International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

Vehicle Detection and Tracking with Deep Learning

Vikram Hegde[1], Patel Suhel Riyaj[2], Mouna V J[3], S J Yasshraj Vijayan[4]
Prof Nidhishree G[5]

[1,2,3,4]BE Student Dept of Artificial Intelligence and Machine Learning Vidyavardhaka College of Engineering Gokulum, Mysore
[5]Assistant Professor Dept of Artificial Intelligence and Machine Learning idyavardhaka College of Engineering, Gokulum, Mysore

Abstract
This project employs YOLOv8, a state-of-the-art object detection model, to detect and track vehicles in real-time for effective traffic management. By processing road video footage, the system addresses challenges like monitoring congestion and identifying accident risks, offering an efficient, scalable solution that outperforms traditional methods, which are often costly and limited in coverage.
With YOLOv8's detection capabilities integrated into a tracking algorithm, the system continuously monitors traffic flow, vehicle density, and unusual behaviour's on the road. This approach provides real-time insights that can support smart city infrastructure by enhancing road safety and optimizing traffic. Ultimately, this project demonstrates the transformative potential of deep learning in urban traffic analysis, contributing to safer and more efficient transportation systems.

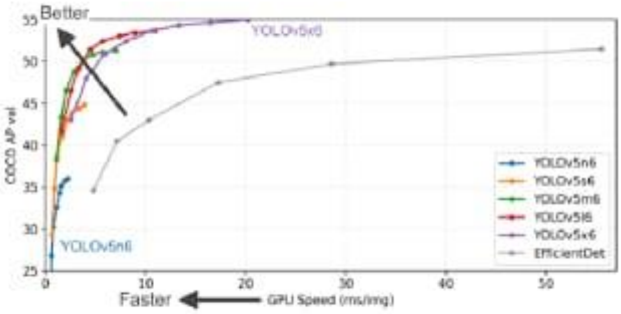
Keywords: Machine Learning, Traffic Management, YOLO Algorithm, Random Forest, Traffic Congestion, Real-Time Detection, Deep Learning, Smart Cities, Image Processing, UAV.

1. INTRODUCTION

Traffic congestion and road safety are major concerns in urban areas, leading to delays, increased emissions, and higher accident risks. Traditional traffic monitoring techniques, such as loop detectors and radar, are costly, limited in coverage, and unable to provide comprehensive, real-time data. Recent advances in deep learning have introduced powerful tools for automated traffic analysis, offering more efficient, scalable solutions that can transform how we manage and understand urban traffic.
Deep learning models, particularly YOLOv8, have shown remarkable effectiveness in object detection tasks. YOLOv8's architecture allows it to process entire images at once, enabling real-time vehicle detection and tracking. By leveraging deep learning's ability to recognize complex patterns in video footage, this project can identify vehicle density, detect accidents, and track movement patterns with high accuracy. This continuous, automated monitoring enables us to address congestion and improve road safety without the need for extensive physical infrastructure.
This project integrates deep learning for real-time traffic analysis, combining YOLOv8's detection capabilities with tracking algorithms to monitor vehicle flow and behaviour's effectively. Through this approach, we aim to contribute to smarter urban transportation and provide valuable insights for traffic

International Journal for Multidisciplinary Research (IJFMR)
E-ISSN: 2582-2160 ● Website: www.ijfmr.com ● Email: editor@ijfmr.com

optimization. This work demonstrates how deep learning can support the development of scalable, AI-driven solutions for urban infrastructure, paving the way for safer and more efficient cities..
In conclusion, leveraging deep learning for traffic analysis offers a scalable, real-time solution to urban traffic challenges, reducing reliance on costly infrastructure. By utilizing YOLOv8 for accurate vehicle detection and tracking, this approach not only improves congestion management and road safety but also lays the foundation for integrating advanced technologies into smart city ecosystems. As deep learning algorithms continue to evolve, the potential for further innovation in traffic optimization andsustainability remains vast, promising a smarter and greener urban future.



2. LITERATURE REVIEW

The rapid advancements in Intelligent Transportation Systems (ITS) have made vehicle detection and tracking an essential field of study. Different methodologies have been explored to improve real-time vehicle detection accuracy under diverse environmental conditions. Traditional techniques like background subtraction and frame differencing form the basis of vehicle detection systems. Background subtraction segments moving objects by comparing each frame against a static background, but this method struggles with changing lighting conditions and background clutter [1]. Frame differencing is another widely used method, particularly effective in detecting objects in dynamic scenes, yet it lacks the robustness required for consistently accurate detection when objects stop moving.
In recent years, deep learning techniques have made significant contributions to vehicle detection and tracking. A notable study utilized YOLOv4 (You Only Look Once, Version 4) alongside SORT (Simple Online and Real-time Tracking) to analyse low-quality CCTV footage for traffic flow estimation. The YOLO framework is known for its ability to detect objects in a single processing pass, which enhances speed, a critical factor in real-time applications. SORT aids YOLO by providing reliable tracking over multiple frames, even in complex traffic environments. However, challenges such as handling partial occlusions and accurate classification across varying vehicle types and lighting conditions still exist in this system [2].
Object detection techniques in surveillance systems have traditionally relied on a mix of statistical and optical flow methods to handle the limitations of simple background subtraction. Statistical models, for instance, attempt to adapt the background dynamically based on changes in pixel values, which can better handle variations like weather changes or shadows. Despite this, statistical methods can still struggle with sudden lighting shifts or when objects have similar colours to the background, leading to false detections [3]. Optical flow methods, another approach, estimate motion by analysing changes in pixel intensity across frames, making it effective in scenarios where the camera is moving. However, optical flow is computationally intensive and tends to generate noise, which can complicate real-time detection

in large-scale traffic monitoring systems [4].

In addition to these advanced techniques, network-based traffic analysis has been explored for ITS applications. This approach captures and examines data flow across networked sensors or CCTV systems, allowing for detailed analysis of congestion patterns, vehicle counts, and incident detection. While originally focused on cybersecurity and intrusion detection, traffic analysis methodologies are now being adapted for physical traffic flow monitoring. However, these methods require extensive data processing power and high-speed internet to achieve the necessary real-time response, which can be a limitation in areas with restricted infrastructure [3].

Some of the recent studies have focused on improving vehicle classification accuracy to enhance ITS capabilities. The classification step typically distinguishes between object types, such as cars, trucks, and buses, using features like size, speed, and shape. YOLO's real-time capability has proven valuable here, but challenges persist in accurately classifying small or partially occluded vehicles. Variants like Faster RCNN, while slower, offer higher accuracy by using a twostage detection process that first proposes regions of interest and then classifies objects, but the added processing time can limit its suitability for real-time applications [1][5].

Lastly, limitations remain in addressing adverse weather and lighting conditions. Studies have shown that models trained with data across multiple environmental conditions, such as night and rainy scenes, generally perform better. However, the resource requirements for gathering and processing this type of data, along with training deep learning models capable of handling such diverse input, remain a major challenge in real-time traffic surveillance [2].

In the approach presented by Krishna et al. [6], the Random Forest algorithm is utilized for predictive traffic analysis in smart cities. This machine learning model works by constructing multiple decision trees based on traffic data collected from sensors embedded in road networks. The system employs two-thirds of the dataset to train and predict congestion patterns, making it feasible for devices with limited computational capabilities. A key advantage of this method is its ability to predict traffic density and notify users well in advance, which can help alleviate congestion in critical areas. However, a limitation of this technique lies in its reliance on historical data, which may not fully capture rapid changes in traffic flow due to unforeseen events. To address these constraints, the authors suggest future work that explores hybrid algorithms to improve model adaptability and accuracy in diverse real-time scenarios [6].

In the study by Kunekar et al. [7], the YOLOv7 deep learning algorithm is leveraged for real-time vehicle detection and traffic signal adjustment at intersections. YOLO's convolutional neural network-based design allows it to process video frames from traffic cameras rapidly, identifying vehicle types and adjusting traffic signals based on density and vehicle type. This approach significantly improves traffic flow at peak times by dynamically assigning green light durations. Despite its high accuracy, the model can struggle with accuracy when detecting vehicles in lowvisibility or adverse weather conditions. Furthermore, YOLOv7's computational demands make it challenging to deploy on devices with limited processing power, which limits its widespread application in resource-constrained environments. Future research may focus on optimizing YOLO for use on edge devices or exploring more lightweight versions of the model, which could expand its applicability across various urban settings [7].

The paper by Niyazahamad et al. [8] introduces an approach that combines UAVs and deep learning, particularly convolutional neural networks (CNNs), for traffic monitoring in wide, densely populated areas. This model addresses the challenges of real-time vehicle tracking from high altitudes by employing pattern recognition and objecttracking algorithms tailored for aerial image analysis. Such a setup is advantageous for overseeing large areas and enabling rapid congestion detection and management. One of the main limitations, however, is its dependence on optimal lighting and weather conditions, as these directly affect the accuracy of the CNN in recognizing vehicles. In lowvisibility scenarios, such as nighttime or foggy weather, the performance can be inconsistent. Further research in this field could aim to improve the CNN's resilience under challenging conditions, potentially through hybrid models or enhanced aerial imaging capabilities [8].

In another study by Mounica et al. [9], image processing techniques are used to analyze traffic flow based on realtime video data from road cameras. The authors apply Convolutional Neural Networks (CNNs) to identify specific patterns, such as accident hotspots and potential violations, helping authorities monitor road conditions and make datainformed decisions to ease congestion. The CNN model used for analysis is adept at recognizing various traffic events and relaying real-time updates, which can support faster emergency response and reduce delays. However, this approach relies on a continuous and high-quality video feed, which can be disrupted due to camera malfunctions or poor weather. In addition, CNN models typically require significant computational resources, which could hinder deployment in some areas. The authors propose that future models might integrate machine learning with additional data sources, such as historical traffic data, to better anticipate traffic trends and handle disruptions more effectively [9].

3. METHODOLOGY

Based on the literature, a general methodology for a vehicle detection and tracking system can be outlined as follows:

1. Data Collection:

Use publicly available datasets or collect video footage from road surveillance cameras to capture various scenarios (e.g., different lighting, weather conditions).

2. Data Preparation:

Annotate the collected data with bounding boxes around vehicles to facilitate model training. This process often involves using annotation tools for accurate labelling.

3. Model Selection:

For vehicle detection, YOLO (You Only Look Once) is chosen for its speed and accuracy in real-time object detection. YOLOv8, the latest version, offers improvements in processing and detection accuracy.

For tracking, methods like Kalman filtering or optical flow may be employed to track vehicles across frames after detection.

4. Model Training:

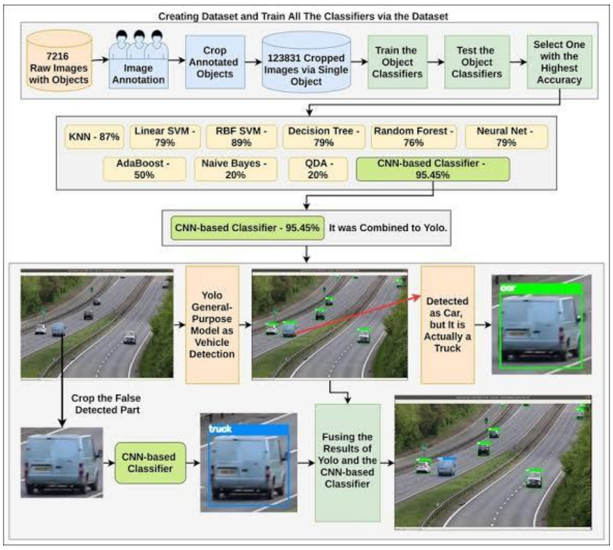
Train the YOLO model on the annotated dataset, optimizing hyperparameters to minimize detection errors. The training process involves feeding the model with images, updating weights based on error rates, and refining detection through iterative training.

5. Model Evaluation:

Evaluate the model using precision, recall, and F1-score to determine detection and tracking accuracy. Additional metrics like processing speed and memory usage are also important for real-time applications.

6. Real-Time Implementation:

Deploy the trained model on a system equipped with real-time video input. The model continuously detects vehicles and tracks their positions across frames, providing outputs such as vehicle count, speed, and movement patterns.



3. DISCUSSION

The integration of YOLOv8 for real-time vehicle detection and tracking represents a significant step toward improving urban traffic management. Unlike traditional traffic monitoring methods, which are costly and limited in scope, this approach leverages deep learning to deliver a scalable, efficient solution that adapts to complex urban environments. YOLOv8's real-time capabilities make it wellsuited for capturing high-density traffic scenes, where quick detection and response are critical for applications such as congestion management and accident detection.

One of the key advantages of YOLOv8 is its ability to balance accuracy and speed, even in challenging conditions. Urban traffic environments often present variables such as variable lighting, occlusions, and heavy congestion. YOLOv8, with its advanced architecture and real-time processing abilities, is particularly adept at handling these complexities, making it a valuable tool for traffic management systems aiming to provide reliable and timely data. This capability is enhanced further by integrating tracking algorithms like SORT or Deep SORT, which maintain consistent vehicle tracking across frames, essential for reliable congestion analysis and accident detection.

The real-time tracking of vehicles also enables a more dynamic understanding of traffic flow, capturing density changes, movement patterns, and potential bottlenecks. This real-time data can inform traffic signal adjustments, rerouting, and even road design decisions in the long term. In particular, the system's ability to detect accidents through sudden vehicle behaviour changes could allow for immediate response measures, potentially reducing secondary accidents and congestion.

Beyond immediate applications, this approach supports broader smart city initiatives. Traffic data gathered from YOLOv8 and related tracking systems could be fed into a larger network of IoT-enabled devices, contributing to adaptive traffic systems that respond to real-time conditions. By providing a continuous, automated monitoring solution, deep learning models like YOLOv8 facilitate a shift toward proactive, AI-driven urban infrastructure management.

In summary, the use of YOLOv8 and associated tracking technologies offers a practical, scalable solution for addressing the complex demands of urban traffic management. By enabling real-time, accurate detection and tracking, this approach not only improves traffic efficiency and safety but also demonstrates the transformative potential of AI in urban planning and management. This project exemplifies how advancements in deep learning can support smarter, more efficient, and safer cities, ultimately benefiting both commuters and the environment.

4. CONCLUSION

After reviewing the various methods and algorithms used in vehicle detection and tracking, it is evident that deep learning approaches, especially YOLO-based models, offer substantial advantages in terms of speed and real-time accuracy. YOLO's single-stage detection framework is highly efficient, making it particularly suitable for continuous video monitoring, even under complex traffic conditions. Given the challenges in handling diverse lighting, occlusions, and varying vehicle types, YOLOv8 emerges as an optimal choice due to its enhanced accuracy and processing capabilities compared to previous versions. Based on these findings, we believe that implementing YOLOv8 will provide the most robust and efficient solution for our vehicle detection and tracking project. We aim to leverage YOLOv8's advanced detection performance to build a system capable of reliable, real-time traffic analysis.

REFERENCES

1. Sri Jamiya S, Esther Rani P. A Survey on Vehicle Detection and Tracking Algorithms in Real-Time Video Surveillance.
2. Nilani Algiriyage et al. Towards Real-time Traffic Flow Estimation using YOLO and SORT from Surveillance Video Footage.
3. Saurav Mishra, Manjusha Pandey, Gargi Srivastava. A Survey on Traffic Analysis in Networks.
4. Swaranjali Dilip Jadhav, Ankita Anil Choudhari, Varsha Subhash Chavan, Gowri Mane, Amol Hindurao Pudale. Vehicle Detection and Tracking System.
5. Sri Jamiya S, Esther Rani P. A Survey on Vehicle Detection and Tracking Algorithms in Real-Time Video Surveillance.
6. Krishna, A., et al. Traffic Analysis Using Machine Learning in Smart Cities. International Journal of Engineering and Advanced Technology (IJEAT)
7. Kunekar, P., et al. Traffic Management System Using YOLO Algorithm. Engineering Proceedings.
8. Niyazahamad S K, et al. Vehicle Detection and Tracking in Aerial Images Using Deep Learning Techniques. International Research Journal of Engineering and Technology (IRJET).
9. Mounica B, et al. Traffic Analysis Using Image Processing. International Journal of Computer Science and Mobile Computing (IJCSMC).
10. 3. Saurav Mishra, Manjusha Pandey, Gargi
11. Srivastava. A Survey on Traffic Analysis in Networks.