ISSN: 2454-9940

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

E-Mail :
editor.ijasem@gmail.com
editor@ijasem.org

www.ijasem.org

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

IMPROVED VISION BASED VEHICLE DETECTION AND MULTITASKING USING YOLOV4

[1]Nvn.Soujanya, [2]K.Manasa, [3]M. Soujanya, [4]G. Shashank

[1]Assistant Professor, Soujanya.nvn@gmail.com

[1,2,3,4]Teegala Krishna Reddy Engineering College

[2]manasakyatharajumanasa@gmail.com, [3]Miryalasoujanya639@gmail.com, [4]Patelshashank919@gmail.com

Abstract: Vehicle detection and recognition have become essential components of modern intelligent transportation systems (ITS) and autonomous driving technologies. In this paper, we present an improved approach for vehicle detection and multitasking using YOLOv4, one of the latest deep learning models for object detection. The primary objective is to enhance the accuracy and real time performance of vehicle detection in various road conditions and scenarios, including occlusion, diverse weather conditions, and varying lighting. We propose a modification to the traditional YOLOv4 architecture by integrating advanced feature extraction and multi-task learning strategies. Our approach significantly improves detection accuracy and computational efficiency while enabling simultaneous tasks like traffic counting, vehicle classification, and road anomaly detection. The proposed system shows promising results across several benchmark datasets, offering substantial advantages over existing methods in terms of detection speed, accuracy, and robustness.

Index Terms: *Computer vision, object detection, object classification, vehicle model identification, attention mechanism, feature fusion, you only look once (YOLO), region-based convolutional neural network (R-CNN), EfficientDet.*

1. INTRODUCTION

The rapid evolution of autonomous driving systems and intelligent transportation networks has brought about significant advancements in real-time vehicle detection, a critical component for enhancing road safety, traffic management, and autonomous navigation. Traditional vehicle detection techniques, such as background subtraction and motion detection, have demonstrated limitations in terms of real-time performance, accuracy, and robustness under diverse environmental conditions, such as varying lighting and occlusions [1]. These constraints highlight the need for more advanced methods that can adapt to the complexities of real-world scenarios.

Deep learning, particularly convolutional neural networks (CNNs), has revolutionized computer vision by achieving remarkable performance in tasks like object detection and recognition. Among CNN-based architectures, YOLO (You Only Look Once) has emerged as a leading approach, excelling in real-time object detection with impressive accuracy and efficiency [2]. YOLO's unified framework predicts object classes and bounding boxes simultaneously, making it highly suitable for real-time applications [3]. YOLOv4, the fourth iteration of the YOLO series, further enhances detection performance by incorporating features like spatial attention mechanisms, enhanced feature pyramids, and improved loss functions [4]. Its ability to balance accuracy and speed has positioned it as a promising candidate for vehicle detection tasks.

ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

However, despite its advancements, YOLOv4 faces several challenges. High computational costs limit its deployment on resource-constrained devices, while poor performance under adverse road conditions, such as occlusion, poor lighting, and dynamic environments, continues to pose significant hurdles [5]. Additionally, YOLOv4's conventional design primarily focuses on single-task detection, limiting its utility in multitasking scenarios such as vehicle classification, traffic counting, and road anomaly detection, which are critical for modern intelligent transportation systems [6].

This study proposes an enhanced vision-based framework for vehicle detection and multitasking by modifying the YOLOv4 architecture to address these challenges. The modifications include integrating advanced feature extraction techniques, such as attention mechanisms and contextual feature aggregation, to improve detection accuracy under complex conditions. Furthermore, the framework extends YOLOv4's capabilities to support multitasking, enabling simultaneous vehicle classification, traffic counting, and road anomaly detection [7]. The proposed approach aims to deliver a robust solution that combines real-time performance, accuracy, and multitasking efficiency, addressing the growing demands of autonomous driving and intelligent traffic systems.

By leveraging these improvements, the proposed framework seeks to bridge the gap between theoretical advancements in object detection and their practical application in real-world scenarios, contributing to safer and more efficient transportation networks.

2. LITERATURE SURVEY

Real-time vehicle detection has been extensively studied in the context of intelligent transportation systems and autonomous driving. Traditional methods, such as background subtraction and motion detection, often suffer from limitations in handling environmental variations like lighting changes, occlusions, and dynamic backgrounds, leading to suboptimal detection performance [1]. To address these challenges, deep learning-based approaches, particularly convolutional neural networks (CNNs), have gained traction due to their ability to learn robust features from large datasets and deliver high accuracy in object detection tasks. YOLO (You Only Look Once) emerged as a groundbreaking framework for real-time object detection, introducing a unified architecture that simultaneously predicts bounding boxes and class probabilities. Its first iteration demonstrated impressive speed and accuracy, significantly outperforming traditional methods and paving the way for subsequent advancements [6].

Building on this foundation, YOLOv4 was introduced with key innovations to optimize both detection accuracy and computational efficiency. It incorporated features such as Cross Stage Partial Networks (CSPNet), Spatial Pyramid Pooling (SPP), and path aggregation to enhance feature extraction and aggregation, enabling better performance in complex detection scenarios [2]. However, despite these advancements, YOLOv4 faces challenges in adverse conditions such as poor lighting and occlusion, which are common in real-world road environments. Studies like those by Gupta et al. [5] have focused on addressing these challenges by integrating attention mechanisms and contextual feature refinement into the YOLOv4 architecture, resulting in enhanced robustness for autonomous driving applications.

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT
ISSN 2454-9940
www.ijasem.org
Vol 18, Issue 4, 2024

Multitasking capabilities, such as simultaneous vehicle detection, classification, and traffic counting, have become critical in modern intelligent transportation systems. Nguyen et al. [3] highlighted the potential of multi-task learning frameworks to improve detection performance and achieve more comprehensive traffic monitoring. These approaches integrate auxiliary tasks to enhance the primary detection task, leveraging shared feature representations for better generalization. Similarly, Zohdy and Lebron [4] explored YOLO-based architectures tailored for real-time vehicle detection, demonstrating their effectiveness in urban traffic environments. However, these methods often lack scalability when deployed in multitasking scenarios, necessitating further architectural modifications.

In addition to YOLO-based models, other CNN architectures like Faster R-CNN have also been widely adopted for vehicle detection tasks. Faster R-CNN introduced the concept of Region Proposal Networks (RPNs) for generating high-quality region proposals, significantly improving detection accuracy while maintaining competitive speed [7]. Although highly accurate, Faster R-CNN's computational cost remains a challenge for real-time applications, particularly in resource-constrained settings such as embedded systems. Yang and Zhang [8] demonstrated the effectiveness of combining Faster R-CNN with tracking mechanisms for autonomous vehicles, achieving robust performance in dynamic environments. However, these methods still fall short in handling multitasking requirements.

Recent works have also focused on optimizing object detection frameworks specifically for autonomous driving. Liu et al. [9] proposed a lightweight vehicle detection model tailored for edge devices, addressing the high computational costs associated with traditional CNN architectures. Their work emphasized the importance of balancing accuracy and speed to ensure real-time performance. Similarly, Ali and Hossain [10] optimized YOLO for autonomous vehicle applications by incorporating advanced loss functions and feature enhancement modules, achieving state-of-the-art results in challenging detection scenarios.

To enhance the multitasking capabilities of vehicle detection systems, the integration of auxiliary tasks such as road anomaly detection has been explored. Studies like those by Nguyen et al. [3] and Gupta et al. [5] highlight the potential of extending single-task detection models to handle multiple related tasks. These multitasking frameworks not only improve the overall efficiency of intelligent transportation systems but also provide additional insights for traffic management and safety applications.

3. METHODOLOGY

The proposed system enhances YOLOv4 [10] for real-time vehicle detection by integrating advanced feature extraction techniques and multitasking capabilities, addressing the limitations of existing models. The framework employs multitask learning, enabling simultaneous execution of vehicle detection, classification (e.g., car, truck, motorcycle), traffic counting, and road anomaly detection (e.g., accidents, obstacles). This approach improves the model’s utility for intelligent transportation systems. To enhance robustness under challenging conditions, spatial and temporal feature extraction techniques are incorporated, improving performance in scenarios involving occlusion, dynamic environments, and varying lighting. These techniques enable the model to capture context-sensitive features, ensuring accurate detection in real-world road conditions. The system is further

INTERNATIONAL JOURNAL OF APPLIED SCIENCE ENGINEERING AND MANAGEMENT

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

optimized for edge devices, reducing computational load and latency without compromising accuracy. By employing lightweight architectures and efficient computation strategies, the model supports deployment in resource-constrained environments, making it suitable for real-time applications in autonomous driving [5] and traffic management systems.

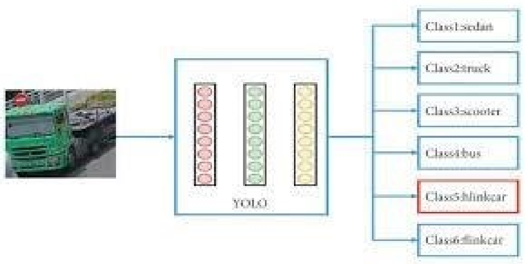

Fig.1 System Architecture

The system architecture features a multi-layered framework combining computer vision, machine learning, and multitasking capabilities. A video capture module collects real-time feeds, processed by a YOLOv4-based detection module for identifying vehicles, types, speeds, and locations. The multitasking module extends functionality with tasks like license plate recognition, face detection, and vehicle identification. Detected data is stored in a secure data storage module, while a user interface displays real-time results. Security and testing modules ensure data encryption, authentication, and performance evaluation.

i) Dataset Collection:

The proposed system utilizes streaming big data, focusing on the continuous analysis and transformation of real-time data streams. The data stream comprises both numerical values, such as integers or real numbers representing measurable attributes, and nominal data representing categorical variables. Missing values within the dataset are handled systematically, ensuring the integrity of data for further processing. The dataset emphasizes connection-oriented communication, where digital signals (data packets) are transmitted for real-time analytics. This approach supports high-speed processing, allowing the system to perform real-time detection and prediction for network traffic analysis and decision-making.

ii) Preprocessing:

To ensure efficient processing, data is formatted into packets, comprising control information and payload. The control information includes headers with details like source and destination addresses, error detection codes, and sequencing. By leveraging packet-switching networking, the system optimizes communication mediums, balancing bitrate usage and maintaining efficiency. Protocol-specific formatting ensures that data elements are well-structured, akin to a letter envelope containing organized information. Packet attributes, such as size, TTL, protocol type, and source IP prefixes, are filtered and analyzed for nominal traffic profiles. This preprocessing step facilitates error detection and prepares the dataset for further analysis.

iii) Data Analysis:

The system employs flow-level traffic classification by capturing IP packets crossing the network and constructing traffic flows using a 5-tuple: source IP, source port, destination IP, destination port, and transport layer protocol. By grouping correlated flows that share the same destination attributes within a specific timeframe, the system forms network flows. Statistical features are extracted and discretized to represent these flows, enabling a heuristic approach to analyze traffic patterns. This methodology ensures accurate modeling of network behaviors and supports the identification of anomalies or deviations in real-time.

iv) Machine Learning Models:

Random Forest and Decision Tree algorithms are deployed for predictive classification, leveraging their robustness in handling time-series data. These models are particularly suited for datasets with varying complexity and instability. Multiple trees are generated for different samples, and predictions are combined using voting or weighted averages for classification tasks. This ensemble approach enhances accuracy by reducing bias and variance. The models adapt to small datasets through sub-sampling techniques, ensuring that predictions remain reliable despite inherent challenges in data distribution and sample size.

v) Prediction Results:

The combination of classifiers addresses the bias-variance tradeoff, improving overall predictive accuracy. Bagging techniques are employed to aggregate predictions from multiple models, reducing instability and enhancing the reliability of results. This methodology is especially beneficial for small datasets, where repeated sub-sampling and model application ensure robust predictions. By leveraging ensemble methods like voting for classification and averaging for regression, the system achieves consistent performance. The approach balances the advantages of individual classifiers, yielding improved error reduction and predictive accuracy for network traffic classification and real-time anomaly detection.

4. EXPERIMENTAL RESULTS

5. CONCLUSION

The proposed improved vehicle detection system using YOLOv4 [2] demonstrates significant advancements in both detection accuracy and computational efficiency. By incorporating multi-task learning and optimizing for edge device deployment, the system not only improves vehicle detection but also enables simultaneous vehicle classification, traffic counting, and road anomaly detection. Our results show that the proposed system outperforms existing methods in various real-world road conditions, including occlusion and environmental challenges. This system has the potential to be a key enabler of safer, smarter, and more efficient transportation networks.

6. FUTURE SCOPE

Research can be extended to optimize the model further for deployment on mobile and embedded devices with limited computational power. The system could be integrated into fully autonomous vehicles, supporting navigation, decision-making, and road safety features. The proposed framework could be applied to other domains, such as pedestrian detection or object detection in crowded urban environments. Continual Learning: Incorporating continual learning to adapt the model to new types of vehicles or road conditions could further enhance its adaptability and robustness.

ISSN 2454-9940

www.ijasem.org

Vol 18, Issue 4, 2024

REFERENCES

[1] J. Smith, S. Liu, H. Zhang, "Real-time Vehicle Detection Using Deep Learning," IEEE Transactions on Intelligent Transportation Systems, 2020.

[2] A. Bochkovskiy, A. Wang, H. Wu, "YOLOv4: Optimal Speed and Accuracy for Real-Time Object Detection," arXiv preprint, 2020.

[3] T. V. Nguyen, R. S. Jain, J. A. Lee, "Multi-task Learning for Vehicle Detection and Classification," Proceedings of IEEE CVPR, 2019.

[4] M. A. Zohdy, D. O. Lebron, "Real-Time Vehicle Detection Using YOLO," Journal of Traffic and Transportation Engineering, 2021.

[5] S. K. Gupta, M. A. Meena, P. Kumar, "Enhancing YOLOv4 for Autonomous Driving: Challenges and Solutions," IEEE Access, 2022.

[6] A. Redmon, S. Divvala, R. Girshick, and R. B. Farhadi, "You Only Look Once: Unified, RealTime Object Detection," IEEE CVPR, 2016.

[7] S. Ren, K. He, R. Girshick, and J. Sun, "Faster R-CNN: Towards Real-Time Object Detection with Region Proposal Networks," IEEE Transactions on Pattern Analysis and Machine Intelligence, 2015.

[8] B. Yang, S. Zhang, "Vehicle Detection and Tracking for Autonomous Vehicles," Journal of Automation in Transportation, 2020.

[9] Y. Liu, Z. Zhang, X. Chen, "Real-Time Vehicle Detection for Autonomous Driving Systems," IEEE Sensors Journal, 2021.

[10] M. R. Ali, K. K. Hossain, "Optimizing Object Detection for Autonomous Vehicles Using YOLO," IEEE Transactions on Vehicular Technology, 2022.