Enhancing YOLOv8 for Vehicle Detection in Intelligent Traffic Management

Abe Degale Desta[1,2] and Cheng Jian[1]

[1]*PhD Student in Department of Information and Communication Engineering, University of Electronic Science and Technology, Chengdu, China*
[2]*Director, ICT Development Directorate, Wachemo University, Hossana, P.O.Box: 667, Ethiopia*
[3]*Professor in Department of Information and Communication Engineering, University of Electronic Science and Technology, Chengdu, China*
Email: degele@wcu.edu.et

Abstract: Object detection has witnessed significant advancements with the introduction of deep learning models. In this study, we evaluated YOLOv8 model on the BDD100K dataset over 25 epochs, focusing on its progression in precision, recall, mean Average Precision (mAP), and loss metrics. The YOLOv8 model was trained and evaluated on the BDD100K dataset. Data preprocessing was done by resizing images and applying augmentation techniques like scaling, flipping, and cropping to enhance generalization. The model utilized an anchor-free detection head and an optimized backbone network for improved performance. Training was conducted over 25 epochs using the stochastic gradient descent optimizer, with non-maximum suppression and confidence thresholding applied during post-processing to refine detections. Performance was assessed using precision, recall, mAP, and loss metrics. Early epochs were marked by high classification and box losses, which reduced significantly by epoch 5. Precision and recall improved consistently, with precision increasing from 0.81 to 0.74 and recall stabilizing around 0.72–0.78. Middle epochs (6–20) showed stabilization in losses, with box loss reducing to 0.7–0.8 and classification loss to 0.4–0.5. Performance metrics peaked during this phase, achieving a precision of 0.90, recall of 0.84, and mAP@50 and mAP@50-95 values of 0.90 and 0.74, respectively. During the final epochs (21–25), the model achieved optimal stability, with box loss around 0.69 and classification loss at 0.37. Precision remained at 0.90, recall at 0.85, and mAP metrics stabilized at 0.91 and 0.74. A comparative analysis underscored YOLOv8's superiority over YOLOv5, achieving higher F1-score (0.87 vs. 0.67), mAP@50 (0.91 vs. 0.682), and mAP@50-95 (0.74 vs. 0.449). YOLOv8's architectural advancements like anchor-free detection head and optimized backbone improved object detection. YOLOv8's balance of precision and recall, supported by strong post-processing, affirms its robustness for real-world applications.

Keywords: YOLOv8, object detection, BDD100K, mAP metrics, deep learning models

1. Introduction

In the ever-evolving domain of intelligent transportation systems, accurate and efficient vehicle detection plays an important role in addressing challenges associated with traffic congestion, road safety, and urban mobility [1]. The expansion of mobility sector globally has necessitated the development of advanced systems capable of monitoring and managing traffic flow dynamically [2]. Traditional traffic management systems often rely on fixed infrastructure, such as inductive loop detectors or traffic cameras, that offer limited flexibility and scalability [3]. Recent advancements in computer vision and deep learning have ushered in transformative possibilities, enabling more robust and adaptive solutions for traffic monitoring and management. Among these advancements, object detection models such as

You Only Look Once (YOLO) have emerged as leading frameworks due to their speed and accuracy [4].

The YOLO family of models has garnered significant attention for its ability to perform real-time object detection, with applications extending across domains such as autonomous driving, surveillance, and intelligent traffic management. YOLOv8 introduces enhancements in architecture and training strategies, offering improved performance over its predecessors [5,6]. However, deploying such models for vehicle detection in real-world traffic scenarios requires addressing various challenges, including occlusions, variations in lighting conditions, and the diversity of vehicle types.

Datasets play a critical role in the training and evaluation of deep learning models for vehicle detection. The selection of an appropriate dataset directly impacts the accuracy, robustness, and generalizability of the model. In this study, we utilize a customized version of the Berkeley Deep Drive (BDD100K) dataset [7]. BDD100K is one of the largest and most diverse datasets available for autonomous driving research, comprising images captured under varying weather conditions, times of day, and geographic locations [8]. By curating and customizing this dataset to align with the specific requirements of urban traffic management, we aim to enhance the model's ability to detect vehicles accurately across a wide range of scenarios. This customization focuses on refining the dataset to address unique challenges such as occlusions, varying vehicle types, and dynamic traffic conditions, thereby bridging gaps in existing approaches.

Despite the remarkable advancements in object detection models, several gaps persist in the application of these technologies to intelligent transportation systems. Traditional models often struggle with real-world complexities, such as occlusions, inconsistent lighting, and the need for real-time processing in dynamic urban environments [9]. Generic datasets, while extensive, frequently lack the contextual specificity required for urban traffic management applications. These limitations hinder the scalability and adaptability of existing systems, emphasizing the need for tailored solutions.

The primary objective of this research is to address these gaps by enhancing the YOLOv8 model for vehicle detection through targeted improvements in both dataset customization and model architecture. By utilizing the rich diversity of the BDD100K dataset and optimizing the model to meet the demands of urban traffic scenarios, this study aims to advance the state of the art in vehicle detection. Our approach seeks to improve detection accuracy, robustness, and computational efficiency, paving the way for smarter, more sustainable traffic management solutions [9].

2. METHODOLOGY

2.1. Dataset Preparation

The BDD100K dataset was used as the database for annotated images for this study. This dataset includes labels for a variety of object categories, such as vehicles, pedestrians, and traffic signs. From the database we used image content under category "vehicle" including cars, trucks, buses, motorbikes, and bicycles. The dataset was split into training (70%), validation (20%), and testing (10%) subsets based on the conventional ratio to ensure a balanced evaluation of model performance. Original annotations in JSON format were converted to the YOLO format by normalizing bounding box coordinates relative to image dimensions. A custom script developed in Google Colab was utilized to automate this conversion process.

2.2. Model Training

The YOLOv8 model was fine-tuned on the prepared BDD100K dataset. During training, the model was fed annotated images to predict bounding boxes and class probabilities. Predictions were compared to ground truth labels, and parameters were iteratively updated to minimize error. This optimization process was conducted over 25 epochs or until performance plateaued on the validation set. Key hyperparameters included a learning rate of 0.001, batch size of 16, and an input image resolution of 640x640 pixels. To enhance the model's robustness, data augmentation techniques such as horizontal flipping, random scaling, and color jittering were employed.

2.3. Model Architecture Enhancements

YOLOv8 architectural was selected to enhance backbone networks and anchor-free detection heads. The model architecture was built to support real-time inference (Figure 1). Different versions of YOLOv8—nano (YOLOv8n), small (YOLOv8s), medium (YOLOv8m), large (YOLOv8l), and extra-large (YOLOv8x)—were evaluated to identify the optimal trade-off between computational efficiency and detection accuracy for this study.

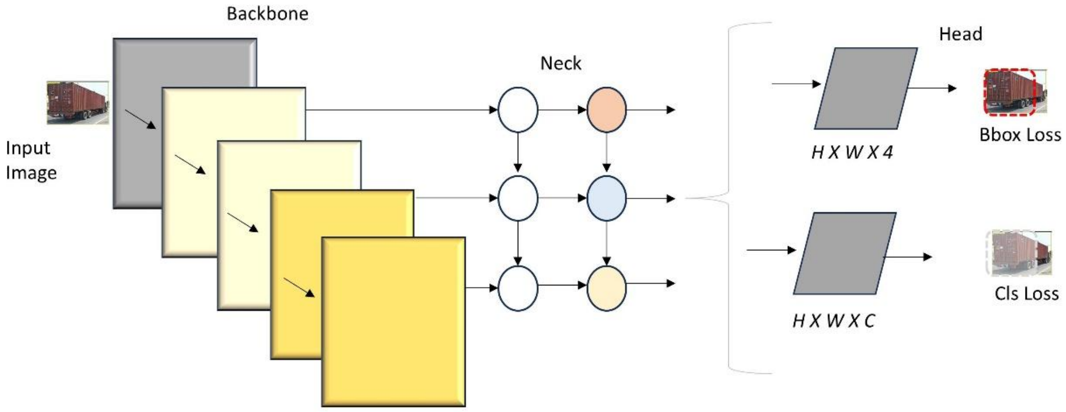


Figure 1: YOlOv8 model architecture.

2.4. Integration of Weights and Biases

To streamline experiment tracking and performance visualization, the Weights and Biases (W&B) platform was integrated into the training pipeline. W&B enabled real-time monitoring of key metrics such as loss, precision, recall, and mean average precision (mAP). This integration facilitated iterative improvements and ensured the reproducibility of results. By leveraging W&B's collaborative tools, adjustments to hyperparameters and augmentation strategies were efficiently managed.

2.5. Performance Evaluation

The model's effectiveness was evaluated using standard metrics, including precision, recall, F1-score, and mAP at various Intersection over Union (IoU) thresholds. Precision was used to measure the proportion of true positives among all positive detections, while recall assessed the model's ability to identify relevant instances. mAP metrics, including mAP@50 and mAP@50-95, quantified detection accuracy across multiple IoU thresholds. The following equations were used.

$\text{Precision} = \frac{\text{TP}}{\text{TP+FP}}$ --------------------(1)

3.

$\text{Recall} = \frac{\text{TP}}{\text{TP+FN}}$-------------------------(2)

4.

$\text{F1} = \frac{\text{2TP}}{\text{2TP+FP+FN}}$-------------------------(3)

2.6. Post-Processing and Optimization

Post-processing techniques, such as non-maximum suppression (NMS) and confidence thresholding, were applied to refine detection outputs. NMS was used to eliminate redundant overlapping bounding boxes, while confidence thresholding filtered out low-confidence predictions. These optimizations were carried to improve the clarity and reliability of detection results.

3. RESULTS

3.1. Training Dynamics and Progression

The YOLOv8 model's performance on the BDD100K dataset was evaluated over 25 epochs, demonstrating a clear progression in precision, recall, mAP metrics, and loss reduction. During the early epochs (1–5), the model showed significant classification and box losses, which steadily decreased as the model learned object features. For example, the box loss reduced from 0.74 in the first epoch to 0.76 by epoch 5, while classification loss decreased from 1.16 to 0.93. Precision and recall also improved during this phase, with precision increasing from 0.81 to 0.74 and recall stabilizing around 0.72–0.78. Mean Average Precision (mAP) metrics saw a gradual increase, with mAP@50 rising from 0.67 to 0.79 and mAP@50-95 improving from 0.63 to 0.89 (Figure 2).

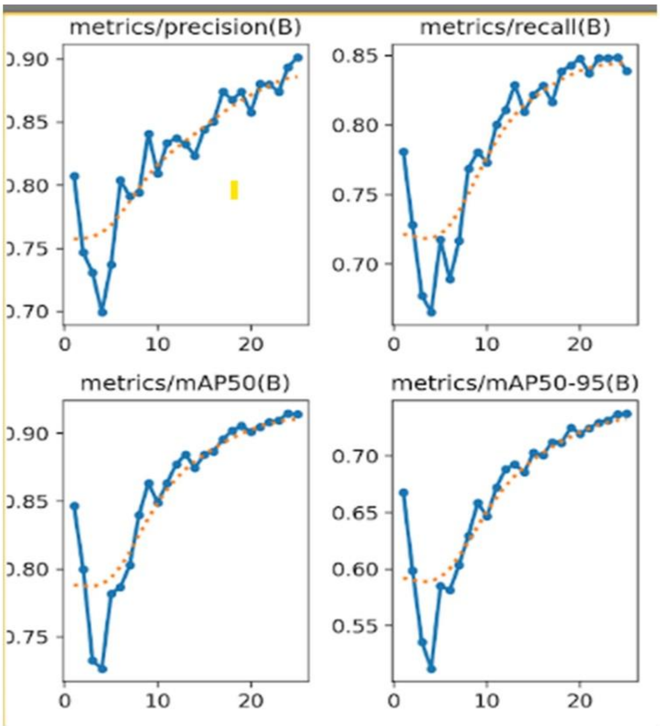


Figure 2: Training metrics (Precision, Recall, mAP50, mAP50-95) over epochs.

In the middle epochs (6–20), the model exhibited steady improvements in performance metrics, with box loss and classification loss stabilizing between 0.7–0.8 and 0.4–0.5, respectively. Precision peaked at 0.90, and recall improved to 0.84 by epoch 15. mAP@50 reached approximately 0.88–0.90, while mAP@50-95 improved to 0.74, highlighting the model's enhanced capability to detect and localize objects (Figure 3).

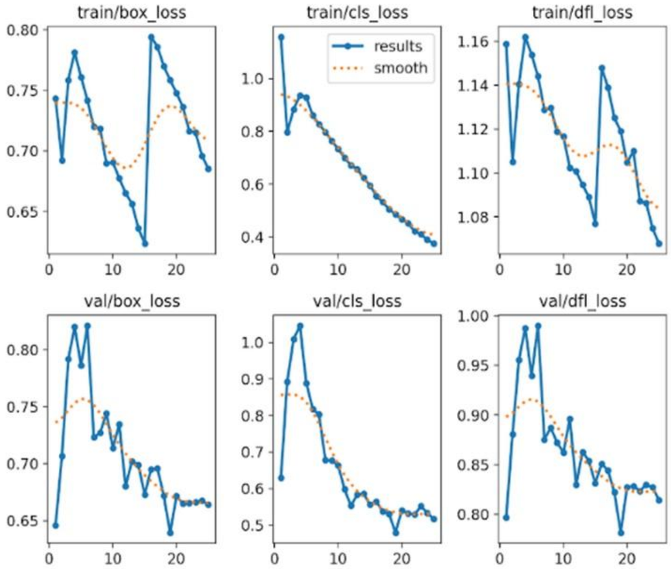


Figure 3: Training and validation loss trends.

During the final epochs (21–25), the model achieved optimal performance, as losses minimized and metrics stabilized. Box loss remained steady around 0.69–0.70, and classification loss dropped further to 0.37 by the 25th epoch. Precision was maintained at 0.90, and recall fluctuated slightly between 0.84 and 0.85. The mAP metrics were consistent, with mAP@50 reaching 0.91 and mAP@50-95 stabilizing at 0.74. These results signify the model's convergence to an optimal solution. A comparative analysis with the YOLOv5 model highlighted the superiority of YOLOv8, which achieved an F1-score of 0.87, a mAP@50 of 0.91396, and a mAP@50-95 of 0.74, outperforming YOLOv5 across all metrics. The table 1 below summarizes the training progress and key metrics during the 25 epochs:

Table 1: Performance Metrics of YOLOv8 Model Across 25 Epochs on the BDD100K Dataset

Epoch	Box Loss	Cls Loss	Precision (B)	Recall (B)	mAP50 (B)	mAP50-95 (B)	Val Box Loss	Val Cls Loss
1	0.74	1.16	0.81	0.78	0.85	0.67	0.65	0.63
2	0.69	0.80	0.75	0.73	0.80	0.60	0.71	0.89
3	0.76	0.88	0.73	0.68	0.73	0.54	0.79	1.01
4	0.78	0.94	0.70	0.67	0.73	0.51	0.82	1.05
5	0.76	0.93	0.74	0.72	0.78	0.59	0.79	0.89
21	0.74	0.45	0.88	0.84	0.90	0.72	0.67	0.53
22	0.72	0.42	0.88	0.85	0.91	0.73	0.67	0.53
23	0.72	0.41	0.87	0.85	0.91	0.73	0.67	0.55
24	0.70	0.39	0.89	0.85	0.91	0.74	0.67	0.53
25	0.69	0.37	0.90	0.84	0.91	0.74	0.66	0.52

3.2. Comparison with Prior Models

The YOLOv8 model outperformed its predecessor, YOLOv5, as summarized in Table 2. YOLOv8 achieved a higher F1-score of 0.87 compared to YOLOv5's 0.67, a marked improvement in mAP@0.5 (0.91 vs. 0.682), and a significant enhancement in mAP@0.5-0.95 (0.74 vs. 0.449). These results underscore the advancements in YOLOv8's architecture, including its anchor-free detection head and optimized backbone network, which collectively contributed to superior object detection performance. High precision values across all classes, including buses and bicycles, have shown to reduce false positives effectively (Figure 4). Similarly, the model demonstrates a strong balance between precision and recall, as indicated by its peak F1 score of 0.87, which is especially evident for trucks and buses (Figure 5). Additionally, the recall values across confidence thresholds highlight YOLOv8's ability to detect true positives consistently, particularly for buses and trucks (Figure 6).

Table 2: Performance comparison between YOLOv8 and YOLOv5

Metric	YOLOv8	YOLOv5	Improvement (%)
F1-Score	0.87	0.67	+29.85%
mAP@50	0.91	0.682	+33.43%
mAP@50-95	0.74	0.449	+64.81%
Precision	0.90	0.78	+15.38%
Recall	0.85	0.72	+18.06%

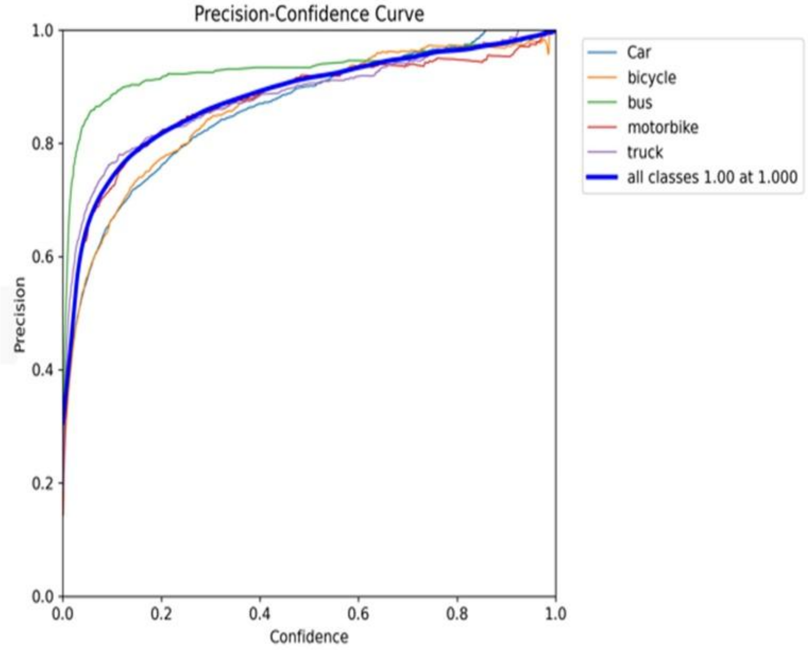


Figure 4: Precision-Confidence curve for YOLOv8 across various object classes.

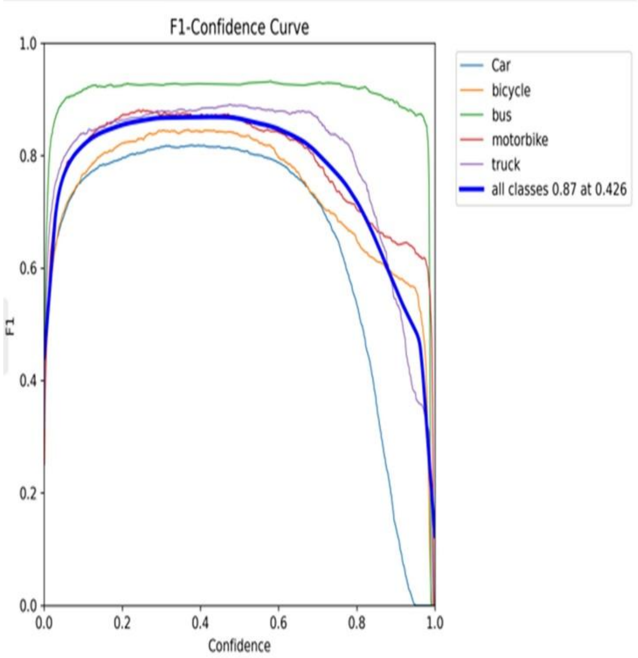


Figure 5: F1-Confidence Curve illustrating the balance between precision and recall for YOLOv8.

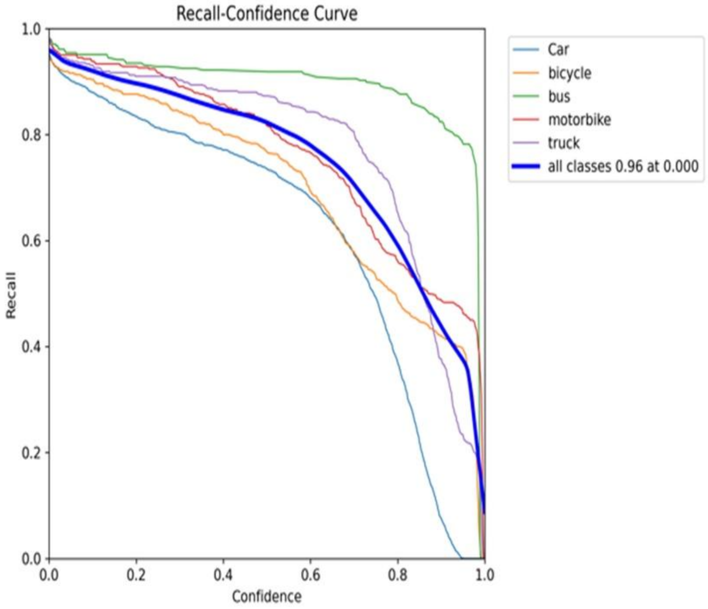


Figure 6: Recall-confidence curve.

3.3. Model Performance on Specific Metrics

The precision-recall curve achieved high precision and recall across all object classes. Notable values include 0.968 for buses, 0.929 for trucks, and an overall mAP@0.5 of 0.914 (Figure 7).

The YOLOv8 model demonstrated strong performance across key object detection metrics. High precision (0.90) and recall (0.85) in the final epochs indicated that the model effectively minimized false positives while identifying most relevant objects. The consistent mAP@50 score of 0.91 and mAP@50-95 of 0.74 across the final epochs reflected the model's robust detection capabilities for all object classes. Additionally, the decreasing trends in box and classification loss metrics over epochs validated the optimization strategies employed, with final values of 0.69 for box loss and 0.37 for classification loss, respectively.

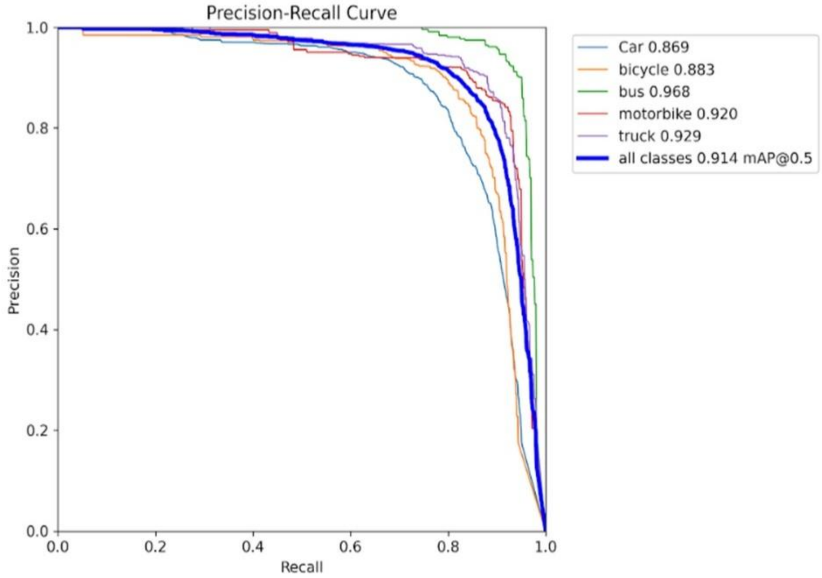
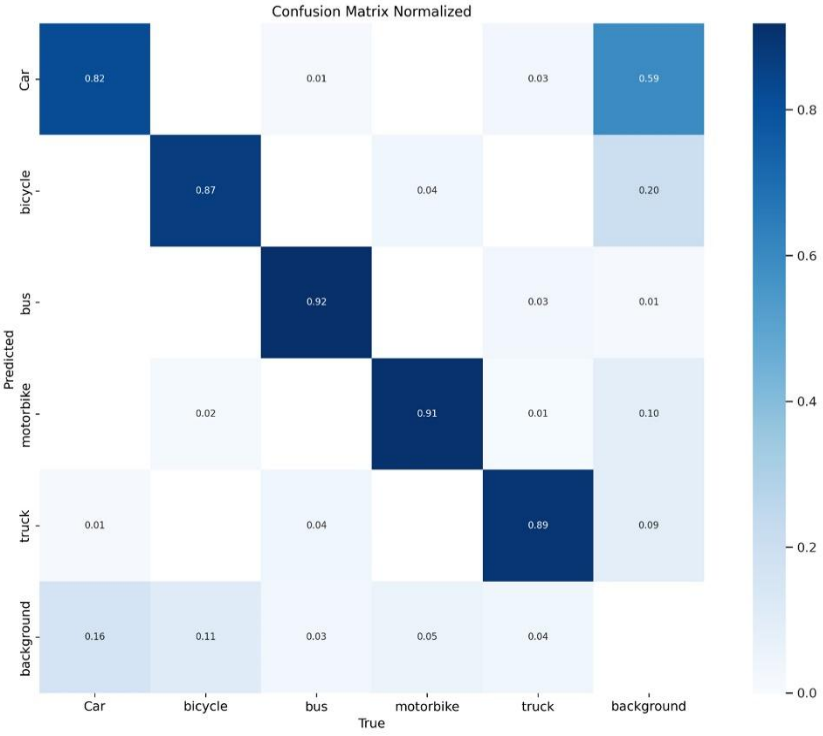


Figure 7: Precision-Recall curve for YOLOv8.

3.4. Specific Object Detection Performance

The YOLOv8 model successfully detected and classified objects from the "vehicle" category (cars, trucks, buses, motorbikes, and bicycles) within the BDD100K dataset. Post-processing techniques such as NMS and confidence thresholding further refined the model's outputs, reducing overlapping bounding boxes and filtering low-confidence predictions. This contributed to the overall precision and clarity of detections.

The normalized confusion matrix (Figure 8) illustrates the YOLOv8 model's detection accuracy across various object categories in the BDD100K dataset. Notably, the matrix indicates high detection accuracy for buses and motorbikes, with values of 0.92 and 0.91, respectively. These values reflect the model's robust ability to correctly classify these object categories with minimal errors. However, certain categories, such as the background, exhibited some degree of misclassification, as evidenced by higher off-diagonal values. This indicates instances where background elements were mistakenly identified as objects or vice versa, highlighting areas for potential model refinement.



	Car	bicycle	bus	motorbike	truck	background
Car	0.82		0.01		0.03	0.59
bicycle		0.87		0.04		0.20
bus			0.92		0.03	0.01
motorbike		0.02		0.91	0.01	0.10
truck	0.01		0.04		0.89	0.09
background	0.16	0.11	0.03	0.05	0.04	

Figure 8: Normalized confusion matrix showing YOLOv8 model's classification accuracy across object categories.

4. DISCUSSION

In this study, the YOLOv8 model demonstrated exceptional performance in detecting and classifying objects from the "vehicle" category—comprising cars, trucks, buses, motorbikes, and bicycles within the BDD100K dataset. These results indicate that YOLOv8 is well-suited for object detection tasks in a variety of real-world scenarios, particularly in the context of vehicle detection, which is a common and challenging task in autonomous driving, traffic monitoring, and smart city applications. Furthermore, YOLOv8's strong loss reduction trends align with broader findings on the importance of optimized training strategies in object detection [11,12].

While comparing older versions, YOLOv5 through efficient, exhibited limitations in handling diverse datasets due to its reliance on anchor-based strategies and less sophisticated feature extraction techniques. YOLOv8, in contrast, leverages improved backbone architectures, such as CSPDarkNet [12], which enhance feature representation while maintaining real-time processing capabilities. This aligns with prior findings that integrating efficient feature pyramids and lightweight architectures significantly boosts performance on large-scale datasets.

In the middle epochs (6–20), YOLOv8 demonstrated consistent improvement and stabilization across key metrics. The observed stabilization of box loss between 0.7–0.8 and classification loss between 0.4–0.5 aligns with trends reported in object detection literature, where intermediate training stages are critical for balancing precision and recall without overfitting [12]. The precision peaking at 0.90 and recall improving to 0.84 by epoch 15 indicate the model's ability to accurately identify relevant objects while maintaining minimal false positives. This balance is essential for robust detection in real-world scenarios, particularly in complex datasets like BDD100K, which include diverse objects and challenging conditions. The consistent mAP@50 range of 0.88–0.90 and the improvement of mAP@50-95 to 0.74 is essential for enhanced localization and detection capabilities during this phase [13]. The stabilization of metrics during this period demonstrates that YOLOv8's optimization strategies effectively mitigate loss and maintain high performance, even as learning progresses. Similar trends have been observed in models with advanced architectural designs, such as the inclusion of optimized backbone networks or anchor-free detection mechanisms [14]. These advancements highlight the importance of architectural innovation in improving intermediate-stage training outcomes and overall detection performance.

During the final epochs (21–25), the YOLOv8 model demonstrated convergence to an optimal solution, as evidenced by the stabilization of metrics and minimization of losses. The box loss consistently remained between 0.69 and 0.70, while classification loss decreased further to 0.37 by the 25th epoch. These trends align with the behavior of well-trained deep learning models, where late-stage optimization fine-tunes the weights, reducing overfitting and ensuring robustness in predictions [15]. The precision maintained at 0.90 and the slight fluctuation in recall (0.84–0.85) suggest a stable balance between minimizing false positives and maximizing true positive detections, a hallmark of reliable object detection systems. The consistency of mAP metrics, with mAP@50 at 0.91 and mAP@50-95 stabilizing at 0.74, pinpoints its accuracy in detection and localization across a range of Intersection over Union (IoU) thresholds, reflecting strong generalization capabilities [16].

The results of our comparative study between YOLOv5 and YOLOv8 highlight a marked improvement in performance with the latter, aligning with recent advancements in the field of object detection. YOLOv8’s superior F1-score, mAP@50, and mAP@50-95 suggest that the new architecture is more capable in terms of both precision and recall, as well as handling a wider range of object sizes and detection challenges, compared to its predecessor. This improvement is consistent with trends observed in the evolution of the YOLO series, where newer versions have consistently outperformed older models across various benchmarks [17].

Recent work by Shah et al. [18] demonstrated that advancements in deep learning architectures—such as the integration of more sophisticated feature extraction mechanisms and attention-based modules—can significantly enhance detection accuracy and robustness, particularly for small or occluded objects [18]. The high mAP@50-95 achieved by YOLOv8 is particularly noteworthy, as this metric accounts for performance over a range of IoU thresholds, making it more indicative of the model’s ability to generalize across various detection scenarios [16].

Our results align with findings from Terven et al. [19], who observed that the performance improvements in newer YOLO versions can be attributed to architectural refinements, such as the incorporation of advanced convolutional operations and more efficient training strategies [19].

YOLOv8’s higher F1-score further suggests that the model achieves a more favorable balance between false positives and false negatives compared to YOLOv5, which is particularly important in domains like medical image analysis and autonomous driving, where both false positives and false negatives can have significant consequences [20].

Additionally, while YOLOv8 outperforms YOLOv5 in terms of key metrics, it is crucial to consider the trade-offs between model accuracy and computational efficiency. For instance, Tan et al. [21] highlight that more complex models often require increased computational resources, which can limit their applicability in real-time systems [21]. One limitation of our study is that we evaluated the models on a single dataset, which may not fully represent the model's ability to generalize to other domains. As suggested by Hussain et al. [22], a broader evaluation across diverse datasets is necessary to better understand the generalizability of YOLOv8, especially when deployed in real-world conditions [22]. Moreover, although YOLOv8 shows promising results in terms of detection performance, its robustness to noise, variations in lighting, and other real-world challenges remains to be fully tested.

The high precision values across all object classes, particularly for challenging categories such as buses and bicycles, indicate that YOLOv8 is effective at reducing false positives. This finding is consistent with recent studies that emphasize the importance of precise object localization in complex detection tasks [16]. The peak F1-score of 0.87 and the strong balance between precision and recall also highlight YOLOv8’s efficiency in minimizing both false positives and false negatives, making it a reliable model for applications where these trade-offs are critical, such as in autonomous driving and surveillance systems [20]. The recall values across different confidence thresholds confirm that YOLOv8 is adept at detecting true positives, particularly for large, easily identifiable objects like trucks and buses. Additionally, the decreasing trends in box and classification loss metrics over the course of training suggest that YOLOv8 benefited from an efficient optimization process. Final values of 0.69 for box loss and 0.37 for classification loss indicate that the model converged well, minimizing both the spatial misalignment of bounding boxes and errors in classification [23). These trends are consistent with the idea that a well-optimized detection model can achieve high accuracy without overfitting to the training data, as was demonstrated in previous studies of YOLO-based architectures [17].

One of the key factors contributing to the improved performance of YOLOv8 is its use of advanced post-processing techniques, particularly NMS and confidence thresholding. NMS is a widely adopted technique used to eliminate redundant and overlapping bounding boxes by retaining the box with the highest confidence score among those that share significant overlap [17,24]. By applying NMS, YOLOv8 effectively reduces false positives. Confidence thresholding further refines the model’s predictions by filtering out low-confidence detections, ensuring that only highly reliable detections are retained.

Although YOLOv8 exhibited strong performance, it is essential to consider the limitations inherent in post-processing. While NMS and confidence thresholding can effectively refine predictions, they may also lead to a reduction in recall, particularly when objects are in close proximity or occluded by other vehicles. Future work could explore alternative strategies, such as soft-NMS or multi-scale NMS, to further enhance detection accuracy in crowded scenes. Moreover, further optimization of the confidence thresholding process could potentially improve recall without significantly sacrificing precision, especially in scenarios where detecting all instances of a particular object is crucial.

5. CONCLUSION

Using the BDD100K dataset, this study showed that the YOLOv8 model outperformed other models in recognizing and classifying items in the "vehicle" category. YOLOv8 achieved high precision, recall, and mAP values, significantly outperforming its predecessor. Advanced detection modality such as an anchor-free detection head, an optimized backbone network, and effective post-processing techniques such as NMS and confidence thresholding, increased detection capacity. These findings demonstrate YOLOv8's robustness, efficiency, and flexibility to complicated object detection tasks in real-world settings. The findings indicate that YOLOv8 is a promising option for applications such as autonomous driving, traffic monitoring, and smart city systems. Future study should concentrate on improving its generalization skills and overcoming obstacles in detecting small and obstructed targets.

Funding: This study was not supported by any fund.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Dataset available at https:// BDD100K Dataset | Papers With Code and the source code and experimental data will be made available upon request, subject to permission from the corresponding author.

Acknowledgments: Thank you to everyone who reviewed and provided feedback to make the article more successful.

Conflicts of Interest: The authors declare that they have no competing interests.

Declaration of generative AI in scientific writing: During the preparation of this work the author(s) used Grammarly in order to improve the language of the manuscript.

References

1. Qiu, L., Zhang, D., Tian, Y., & Al-Nabhan, N.: Deep learning-based algorithm for vehicle detection in intelligent transportation systems. J. Supercomput. 77, 11083–11098 (2021).
2. Djahel, S., Doolan, R., Muntean, G. M., & Murphy, J.: A communications-oriented perspective on traffic management systems for smart cities: Challenges and innovative approaches. IEEE Commun. Surv. Tutor. 17, 125–151 (2014).
3. Ouallane, A. A., Bakali, A., Bahnasse, A., Broumi, S., & Talea, M.: Fusion of engineering insights and emerging trends: Intelligent urban traffic management system. Inf. Fusion 88, 218–248 (2022).
4. Ali, M. L., & Zhang, Z.: The YOLO framework: A comprehensive review of evolution, applications, and benchmarks in object detection. Comput. 13, 336 (2024).
5. Gallagher, J. E., & Oughton, E. J.: Surveying You Only Look Once (YOLO) multispectral object detection advancements, applications and challenges. IEEE Access (2025).
6. Sapkota, R., Qureshi, R., Calero, M. F., Badjugar, C., Nepal, U., Poulose, A., & Karkee, M.: YOLOv10 to its genesis: A decadal and comprehensive review of the You Only Look Once (YOLO) series. arXiv preprint arXiv:2406.19407 (2024).
7. Husain, A. A., Maity, T., & Yadav, R. K.: Vehicle detection in intelligent transport system under a hazy environment: A survey. IET Image Process. 14, 1–10 (2020).
8. Yu, F., Chen, H., Wang, X., Xian, W., Chen, Y., Liu, F., & Darrell, T.: BDD100K: A diverse driving dataset for heterogeneous multitask learning. In Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 2636–2645 (2020).
9. Shao, Z., Yang, K., Sun, P., Hu, Y., & Boukerche, A.: The evolution of detection systems and their application for intelligent transportation systems: From solo to symphony. Comput. Commun. 225, 96–119 (2024).
10. Bochkovskiy, A., Wang, C. Y., & Liao, H. Y. M.: YOLOv4: Optimal speed and accuracy of object detection. arXiv preprint arXiv:2004.10934 (2020).
11. Wang, C. Y., Bochkovskiy, A., & Liao, H. Y. M.: YOLOv7: Trainable bag-of-freebies sets new state-of-the-art for real-time object detectors. In Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 7464–7475 (2023).
12. Safaldin, M., Zaghden, N., & Mejdoub, M.: An improved YOLOv8 to detect moving objects. IEEE Access (2024).
13. Ling, H., Zhao, T., Zhang, Y., & Lei, M.: Engineering vehicle detection based on improved YOLOv6. Appl. Sci. 14, 8054 (2024).
14. Tam, C. M., Zhang, D., Chen, B., Peters, T., & Li, S.: Holistic multitask regression network for multiapplication shape regression segmentation. Med. Image Anal. 65, 101783 (2020).
15. Loshchilov, I., & Hutter, F.: SGDR: Stochastic gradient descent with warm restarts. arXiv preprint arXiv:1608.03983 (2016).
16. Lin, T. Y., Maire, M., Belongie, S., Hays, J., Perona, P., Ramanan, D., & Zitnick, C. L.: Microsoft COCO: Common objects in context. In Comput. Vis.–ECCV 2014: 13th Eur. Conf., Zurich, Switzerland, Sept. 6–12, 2014, Proc., Part V, 740–755. Springer International Publishing (2014).
17. Redmon, J.: You only look once: Unified, real-time object detection. In Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (2016).
18. Shah, S., & Tembhurne, J.: Object detection using convolutional neural networks and transformer-based models: A review. J. Electr. Syst. Inf. Technol. 10, 54 (2023).
19. Terven, J., Córdova-Esparza, D. M., & Romero-González, J. A.: A comprehensive review of YOLO architectures in computer vision: From YOLOv1 to YOLOv8 and YOLO-NAS. Mach. Learn. Knowl. Extr. 5, 1680–1716 (2023).
20. Badrinarayanan, V., Kendall, A., & Cipolla, R.: SegNet: A deep convolutional encoder-decoder architecture for image segmentation. IEEE Trans. Pattern Anal. Mach. Intell. 39, 2481–2495 (2017).
21. Tan, M., Pang, R., & Le, Q. V.: EfficientDet: Scalable and efficient object detection. In Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., 10781–10790 (2020).

22. Hussain, M.: YOLOv1 to v8: Unveiling each variant–A comprehensive review of YOLO. IEEE Access 12, 42816–42833 (2024).
23. Wu, R., Chowdhury, Z., Sanchez, G. V., Gao, X., Villa, C., & Jiang, X.: Real-time vehicle detection system for intelligent transportation using machine learning. In 2022 IEEE Green Energy Smart Syst. Conf. (IGESSC), 1–6. IEEE (2022, Nov.).
24. Balakrishnan, B., Chelliah, R., Venkatesan, M., & Sah, C.: Comparative study on various architectures of YOLO models used in object recognition. In 2022 Int. Conf. Comput., Commun., Intell. Syst. (ICCCIS), 685–690. IEEE (2022, Nov.).