
YOLO-LFA: Lightweight Feature Attention for Real-Time Vehicle Detection

Jasmine Selvakumari Jeya I, Priyadarshi Nihal, Pratyush Dubey, Dev Tailor,
Dattatreya, Shivalik Mathur

School of Computing Science and Engineering, VIT Bhopal University, Sehore, Madhya Pradesh, India, wjasminejeya@gmail.com, priyadarshinihal2022@vitbhopal.ac.in, pratyushdubey2022@vitbhopal.ac.in, devtailor2022@vitbhopal.ac.in, dattatreya2022@vitbhopal.ac.in, shivalikmathur2022@vitbhopal.ac.in

Abstract.

Real time vehicle detection has significant importance in the context of intelligent transportation and self-driving cars. The key challenge, however, is to prolong high prediction accuracy as utilizing lightweight architectures in live environment. To address the challenge, the authors propose the YOLO-LFA, which utilizes lightweight architectures for object detection. The model combines BiFPN-based multi-scale feature fusion, C2PSA, SE channel attention, DySample, and SEAM for improved feature quality and detection performance. The framework has been evaluated on the UA-DETRAC dataset and has displayed potential outputs compared to existing framework, such as YOLOv8, YOLO-MSN, and YOLO-BASW. The model has shown refined detection performance without compromising the lightweight framework. The model has considerable capacity in the context of traffic monitoring, self-driving cars and intelligent transportation systems, which desire efficient and accurate real-time vehicle detection.

Keywords. Vehicle Detection, Intelligent Transportation, YOLO-LFA, Self-driving Cars, Lightweight Framework.

1. INTRODUCTION

Real-time vehicle detection is essential for intelligent transportation systems, traffic monitoring, and autonomous vehicles. YOLO models provide a strong balance between detection accuracy and speed, but edge deployment remains challenging due to computational and memory constraints. This study proposes YOLO-LFA (Lightweight Feature Attention), integrating BiFPN, C2PSA, SE, DySample, and SEAM to improve detection accuracy while maintaining computational efficiency.

2. RELATED WORK

YOLOv5–YOLOv10 have significantly improved real-time object detection through enhanced backbones, feature aggregation, and anchor-free prediction [2–4]. Multi-scale methods such as FPN, PANet, and BiFPN [7] improve detection of vehicles at different sizes, while attention mechanisms such as SE [8] and CBAM enhance feature representation and localization. Building on these techniques, YOLO-LFA combines multi-scale fusion and attention mechanisms for efficient and accurate real-time vehicle detection.

3. KEY CONTRIBUTION

The key innovations of this study are as listed. A lightweight identification architecture, YOLO-LFA, is proposed for efficient real-time vehicle detection. Moreover, the model combines the benefits of BiFPN [7] and DySample [9] for the improvement of feature aggregation at multiple scales. Furthermore, the model includes C2PSA spatial attention and SE channel attention [8] for the improvement of feature extraction. Lastly, the SEAM detection head is used for the improvement of detection performance. The proposed model is validated through comparative experiments using YOLOv8 Base [2], YOLO-MSN [5], and YOLO-BASW [6].

4. METHODOLOGY

YOLO-LFA uses a backbone, BiFPN-based neck, and SEAM detection head to achieve accurate and lightweight vehicle detection. BiFPN enables multi-scale feature fusion, while DySample improves upsampling. C2PSA and SE blocks enhance spatial and channel attention, respectively, improving detection under occlusion and dense traffic.

The model was trained on 9,816 UA-DETRAC images at 640×640 resolution using augmentation techniques such as Mosaic, MixUp, Copy-Paste, and random erasing. Training was performed for 60 epochs on an NVIDIA T4 GPU using PyTorch, with Adam optimizer, a learning rate of 0.001, weight decay of 0.0005, and cosine learning-rate scheduling. The combined modules provide a balance between detection accuracy and computational efficiency for real-time ITS applications.

5. RESULT AND DISCUSSION

The improvement in YOLO-LFA can be explained by the incorporation of lightweight attention mechanisms along with dynamic feature aggregation. BiFPN [7] helps in enhancing multi-scale feature fusion, while DySample [9] helps in feature alignment during upsampling. Furthermore, the incorporation of spatial and channel attention [8] helps in focusing on specific features in complex traffic scenes. Table 1 exhibits a statistical evaluation of the enhanced YOLO-LFA model with related models in terms of parameters, mAP₅₀ and precision.

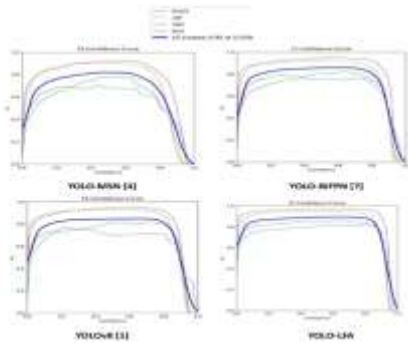


Figure 5.1. F1 Score Comparison across models

Table 1. Comparison of proposed model to existing models

Models	Parameters (10^6)	mAP50	Precision
YOLOv8 [2]	3.01	0.9090	0.8689
YOLO-MSN [5]	2.82	0.9120	0.8974
YOLO-BiFPN [7]	3.34	0.9220	0.9140
YOLO-LFA	2.76	0.9338	0.9242

The proposed YOLO-LFA yields the foremost detection functioning with a mAP₅₀ of 0.9338 and precision of 0.9242 while using the lowest number of parameters (2.76M). These output represent that YOLO-LFA presents a better tradeoff between model processing overhead and detection accuracy contrast to YOLOv8, YOLO-MSN and YOLO-BiFPN. Figure 1 shows how the F1-confidence curves of different YOLO models (YOLO-MSN, YOLO-BiFPN, YOLOv8, and YOLO-LFA) compare for different types of objects. The curves show that YOLO-LFA and YOLO-BiFPN have higher and more stable F1 scores over a wider range of confidence levels. This means that they are better at finding things. On the other hand, YOLO-MSN and YOLOv8 have F1 values that are lower or less stable, especially at higher confidence levels.

6. CONCLUSIONS

This study examines the constraints of facilitating accurate vehicle detection while preserving pace or model size, enabling setup in realtime Intelligent Transportation Systems (ITS) and edge environments. A lightweight detection framework named YOLO-LFA is proposed, which integrates BiFPN feature fusion, DySample dynamic upsampling, C2PSA spatial attention, SE channel attention, and the SEAM detection head to refine attribute representation and crossscale interactions. Experimental outcomes illustrate that the suggested framework obtains robust and efficient vehicle detection while maintaining strong computational performance. The current evaluation is conducted on the UA-DETRAC dataset. Future direction will focus on testing the framework on extended datasets and further improving it for real-world Edge-based system implementation.

7. REFERENCES

- [1] S.M. Parate, I.J.S. Jeya, ‘AttBi-ResU: A smart 3D vision-based lane detection system for autonomous vehicles’, IEEE Access, 2025.
- [2] Q. Fan, Y. Li, M. Deveci, K. Zhong, S. Kadry, ‘LUD-YOLO: A novel lightweight object detection network for unmanned aerial vehicle’, Information Sciences, vol. 686, p. 121366, 2025.
- [3] A.A. Murat, M.S. Kiran, ‘A comprehensive review on YOLO versions for object detection’, Engineering Science and Technology, an International Journal, vol. 70, p. 102161, 2025.
- [4] Q. Wu, X. Nie, ‘Improved YOLOv10: A real-time object detection approach in complex environments’, Sensors, vol. 25, no. 22, 2025.
- [5] J. Liu, Y. Xie, Y. Zhang, H. Li, ‘Vehicle flow detection and tracking based on an improved YOLOv8n and ByteTrack framework’, World Electric Vehicle Journal, vol. 16, no. 1, pp. 1-13, 2024.

- [6] Z. Zhao, S. Chen, P. Zhang, J. Chen, L. Liao, 'Research on the construction of real-time vehicle detection model based on YOLOv8-BASW', IEEE Access, vol. 13, pp. 1-15, 2025.
- [7] J. Doherty, B. Gardiner, E. Kerr, N. Siddique, 'BiFPN-YOLO: One-stage object detection integrating bi-directional feature pyramid networks', Pattern Recognition, vol. 160, 2025.
- [8] X. Chai, M. Zhao, J. Li, J. Li, 'Image small target detection in complex traffic scenes based on YOLOv8 multiscale feature fusion', Alexandria Engineering Journal, vol. 126, pp. 578-590, 2025.
- [9] C.K. Ha, H. Nguyen, L.H. Le, 'AAB-FusionNet: A real-time object detection model for UAV edge computing platforms', MethodsX, vol. 15, p. 103654, 2025.

Biographies



Jasmine Selvakumari Jeya I is currently a Senior Associate Professor with the School of Computing Science and Engineering, VIT Bhopal University, Madhya Pradesh. She brings over 24 years of experience in teaching and research. She has an impressive portfolio with more than 60 research articles published, 30 patents filed, and two patents granted.



Priyadarshi Nihal is currently pursuing the B.Tech. degree in Computer Science and Engineering at VIT Bhopal University, Bhopal, India. His research interests include artificial intelligence, computer vision, and intelligent perception systems, with a focus on deep learning-based object detection and real-time visual analytics.



Pratyush Dubey is presently pursuing the B.Tech. degree in Computer Science and Engineering at VIT Bhopal University, Bhopal, India. His research interests include artificial intelligence, computer vision, and intelligent systems, with a focus on deep learning techniques for computer vision applications.



Shivalik Mathur is currently pursuing the B.Tech. degree in Computer Science and Engineering at VIT Bhopal University, Bhopal, India. His research interests include software engineering, cloud computing, DevOps practices, and distributed systems.



Dattatrey is presently pursuing the B.Tech. degree in Computer Science and Engineering at VIT Bhopal University, Bhopal, India. His research interests include machine learning, deep learning, natural language processing, and computer science fundamentals such as operating systems, database management systems, and distributed computing.