intelligent Transportation Systems, 16 (6), 3170-3181.

Hussain, M., 2024. YOLOv1 to v8: Unveiling each variant-a comprehensive review of YOLO. IEEE Access, 12, 42816-42833.

Jegham, N., Koh, C.Y., Abdelatti, M., Hendawi, A., 2024. Evaluating the evolution of YOLO (you only look once) models: A comprehensive benchmark study of YOLO11 and its predecessors. arXiv preprint arXiv:2411.00201.

Jiang, T., Zhong, Y., 2025. ODVerse33: Is the new YOLO version always better? A multi-domain benchmark from YOLO v5 to v11. arXiv preprint arXiv:2502.14314.

Jiang, X., Sun, K., Ma, L., Qu, Z., Ren, C., 2022. Vehicle logo detection method based on improved YOLOv4. Electronics, 11 (20), 3400.

Kaur, P., Khehra, B.S., Mavi, E.B.S., 2021. Data augmentation for object detection: A review. IEEE International Midwest Symposium on Circuits and Systems (MWSCAS), IEEE, 537-543.

Lee, Y.-H., Kim, H.-J., 2024. Comparative analysis of YOLO series (from V1 to V11) and their application in computer vision. Journal of the Semiconductor & Display Technology, 23 (4), 190-198.

Lu, P., Wu, F., Hsiao, S.-W., Tang, J., 2025. A feature curve-based method for balancing brand identity and emotional imagery in automobile frontal form design. Advanced Engineering Informatics, 65.

Ma, X., Boukerche, A., 2020. An AI-based visual attention model for vehicle make and model recognition. IEEE Symposium on Computers and Communications (ISCC), IEEE, 1-6.

Ma, Y., Chowdhury, M., Sadek, A., Jeihani, M., 2009. Real-time highway traffic condition assessment framework using vehicle-infrastructure integration (VII) with artificial intelligence (AI). IEEE Transactions on Intelligent Transportation Systems, 10 (4), 615-627.

Montserrat, D.M., Lin, Q., Allebach, J., Delp, E.J., 2017. Training object detection and recognition CNN models using data augmentation. Electronic Imaging, 29, 27-36.

Murendeni, R., Mwanza, A., Obagbuwa, I.C., 2024. Using a YOLO deep learning algorithm to improve the accuracy of 3D object detection by autonomous vehicles. World Electric Vehicle Journal, 16 (1), 9.

Naftali, M.G., Sulistyawan, J.S., Julian, K., 2022. Comparison of object detection algorithms for street-level objects. arXiv preprint arXiv:2208.11315.

Rahman, R., Bin Azad, Z., Hasan, M.B., 2022. Densely-populated traffic detection using YOLOv5 and non-maximum suppression ensembling. Proceedings of the International Conference on Big Data, IoT, and Machine Learning: BIM 2021, Springer, 567-578.

Shandilya, S.K., Srivastav, A., Yemets, K., Datta, A., Nagar, A.K., 2023. YOLO-based segmented dataset for drone vs. bird detection for deep and machine learning algorithms. Data in Brief, 50, 109355.

Swathi, Y., Challa, M., 2024. YOLOv8: Advancements and innovations in object detection. International Conference on Smart Computing and Communication, Springer, 1-13.

Usama, M., Anwar, H., Anwar, S., 2025. Vehicle and license plate recognition with novel dataset for toll collection. Pattern Analysis and Applications, 28 (2), 57.

Varghese, R., Sambath, M., 2024. YOLOv8: A novel object detection algorithm with enhanced performance and robustness. 2024 International Conference on Advances in Data Engineering and Intelligent Computing Systems (ADICS), IEEE, 1-6.
Yang, B., 2023. Research on vehicle detection and recognition technology based on artificial intelligence. Microprocessors and Microsystems, 104937.
Zhang, Y., Guo, Z., Wu, J., Tian, Y., Tang, H., Guo, X., 2022. Real-time vehicle detection based on improved YOLO v5. Sustainability, 14 (19), 12274.