International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

A HYBRID ML–DL FRAMEWORK FOR REAL-TIME VEHICLE DETECTION, CLASSIFICATION, AND TRACKING IN INTELLIGENT TRAFFIC SURVEILLANCE

Riddhi Mehta[1]* , Dr. Ankit Shah[2]

[1] Research Scholar, Faculty of Engineering and Technology, Sankalchand Patel University, Visnagar, Gujarat, India

E-mail: mehta.riddhi85@gmail.com

[2] Professor, Adani Institute of Digital Technology Management, Gandhinagar, Gujarat, India

E-mail:shah_ankit101@yahoo.co.in

Abstract

The exponential rise in urban traffic has created a pressing demand for intelligent surveillance systems capable of real-time vehicle monitoring, detection, and analysis. Conventional image processing approaches often struggle to maintain accuracy in complex scenarios involving illumination changes, occlusions, and dense traffic conditions. To overcome these limitations, this study proposes a hybrid Machine Learning and Deep Learning (ML/DL) framework for efficient vehicle detection, classification, and tracking. The system employs a YOLO-based deep learning model for precise object localization, a Convolutional Neural Network (CNN) for vehicle type classification, and a multi-object tracking algorithm to ensure consistent tracking across consecutive frames. Experimental evaluations indicate that the proposed method achieves superior accuracy, robustness, and adaptability compared to traditional techniques. The outcomes of this research support the advancement of intelligent traffic surveillance systems, contributing to improved traffic management, law enforcement, and urban mobility optimization.

Math. Subject Classification: 68T07, 68T45, 68T05, 68T10, 68U10

Keywords: Surveillance, Vehicle Classification, YOLO, Machine Learning, Deep Learning, Tracking.

1. Introduction

Computer vision has become a key field in artificial intelligence, enabling applications in robotics, automation, motion analysis, and human–machine interaction [1]. Vehicle detection and recognition, a critical aspect of computer vision, are essential for intelligent transportation systems (ITS), traffic monitoring, and urban management [2]. Automated systems for detecting, classifying, and tracking vehicles address the limitations of manual monitoring, providing solutions for efficient traffic flow, safety, and law enforcement [3].

Deep learning models, particularly the YOLO (You Only Look Once) series, have proven highly effective for real-time vehicle detection [4]. YOLOv5 and YOLOv8 offer fast, accurate single-shot detection, while tracking algorithms like DeepSORT maintain vehicle identities across consecutive frames [5]. This integration enables classification of vehicles

Received: August 10, 2025

into categories such as cars, buses, trucks, and motorcycles, supporting applications including congestion analysis, accident prevention, and toll management [6].

Challenges such as occlusion, scale variations, poor lighting, and low-resolution footage continue to affect detection accuracy [7]. This study proposes a vehicle surveillance framework combining YOLOv8 and tracking algorithms to achieve reliable detection, classification, and monitoring in real-world traffic scenarios [8-10]. Experimental results on real-time traffic videos demonstrate the system's accuracy, robustness, and computational efficiency, highlighting its potential to improve traffic management, urban safety, and smart city initiatives [11].

2. Methodology

The effectiveness of the suggested YOLOv8 for vehicle detection and categorisation in experiments is assessed in this section. After detection and classification, using OCR algorithm we detect and read vehicle number plate which stored into CSV file with the details of vehicle. to track the vehicle using Google API which shows the travelling area covered by vehicle over the Google map, we used this CSV file to read the number plate and latitude and longitude. The descriptions of the dataset are shown first. Next, the network architecture was used to describe the algorithm. The outcome of the vehicle detection and classification process is finally shown.

2.1 Datasets

Two primary datasets were employed in this study to evaluate the proposed YOLOv8-based detection and tracking system: the SPU Parking Video Dataset and the Mehsana Traffic Video Dataset. Both datasets were curated to include a wide range of real-world conditions, ensuring that the trained model generalizes effectively across different environments.

(a) SPU parking Video Dataset:

This dataset was captured at the Sankalchand Patel University parking area using a mobile camera to record vehicles during entry and exit under steady lighting. It contains approximately 300 frames standardized to 640×640 pixels, covering various vehicle types such as cars, two-wheelers, and small trucks. Annotations were created manually using the LabelImg tool in YOLO format, and 20% of samples were revalidated by independent reviewers to ensure accuracy. The dataset represents daytime conditions with moderate shadows, ideal for evaluating detection precision in controlled environments.

(b) Mehsana Traffic Video Dataset:

This dataset was recorded using a fixed roadside camera in Mehsana, Gujarat, to capture multi-lane urban traffic. It includes around 7,800 frames resized to 640×640 pixels and features diverse categories such as cars, buses, trucks, autorickshaws, motorcycles, and pedestrians. The dataset covers varying traffic densities, illumination levels, and mild weather changes. About 22% of frames contain occlusions, offering realistic challenges for detection and tracking tasks. Bounding boxes were annotated in Roboflow using YOLOv8[17] format, verified by multiple annotators with an IoU threshold of 0.85 to maintain labeling consistency.

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

2.2 YOLO v8 Algorithm

YOLO (You Only Look Once) is a single-stage detection model that simultaneously performs object localization and classification, removing the need for region proposal networks. It divides the input image into an S×S grid, where each cell predicts bounding boxes with confidence and class probabilities. Non-Maximum Suppression (NMS) is then used to eliminate redundant detections. In this research, YOLOv8 is utilized for vehicle detection because of its superior accuracy, real-time speed, and resilience in dense or varying traffic conditions. The model employs a CSPDarknet-inspired backbone with C2f modules for efficient feature extraction, an FPN-PAN neck for multi-scale fusion, and a decoupled head that separates classification and localization. Its anchor-free architecture reduces computation while maintaining precision, making YOLOv8 a lightweight, high-performance solution for real-time vehicle detection and tracking in highway surveillance systems.

2.3 Tracking Module

To maintain consistent identities of detected vehicles across consecutive frames, the proposed system employs a hybrid tracking mechanism that integrates Kalman filtering for motion prediction and Intersection-over-Union (IOU) matching for data association. This combination allows the system to achieve reliable multi-object tracking, even under partial occlusions or temporary detection failures.

(a) Kalman Filter Model

The Kalman filter predicts the future position and velocity of each detected vehicle based on a linear state-space model. Each tracked vehicle is represented by a state vector:

$x_k=[u,v,s,r,\dot{u},\dot{v},\dot{s}]^T$

where

u,v represent the bounding box center coordinates,

s denotes the scale (area of the bounding box),

r is the aspect ratio,

$\dot{u},\dot{v},\dot{s}$ correspond to the respective velocities of these quantities.

The state transition model assumes constant velocity motion between frames and is defined as:

$x_k = F\, x_{(k-1)} + w_k$

where F is the transition matrix, and wk represents process noise modeled as Gaussian with zero mean.

The measurement model captures observations from YOLOv8 detections:

$z_k = H\, x_k + v_k$

where zk=[u,v,s,r]T and vk denotes measurement noise.

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The Kalman filter performs two main operations per frame:

- Prediction: Estimate the vehicle’s next position and velocity using motion dynamics.

- Correction: Update the estimate with the new detection that best matches the predicted state.

This probabilistic filtering framework effectively smooths trajectory fluctuations and compensates for short-term detection dropouts, enabling continuity of tracking even in brief occlusion periods.

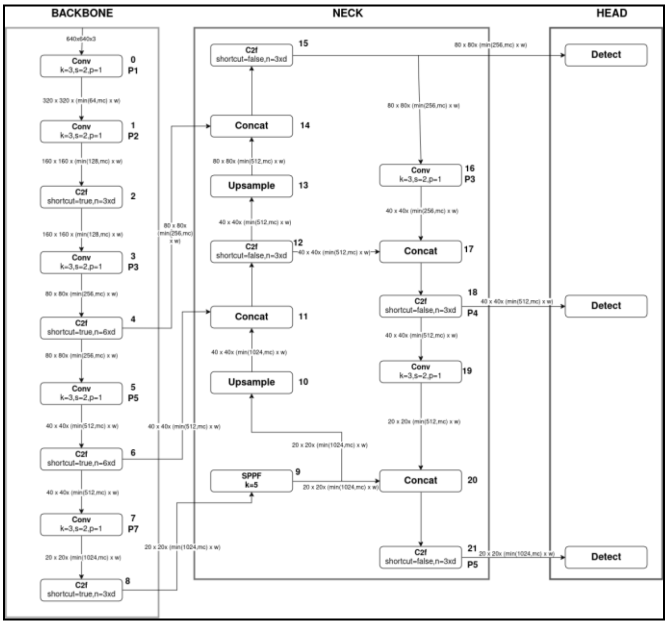


Fig. 1. YOLO v8 Architecture Design [11]

(b) Occlusion Handling

During brief occlusions or missed detections, the Kalman filter predicts the vehicle’s next position using motion estimation, while the IOU-based association re-links detections once visibility is restored. For longer occlusions or sudden motion shifts, an adaptive filtering strategy—based on Cai et al. [27]—dynamically adjusts process noise to maintain stable tracking under rapid movement and complex traffic conditions.

(c) Tracking Output

Each active vehicle track is recorded in a structured CSV file containing bounding box coordinates, class label, detected license plate (via OCR), confidence score, timestamp, and geolocation (latitude, longitude). These records are later utilized for route reconstruction and Google Maps visualization, allowing real-time monitoring and analysis of vehicle movement patterns.

3. Proposed Model

3.1 System workflow

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The process flow illustrated in the diagram represents a complete pipeline for automated vehicle detection and tracking from video input. The system begins by acquiring the video and performing **preprocessing** to enhance frame quality. The video is then divided into individual frames for **data acquisition**, followed by **background image generation** to distinguish moving vehicles from static elements. **Object detection** identifies vehicles within each frame, while **license plate detection** isolates the region containing the plate. After detecting the plate, **character segmentation** separates individual characters, and **text extraction** retrieves the alphanumeric details using OCR techniques. These recognized plate numbers are then **stored in a CSV file** for recordkeeping. The system also performs **vehicle tracking** by matching detected vehicles across consecutive frames, generating continuous movement data. Finally, **vehicle details** are retrieved and correlated with stored records, enabling comprehensive monitoring and analysis of traffic activity.

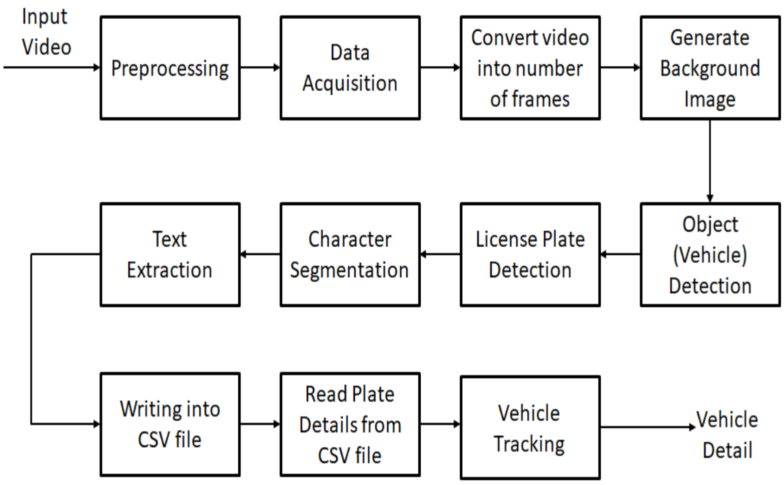


Fig. 2. Proposed Model Workflow

Fig. 3.Vehicle Recognition, Categorisation and number plate detection from a Dataset of SPU Parking Videos

Fig. 4. Vehicle Recognition, Categorisation and number plate detection from a Dataset of Mehsana Traffic Videos

Database and cross-camera tracking

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

Detection results are stored in a structured csv database, containing details such as timestamp, bounding box coordinates, class label, plate number, and camera location. By cross-referencing license plate data across multiple camera streams, the system supports multi-camera vehicle tracking.

Geospatial visualization

For route analysis, the system integrates with the google maps API. Each vehicle's trajectory is plotted based on its detections across cameras, allowing real-time visualization of movement patterns.

4. Results

We used two datasets for the experiment, which include both photographs and videos of vehicles, and our suggested algorithm was able to identify them. The experimental outcome is displayed in Figures 2 and 3. Here, we classified and detected vehicles from both video and image datasets using the YOLO v8 method.

In Figures 3 and 4, the proposed object detection system demonstrates its ability to automatically identify and classify multiple vehicles within parking and roadway environments. Each detected vehicle is highlighted using color-coded bounding boxes, where trucks are outlined in green with a confidence score of 0.453, while cars and buses are enclosed in orange with confidence values ranging from 0.731 to 0.834 for cars and 0.430 for buses. These confidence values represent the model's level of certainty in classification, with higher scores indicating stronger reliability. Such automated detection not only facilitates accurate categorization of vehicles but also serves as a foundation for traffic flow analysis, congestion monitoring, and road safety applications.

Figure 5 illustrates the statistical distribution of detection confidence scores obtained from the two evaluated datasets. As shown in Figure 5(a), the Traffic Videos Dataset exhibits a bimodal distribution, with peaks around 0.3 and 0.8.

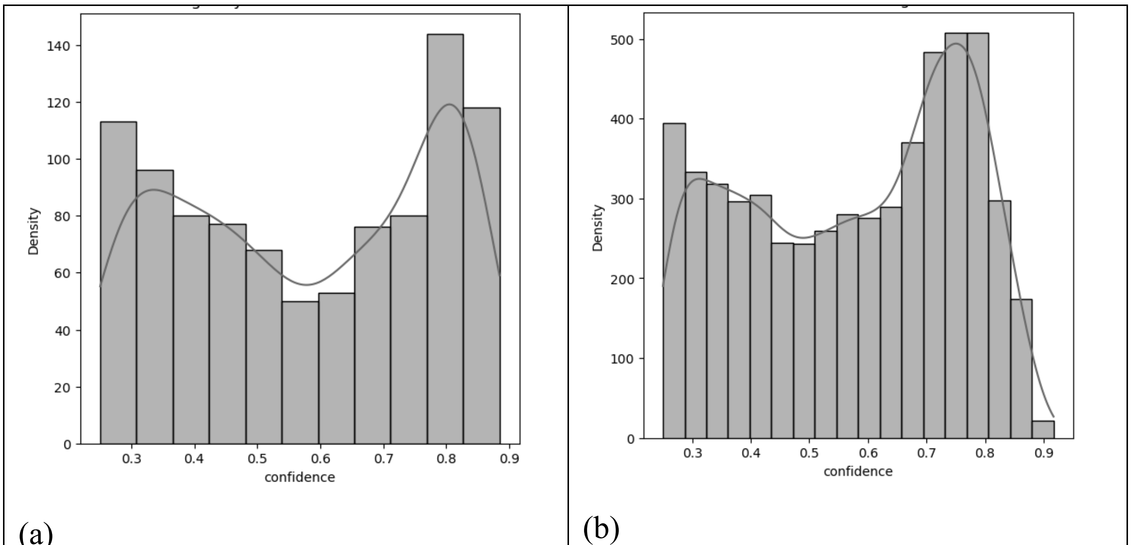


Fig. 5. Confidence score distribution for vehicle detection in (a) Traffic Videos Dataset and (b) Parking Video Dataset.

Received: August 10, 2025

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

The bimodal peaks at 0.3 and 0.8 indicate YOLOv8's ability to detect vehicles effectively under both low and high visibility conditions. Lower peaks arise from occlusion, blur, or poor lighting, while higher peaks correspond to clear, well-lit scenes, proving the model's adaptability. In Figure 5(b), the Parking Video Dataset shows a tighter confidence range (0.7–0.9), reflecting consistent detection in stable environments. As shown in Table 1, trucks, autos, and buses in the Traffic Videos dataset achieved confidence scores of 0.45, 0.83, and 0.43, while cars and trucks in the Parking Video dataset scored 0.82 and 0.86. This indicates stronger truck detection in the Parking dataset and slightly better car detection in traffic videos. Table 2 results further confirm YOLOv8's superior precision, recall, F1-score, and mAP, validating its robust performance across both datasets.

Table 1. Confidence Scores for Vehicle Detection

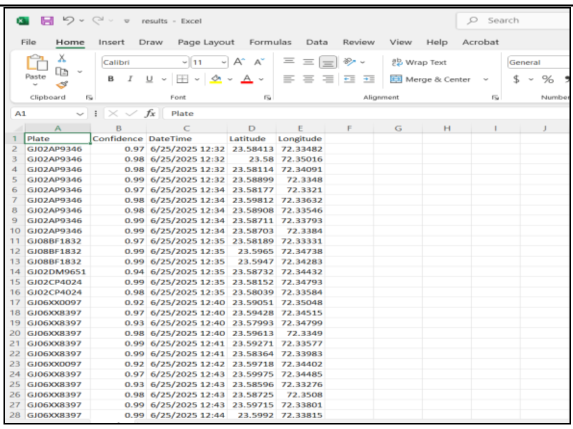
Dataset	Vehicle Type	Confidence Score
Mehsana Traffic Video Dataset	Trucks	0.45
	Cars	0.83
	Buses	0.43
Parking Video	Cars	0.82
	Trucks	0.86

Table 2. Detection Performance Metrics for YOLOv8

Dataset	Precision	Recall	F1-Score	mAP
Mehsana Traffic Video Dataset	0.91	0.88	0.89	0.93
Parking Video	0.94	0.90	0.92	0.95

These results show that YOLOv8 achieves high precision (0.91–0.94) and recall (0.88–0.90), indicating both low false positives and low false negatives. The overall mAP above 0.93 highlights its strong localization and classification performance, even in moderate occlusion and low-resolution conditions.

International Journal of Applied Mathematics
Volume 38 No. 10s, 2025
ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)



	A	B	C	D	E
1	Plate	Confidence	DateTime	Latitude	Longitude
2	GJ02AP9346	0.97	6/25/2025 12:32	23.58413	72.33482
3	GJ02AP9346	0.98	6/25/2025 12:32	23.58	72.35016
4	GJ02AP9346	0.98	6/25/2025 12:32	23.58114	72.34091
5	GJ02AP9346	0.99	6/25/2025 12:32	23.58899	72.3348
6	GJ02AP9346	0.97	6/25/2025 12:34	23.58177	72.3321
7	GJ02AP9346	0.98	6/25/2025 12:34	23.59812	72.33632
8	GJ02AP9346	0.98	6/25/2025 12:34	23.58908	72.33546
9	GJ02AP9346	0.99	6/25/2025 12:34	23.58711	72.33793
10	GJ02AP9346	0.99	6/25/2025 12:34	23.58703	72.3384
11	GJ08BF1832	0.97	6/25/2025 12:35	23.58189	72.33331
12	GJ08BF1832	0.99	6/25/2025 12:35	23.5965	72.34738
13	GJ08BF1832	0.99	6/25/2025 12:35	23.5947	72.34283
14	GJ02DM9651	0.94	6/25/2025 12:35	23.58732	72.34432
15	GJ02CP4024	0.99	6/25/2025 12:35	23.58152	72.34793
16	GJ02CP4024	0.98	6/25/2025 12:35	23.58039	72.33584
17	GJ06XX0097	0.92	6/25/2025 12:40	23.59051	72.35048
18	GJ06XX8397	0.97	6/25/2025 12:40	23.59428	72.34515
19	GJ06XX8397	0.93	6/25/2025 12:40	23.57993	72.34799
20	GJ06XX8397	0.98	6/25/2025 12:40	23.59613	72.3349
21	GJ06XX8397	0.99	6/25/2025 12:41	23.59271	72.33577
22	GJ06XX8397	0.99	6/25/2025 12:41	23.58364	72.33983
23	GJ06XX0097	0.92	6/25/2025 12:42	23.59718	72.34402
24	GJ06XX8397	0.97	6/25/2025 12:43	23.59975	72.34485
25	GJ06XX8397	0.93	6/25/2025 12:43	23.58596	72.33276
26	GJ06XX8397	0.98	6/25/2025 12:43	23.58725	72.3508
27	GJ06XX8397	0.99	6/25/2025 12:43	23.59715	72.33801
28	GJ06XX8397	0.99	6/25/2025 12:44	23.5992	72.33815

Fig. 6. CSV File details

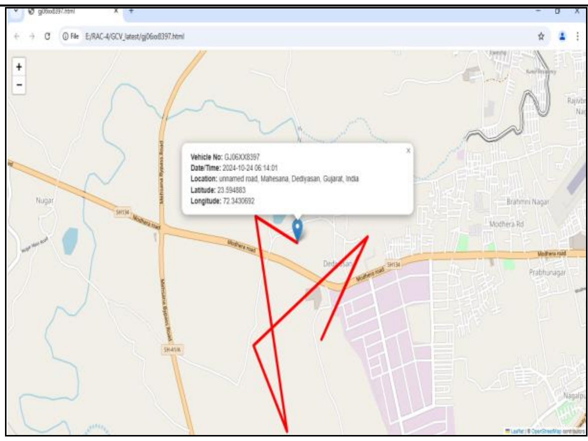


Fig. 7. Vehicle Tracking from CSV file

In Figure 6, the extracted license plate information is systematically stored in a CSV file along with key attributes such as detection confidence score, geolocation coordinates (latitude and longitude), and timestamp. To ensure the reliability of the recognition stage, the system validates detected plate numbers against predefined Indian vehicle registration formats using regular expressions. Special care is taken to distinguish visually similar characters, such as the digit '0' and the letter 'O', to minimize recognition errors. Once validated, the stored geospatial data is utilized for vehicle tracking through the Google Maps API, enabling real-time route visualization based on the recorded positions, which shown in the Figure 7.

5. Conclusion

This research presents an integrated vehicle surveillance framework that unifies detection, license plate recognition, and geospatial tracking within a single automated system. Using YOLOv8 for real-time vehicle identification and OCR-based number plate extraction, the method stores each detection with corresponding time and location data to enable multi-camera tracking. Integration with Google Maps API allows dynamic visualization of vehicle routes, converting detection data into actionable traffic insights. Experimental analysis demonstrated consistent performance across diverse lighting and occlusion scenarios, with validation against Indian plate formats enhancing recognition reliability. Overall, the system delivers a scalable and efficient solution for intelligent traffic monitoring, supporting real-time observation, post-incident analysis, and potential future extensions such as edge-based processing, predictive analytics, and privacy-aware deployment.

References

[1] Mehta, Riddhi, and Ankit Shah, Real-Time Vehicle Detection and Classification Using Deep Learning based Approach,2024 Dec 10;(18s):284-289.

[2] Arora, S., Mittal, R., Arora, D., & Shrivastava, A. K. (2024). A Robust Approach for Licence Plate Detection Using Deep Learning. Inteligencia Artificial, 27(73), 129-141.

[3] Singh, N., Saini, P., Shubham, O., Awasthi, R., Bharti, A., & Kumar, N. (2024). Improved YOLOv5l for vehicle detection: an application to estimating traffic density and identifying over speeding vehicles on highway scenes. Multimedia Tools and Applications, 83(2), 5277-5307.

Received: August 10, 2025

ISSN: 1311-1728 (printed version); ISSN: 1314-8060 (on-line version)

[4] Wieczorek G, Tahir SB, Akhter I, Kurek J. Vehicle detection and recognition approach in multi-scale traffic monitoring system via graph-based data optimization. Sensors. 2023 Feb 3;23(3):1731.

[5] Ge DY, Yao XF, Xiang WJ, Chen YP. Vehicle detection and tracking based on video image processing in intelligent transportation system. Neural Computing and Applications. 2023 Jan;35(3):2197-209.

[6] Wang J, Dong Y, Zhao S, Zhang Z. A high-precision vehicle detection and tracking method based on the attention mechanism. Sensors. 2023 Jan 8;23(2):724.

[7] Ali S, Jalal A, Alatiyyah MH, Alnowaiser K, Park J. Vehicle detection and tracking in UAV imagery via YOLOv3 and Kalman filter. InCMC 2023 Jun 9.

[8] Basak S, Suresh S. Vehicle detection and type classification in low resolution congested traffic scenes using image super resolution. Multimedia Tools and Applications. 2024 Mar;83(8):21825-47.

[9] Farid A, Hussain F, Khan K, Shahzad M, Khan U, Mahmood Z. A fast and accurate real-time vehicle detection method using deep learning for unconstrained environments. Applied sciences. 2023 Feb 27;13(5):3059.

[10] Li S, Chen J, Peng W, Shi X, Bu W. A vehicle detection method based on disparity segmentation. Multimedia Tools and Applications. 2023 May;82(13):19643-55.

[11] Mehta R, Shah A. An Insight into Real Time Vehicle Detection and Classification Methods using ML/DL based Approach. Procedia Computer Science. 2024 Jan 1;235:598-605.

[12] Zhang Y, Guo Z, Wu J, Tian Y, Tang H, Guo X. Real-time vehicle detection based on improved yolo v5. Sustainability. 2022 Sep 27;14(19):12274.

[13] Olorunshola OE, Irhebhude ME, Evwiekpaefe AE. A comparative study of YOLOv5 and YOLOv7 object detection algorithms. Journal of Computing and Social Informatics. 2023 Feb 8;2(1):1-2.

[14] Sathiyabhama, B., et al. "Tracing of vehicle region and number plate detection using deep learning." 2020 International conference on emerging trends in information technology and engineering (ic-ETITE). IEEE, 2020.

[15] Malaainine, Mohamed El Imame, Hatim Lechgar, and Hassane Rhinane. "YOLOv2 Deep Learning Model and GIS Based Algorithms for Vehicle Tracking." Journal of Geographic Information System 13.4 (2021): 395-409.

[16] Yang, Hao Frank, et al. "Cooperative multi-camera vehicle tracking and traffic surveillance with edge artificial intelligence and representation learning." Transportation research part C: emerging technologies 148 (2023): 103982.

[17] Y. Liu and S. Shen, "Vehicle Detection and Tracking Based on Improved YOLOv8," in IEEE Access, vol. 13, pp. 24793-24803, 2025, doi: 10.1109/ACCESS.2025.3538556.

[18] Vikruthi, Sriharsha, et al. "Tracking and Detection of Moving Vehicles using Hybrid Optical Flow Technique in Complex Transport Atmosphere." 2024 International Conference on Expert Clouds and Applications (ICOECA). IEEE, 2024.

[19] Anusha, N., et al. "Cloud-Enabled Neural Networks for Intelligent Vehicle Emissions Tracking and Analysis." 2024 International Conference on Automation and Computation (AUTOCOM). IEEE, 2024.

[20] Dong, Zheng, et al. "Watchdog: Real-time vehicle tracking on geo-distributed edge nodes." ACM Transactions on Internet of Things 4.1 (2023): 1-23.

[21] Kanagamalliga, S., et al. "Traffic Management through Cutting-Edge Vehicle Detection, Recognition, and Tracking Innovations." Procedia Computer Science 233 (2024): 793-800.

[22] Ashraf, Muhammad Hassaan, et al. "HVD-net: a hybrid vehicle detection network for vision-based vehicle tracking and speed estimation." Journal of King Saud University-Computer and Information Sciences 35.8 (2023): 101657.

[23] Liu, Xiaoxu, Wei Qi Yan, and Nikola Kasabov. "Moving vehicle tracking and scene understanding: A hybrid approach." Multimedia Tools and Applications 83.17 (2024): 51541-51558.

[24] Maraş, Bahri, Nafiz Arica, and Ayşın Ertüzün. "A robust vehicle tracking in low-altitude UAV videos." Machine Vision and Applications 34.5 (2023): 77.

[25] Ay, Sevinç, and Murat Karabatak. "A new automatic vehicle tracking and detection algorithm for multi-traffic video cameras." Traitement du Signal 40.2 (2023): 457.

[26] Shroff, Kaushik. "Affordable Vehicle Tracking System."

[27] Cai, Hui, Haifeng Lin, and Dapeng Liu. "TrafficTrack: rethinking the motion and appearance cue for multi-vehicle tracking in traffic monitoring." Multimedia Systems 30.4 (2024): 216.

[28] Wang, Jun, et al. "Learning convolutional self-attention module for unmanned aerial vehicle tracking." Signal, Image and Video Processing 17.5 (2023): 2323-2331.

[29] Shibghatullah, Abdul S., et al. "Vehicle tracking application based on real time traffic." International Journal of Electrical and Electronic Engineering & Telecommunications 11.1 (2022): 67-73.

[30] Sreedhar, Sreela, Abin Oommen Philip, and M. U. Sreeja. "Autotrack: a framework for query-based vehicle tracking and retrieval from CCTV footages using machine learning at the edge." International Journal of Information Technology 15.7 (2023): 3827-3837.

[31] Chu, J., Wu, J., Zhang, W. *et al.* WBS-YOLOv8: A Lightweight Infrared Object Detection Algorithm for Nighttime Autonomous Driving. *SIViP* 19, 619 (2025). https://doi.org/10.1007/s11760-025-04214-4