

# Advanced Vehicle Detection and Modification Identification System for Intelligent Transportation Systems

<sup>1</sup>Navya N C, <sup>2</sup>Athira Raj <sup>3</sup>Athulbabu P, <sup>4</sup>Rohan K P, <sup>5</sup>Vaishnav M K

<sup>1</sup>Assistant Professor, <sup>2</sup>Student, <sup>3</sup>Student, <sup>4</sup>Student, <sup>5</sup>Student <sup>1</sup>Department of Computer Science and Engineering,  
<sup>1</sup>St. Thomas College of Engineering and Technology, Kannur, Kerala, India.

**Abstract :** With the rapid increase in vehicles on roads, intelligent transportation systems (ITS) require faster and more accurate vehicle detection techniques to improve traffic management and safety. Advances in computer vision and deep learning, particularly convolutional neural networks (CNNs), have significantly improved detection performance compared to traditional methods. Two-stage detectors such as Faster R-CNN provide high accuracy, while one-stage models like YOLO and SSD offer better real-time performance, making the choice dependent on application requirements. However, challenges remain in complex traffic environments, including poor weather conditions, night-time detection, and overlapping vehicles, where minimizing false positives and missed detections is essential. An important extension of vehicle detection is the identification of modified vehicles using artificial intelligence. By leveraging computer vision and deep learning techniques, AI systems can detect visual anomalies such as unauthorized structural changes, non-standard lighting, altered body parts, or illegal accessories. Integrating vehicle modification detection into Intelligent Transportation Systems enhances traffic enforcement, safety monitoring, and regulatory compliance, thereby improving the overall effectiveness of smart transportation systems.

**IndexTerms**—Intelligent Transportation Systems (ITS), Vehicle Detection, Illegal Vehicle Modification Detection, Deep Learning, Convolutional Neural Networks (CNN), Faster R-CNN, YOLO, Automatic Number Plate Recognition (ANPR)

**IndexTerms - Indian Sign Language (ISL), American Sign Language (ASL), MediaPipe , OpenCV, Convolutional Neural Networks (CNNs), Transformers, Multimodal Gesture Recognition, Real-Time Translation.**

## I. INTRODUCTION

The rapid growth of vehicular traffic has placed significant demands on modern Intelligent Transportation Systems (ITS), necessitating advanced solutions for traffic monitoring, law enforcement, and road safety. Accurate vehicle detection forms the backbone of ITS applications such as traffic flow analysis, automated surveillance, and violation detection. Recent advances in computer vision and deep learning—particularly convolutional neural networks (CNNs)—have significantly improved vehicle detection performance, enabling robust real-time analysis using models such as Faster R-CNN, YOLO, and SSD. However, most existing systems are limited to identifying the presence of vehicles and do not address the growing issue of illegal vehicle modifications. Unauthorized vehicle modifications, including altered exhaust systems, non-standard body kits, excessive noise emissions, and improper lighting, pose serious safety, environmental, and legal challenges. In regions such as Kerala, India, these modifications are strictly prohibited under the Motor Vehicles Act, 1988, and reinforced by Supreme Court directives. Despite this, enforcement largely relies on manual inspection, which is time-consuming, inconsistent, and impractical in high-traffic or low-visibility conditions such as nighttime or adverse weather. To address these limitations, this paper proposes a multi-modal AI-based vehicle monitoring framework that extends conventional vehicle detection by integrating visual analysis for modification identification. Automatic Number Plate Recognition (ANPR) is incorporated to associate detected violations with vehicle registration data, enabling reliable verification and enforcement support. By fusing multiple data modalities and leveraging advanced deep learning models, the proposed approach enhances detection accuracy, reduces false positives, and improves robustness under challenging real-world conditions. The system is designed to support real-time operation and seamless integration with existing traffic surveillance infrastructure, offering a scalable and practical solution for intelligent traffic enforcement and road safety enhancement.

### A. Problem Statement

With the increasing number of vehicles on the road, unauthorized and illegal vehicle modifications have become a serious concern for traffic regulation and public safety. Many vehicle owners alter their vehicles by modifying number plates, exhaust systems, lighting, or structural components in ways that violate legal standards. These modifications can lead to increased noise pollution, reduced vehicle safety, difficulty in identification, and challenges in law enforcement. Current methods for detecting such violations rely heavily on manual inspection by traffic authorities, which is time-consuming, inconsistent, and limited in coverage.

It is impractical for officers to monitor every vehicle continuously, especially in high-traffic areas. As a result, many violations go undetected, reducing the effectiveness of traffic regulations. Additionally, modified number plates or hidden identifiers make it difficult for traditional surveillance systems to accurately recognize vehicles, further complicating enforcement efforts.

Therefore, there is a need for an intelligent, automated system that can monitor vehicles in real time, detect unauthorized modifications using AI and computer vision techniques, and assist authorities in enforcing traffic laws efficiently. Such a system should be capable of identifying different types of modifications and generating alerts for quick action, thereby improving road safety and regulatory compliance.

## II. LITERATURE SURVEY

A literature review serves as a systematic and comprehensive examination of existing scholarly works, including research articles, books, dissertations and conference proceedings, that are relevant to a specific topic, research question, or area of inquiry. It transcends a mere summary by actively involving critical analysis and synthesis of these prior publications, aiming to provide a robust conceptual framework for new endeavors. The fundamental purpose of conducting a literature review is to effectively map out the intellectual landscape of a given field, demonstrating a thorough understanding of what has already investigated, thereby ensuring that current research is well informed and avoids redundancy. This meticulous process allows researchers to identify key theories, established methodologies and prevailing findings, while simultaneously pinpointing inconsistencies, ongoing debates and, crucially, the discernible gaps in current knowledge that warrant further investigation. By rigorously evaluating the strengths and weaknesses, biases and limitations of previous studies, a literature review not only provides a solid foundation and essential contextual framework for new research but also critically helps to justify the necessity, originality and potential contribution of one's own academic work. It actively guides the formulation of precise research questions and hypotheses, ensuring they are grounded in existing scholarship. Ultimately, it establishes the researcher's credibility within the academic discourse by illustrating their comprehensive command of the subject matter and their sophisticated ability to integrate new insights harmoniously within the dynamic and evolving body of scholarly work.

### A. Vehicle Detection Based on Adaptive Multimodal Fusion and Cross-Modal Vehicle Index using RGB-T Images[1].

The paper "Vehicle Detection Based on Adaptive Multimodal Feature Fusion and Cross-Modal Vehicle Index Using RGB-T Images" proposes an advanced approach for accurate vehicle detection in aerial images captured under varying lighting and environmental conditions. Traditional RGB-based detection struggles in low-light or nighttime scenes, while thermal infrared (TIR) imaging loses texture and detail in bright conditions. To overcome this, the authors introduce an Adaptive Multimodal Feature Fusion and Cross-Modal (AFFCM) framework that intelligently combines RGB and TIR image features. The model employs a Softpooling Channel Attention (SCA) mechanism to assign adaptive importance to each modality based on scene illumination and a Multimodal Adaptive Feature Fusion (MAFF) module that merges features across multiple scales to detect vehicles of different sizes effectively.

Additionally, the paper introduces a Cross-Modal Vehicle Index (CMVI) to generate spatial weight maps that highlight potential vehicle regions while suppressing background clutter, thereby reducing false detections. Experiments conducted on the Drone Vehicle dataset demonstrate significant performance improvements achieving up to 14.44% higher mean Average Precision (mAP) compared to single-modality approaches. This method proves especially effective under low-illumination conditions, balancing accuracy and efficiency through adaptive fusion. However, its reliance on well-aligned RGB-T image pairs and dataset-specific tuning may limit generalization to other scenarios. Overall, the study highlights the power of adaptive multimodal fusion in enhancing vehicle detection across complex real-world environments.

### B. Softpooling Channel Attention (SCA)[2].

The paper is designed to adaptively assign importance to features from both RGB and Thermal (TIR) images during fusion. Instead of relying on simple pooling methods like max or average pooling, softpooling captures more detailed global information by considering all pixel values with weighted contributions. This helps retain richer feature information from both modalities, which is crucial for detecting vehicles under varying processed through fully connected layers and a Sigmoid activation to generate channel-wise attention weights. These weights determine how much each channel from RGB and TIR should contribute, allowing the model to emphasize the more useful features—forexample, thermal information at night or RGB details in daylight. This adaptive weighting improves the fusion process and enhances detection accuracy across diverse environments.

### C. Multimodal Adaptive Feature Fusion (MAFF)[3].

The MAFF module plays a central role in effectively combining the RGB and Thermal(TIR) features extracted from images. After the Softpooling Channel Attention (SCA) module calculates the importance of each feature channel, the MAFF module uses these attention weights to selectively fuse the most informative features from both modalities. This ensures that the model focuses more on the features that contribute most to accurate vehicle detection, such as thermal cues at night or RGB texture details during the day. Additionally, MAFF employs a multi-scale fusion strategy, meaning it integrates features from different levels of the neural network to handle vehicles of various sizes and distances. By adaptively weighting and merging these multi-scale, multi-modal features, the module enhances both the representation quality and the detection precision. Overall, the MAFF module ensures that RGB and TIR data complement each other dynamically, making the detection process more robust across complex lighting and environmental conditions.

## III. METHODOLOGY

The proposed system follows a structured approach using computer vision and deep learning techniques to detect unauthorized vehicle modifications in real time. Initially, a dataset consisting of images and videos of both normal and modified vehicles is collected from sources such as traffic cameras, public datasets, and custom recordings. The collected data is then preprocessed by cleaning, resizing, and annotating relevant vehicle components such as number plates, exhaust systems, and lighting using labeling tools. After preprocessing, the dataset is divided into training and testing sets. A deep learning model, such as YOLO (You Only Look Once), is selected for its ability to perform real-time object detection efficiently. The model is trained on the labeled dataset to identify different vehicle parts and detect anomalies or modifications. Once trained, the model is integrated with a live camera feed, where each video frame is analyzed to detect vehicles and their components. A rule-based logic is applied to compare detected features with standard regulations, enabling the system to identify illegal modifications such as non-standard number plates, excessive lighting, or altered exhaust systems. When a violation is detected, the system generates an alert and stores relevant details such as the captured image, time, and location, which can be accessed by authorities. Finally, the system is evaluated based on performance metrics such as accuracy, precision, and recall under various conditions to ensure reliability and effectiveness.

## IV. PROPOSED SYSTEM

### A. Objectives

The primary objective of the proposed system is to develop an intelligent AI-based camera system capable of automatically detecting unauthorized vehicle modifications in real time. The system aims to reduce the dependency on manual inspection by traffic authorities and improve the efficiency and accuracy of monitoring vehicles on roads. It is designed to identify various types of illegal modifications, including non-standard number plates, altered exhaust systems, and excessive or unauthorized lighting, using advanced computer vision and deep learning techniques. Another key objective is to provide a reliable and scalable solution that can operate continuously in different environmental conditions such as varying lighting and traffic density. The system also seeks to generate instant alerts and maintain records of detected violations, including images, timestamps, and location details, to assist authorities in enforcement and decision-making. Overall, the proposed system aims to enhance road safety, ensure regulatory compliance, and contribute to smarter traffic management through automation and real-time monitoring.

### B. System Architecture

The system architecture begins with a camera or video input source that continuously captures real-time footage of vehicles. The captured video is divided into individual frames using the frame capture module. These frames are then passed through a preprocessing stage where image enhancement techniques such as resizing, noise reduction, and normalization are applied to improve detection accuracy. The processed frames are fed into a deep learning-based vehicle detection module, such as YOLO, which identifies vehicles and their key components like number plates, lights, and exhaust systems.

Next, the feature extraction module isolates these components for detailed analysis. The extracted features are then evaluated using a modification detection logic, which compares them against predefined standards to identify illegal modifications such as non-standard number plates, excessive lighting, or modified exhaust systems. If any violation is detected, the system triggers the alert and storage module, which records relevant details including images, timestamps, and location information. Finally, this information is displayed on a user interface or dashboard, enabling authorities to monitor violations and take necessary actions efficiently.

### C. Functional Modules

The proposed AI-based vehicle modification detection system is divided into several functional modules, each responsible for a specific task in the overall workflow. The first module is the **Input Acquisition Module**, which captures real-time video or images from surveillance cameras or other input sources. This is followed by the **Preprocessing Module**, where the captured frames are enhanced through resizing, noise reduction, and normalization to improve the quality of input data for further analysis.

The next module is the **Vehicle Detection Module**, which uses deep learning models such as YOLO to identify vehicles within each frame. Once a vehicle is detected, the **Feature Extraction Module** extracts important components such as number plates, lights, and exhaust systems. These extracted features are then passed to the **Modification Detection Module**, where predefined rules and trained models are used to identify any unauthorized modifications by comparing them with standard vehicle specifications.

If a violation is detected, the **Alert and Reporting Module** is activated, which generates alerts and stores relevant information such as images, timestamps, and location details. Finally, the **User Interface Module** provides a dashboard for authorities to view detected violations, monitor system performance, and take appropriate actions. Together, these modules ensure efficient, automated, and real-time detection of vehicle modifications.

### D. Data Flow and Implementation

The proposed system follows a systematic data flow starting from real-time video acquisition to final alert generation and reporting. Initially, video data is captured from surveillance cameras or input devices and converted into individual frames for processing. These frames undergo preprocessing steps such as resizing, noise reduction, and normalization to enhance image quality and ensure consistency. The processed frames are then passed to a deep learning-based vehicle detection model, such as YOLO, which identifies vehicles and localizes their key components including number plates, lights, and exhaust systems.

Once the relevant features are extracted, the system applies a modification detection algorithm that analyzes these components and compares them with predefined legal standards to identify any unauthorized changes. If a violation is detected, the system triggers the alert mechanism, which records the event along with supporting data such as captured images, timestamps, and location information. This data is stored in a database for future reference and analysis. The implementation is carried out using technologies such as Python, OpenCV, and deep learning frameworks, enabling efficient real-time processing and scalability. Finally, the processed information is displayed through a user interface or dashboard, allowing authorities to monitor violations and take necessary actions effectively.

### V. PROPOSED SYSTEM DESIGN

The proposed system is designed as an AI-powered surveillance solution that automatically detects unauthorized vehicle modifications in real time using computer vision and deep learning techniques. The system consists of a camera-based input unit that continuously captures video footage of vehicles on the road. The captured data is processed through a series of modules, starting with preprocessing, where image quality is enhanced for accurate analysis. A deep learning model, such as YOLO, is then used to detect vehicles and identify their key components, including number plates, lights, and exhaust systems.

The system incorporates a feature extraction mechanism to isolate these components, followed by a modification detection module that compares them against predefined legal standards to identify violations. Upon detecting any unauthorized modification, the system triggers an alert and records essential details such as images, timestamps, and location information. The design also includes a storage system to maintain violation records and a user-friendly dashboard that allows authorities to monitor and manage detected cases efficiently. Overall, the system is designed to be scalable, accurate, and capable of operating in real-time under various environmental conditions, ensuring improved road safety and efficient traffic law enforcement.

#### User Interface Layer

The User Interface Layer is the front-end component of the proposed system that enables interaction between the system and the end users, such as traffic authorities or administrators. It provides a user-friendly dashboard where detected vehicle modification violations are displayed in an organized manner. The interface presents important details including captured images of vehicles, type of modification detected, timestamp, and location information. It also allows users to view live camera feeds, monitor system

activity, and search or filter past records for analysis.

Additionally, the interface may include features such as alert notifications for real-time violations, report generation, and data visualization through charts or logs. The design of the interface focuses on simplicity, clarity, and ease of use, ensuring that users can quickly interpret the information and take necessary actions. Technologies such as web-based frameworks (e.g., React) can be used to develop a responsive and scalable interface that can be accessed across different devices.

---

## API Gateway / Reverse Proxy

The API Gateway / Reverse Proxy acts as an intermediary layer between the client-side user interface and the backend services of the proposed system. It is responsible for managing all incoming requests from users and routing them to the appropriate backend modules, such as the detection system, database services, or alert mechanisms. This layer helps in improving system security by hiding internal service details and enabling authentication and authorization for access control.

Additionally, it enhances system performance through load balancing, request filtering, and caching mechanisms, ensuring efficient handling of multiple user requests simultaneously. The API Gateway also standardizes communication by providing a unified interface for all services, making the system more scalable and easier to maintain. Technologies such as Nginx or Express-based gateways can be used to implement this layer effectively, ensuring smooth and secure communication between the frontend and backend components of the vehicle modification detection system.

## Input Layer

The Input Layer is the initial stage of the proposed system, responsible for capturing real-time data from external sources such as surveillance cameras, CCTV systems, or recorded video feeds. This layer continuously acquires video streams or images of vehicles in various traffic environments, ensuring a steady flow of data into the system. It supports different input formats and resolutions, making it adaptable to multiple camera setups and conditions such as day/night or varying weather.

The captured data is then forwarded to the processing pipeline, where it is converted into individual frames for further analysis. The efficiency and quality of the Input Layer play a crucial role in the overall performance of the system, as accurate detection depends on clear and consistent input data. This layer is designed to operate in real time, ensuring that vehicle movements are captured without delay, enabling timely detection of unauthorized modifications.

## Processing Pipeline

The Processing Pipeline is the core component of the proposed system where the captured input data is analyzed and transformed into meaningful information. Once video frames are obtained from the input layer, they undergo preprocessing steps such as resizing, normalization, and noise reduction to improve image quality and ensure consistency. The processed frames are then passed through a deep learning-based detection model, such as YOLO, which identifies vehicles and localizes their key components like number plates, lights, and exhaust systems.

Following detection, the system performs feature extraction to isolate these components for detailed analysis. A modification detection logic is then applied, where the extracted features are compared with predefined legal standards to identify any unauthorized modifications. The pipeline is designed to operate in real time with optimized performance, ensuring minimal delay between data capture and result generation. Overall, this pipeline enables accurate, efficient, and automated analysis of vehicle data for detecting violations.

---

## Design Summary

The proposed system is an AI-based solution that captures vehicle data through cameras and processes it using deep learning models to detect unauthorized modifications. It analyzes vehicle components, identifies violations, and stores the results in a

database. A user-friendly dashboard displays alerts and records, while an API layer ensures secure and efficient communication. The system is designed for real-time, accurate, and automated traffic monitoring.

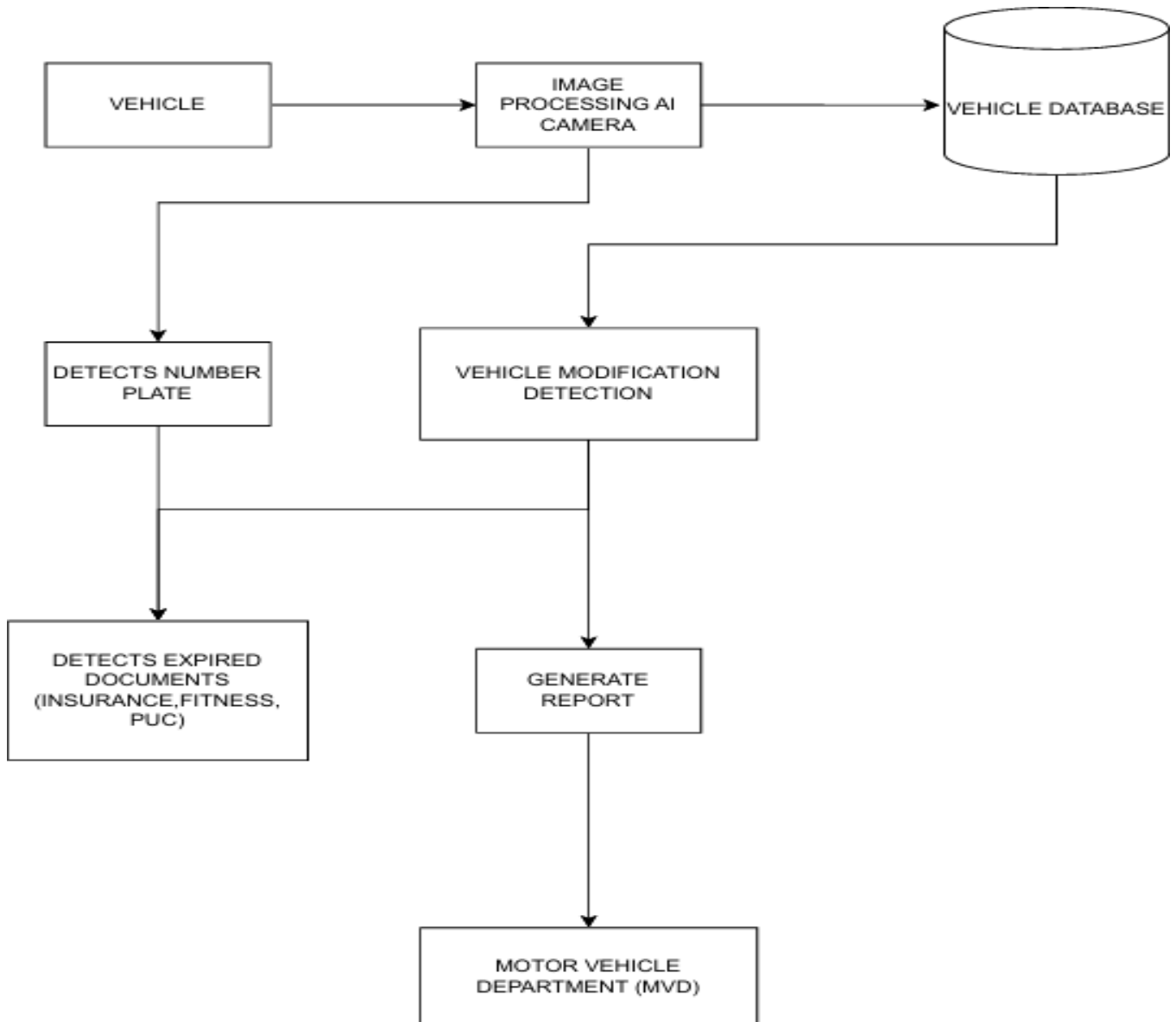


Figure 1. Architecture Diagram

The system architecture consists of multiple layers including input, processing, and output to ensure smooth data flow. It captures video from cameras, processes it using AI models to detect vehicle modifications, and analyzes the results. Finally, it stores violation data and displays alerts through a user-friendly dashboard for authorities.

**Actors**

- **User:** The primary actor who interacts with the system to perform sign language translation.
- **Admin:** Responsible for managing system settings, datasets, and models.

- **System:** Represents the internal processing unit that performs recognition and translation tasks.

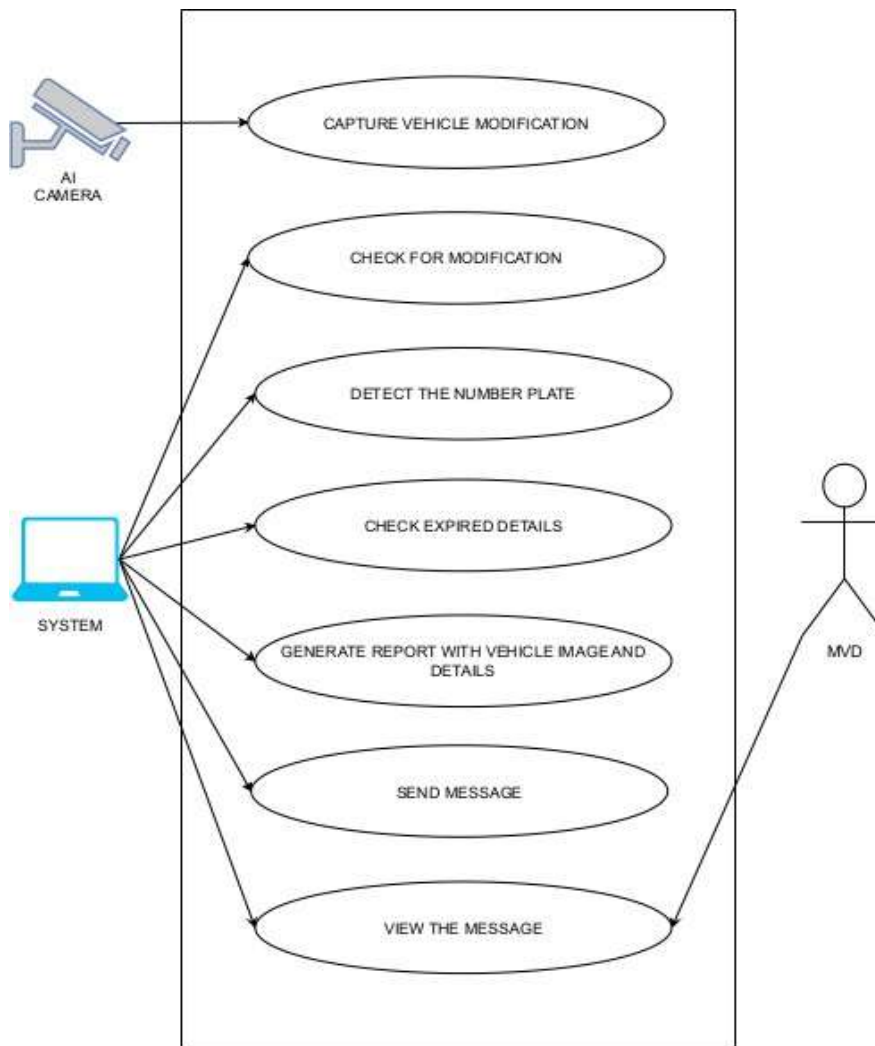


Figure 2. Use Case Diagram

### 1. Login/Register

The user (admin or traffic authority) logs into the system or creates a new account to securely access system features such as monitoring and reports.

### 2. Capture Vehicle Data

The system captures real-time video or images of vehicles through cameras installed on roads or surveillance systems.

### 3. User Management

The admin manages user accounts, roles, and permissions. This includes adding, updating, or removing users and controlling system access.

#### 4. System Configuration

The admin configures system settings such as detection parameters, thresholds, camera settings, and alert rules to ensure accurate performance.

#### 5. Dataset and Model Management

The admin manages datasets and AI models used for vehicle and modification detection. This includes training, updating, and maintaining models for better accuracy.

#### 6. Detect Vehicle

The system identifies vehicles in the captured video frames using deep learning models like YOLO.

#### 7. Extract Vehicle Features

The system extracts important components such as number plates, lights, and exhaust systems from detected vehicles.

#### 8. Detect Vehicle Modifications

The system analyzes extracted features and compares them with standard regulations to identify unauthorized modifications.

#### 9. Generate Alert

When a modification is detected, the system generates an alert with relevant details such as image, time, and type of violation.

#### 10. Store and Manage Data

The system stores detected violations, images, timestamps, and other details in a database for future reference and analysis.

#### 11. View Reports / Dashboard

The user (authority/admin) can view real-time alerts, past records, and reports through a dashboard for monitoring and decision-making.

#### 12. Display Results

The system displays the final output, including detected violations and vehicle details, in a clear and user-friendly format.

### VI. RESULTS

The system begins with user authentication through login or registration, after which real-time vehicle data is captured using cameras. The captured data is processed to detect vehicles and extract key features such as number plates, lights, and exhaust systems. Using trained AI models, the system analyzes these features to identify any unauthorized modifications based on predefined rules. When a violation is detected, an alert is generated and relevant details such as images, timestamps, and violation type are stored in a database. The admin manages users, system configurations, and model updates to ensure accurate performance. Finally, the results are displayed through a user-friendly dashboard, allowing authorities to monitor, analyze, and take necessary actions efficiently.

The proposed system operates by capturing live video streams of vehicles and processing them through an AI-based detection pipeline. After user authentication, the system continuously identifies vehicles in each frame and extracts important components

such as number plates and external features. These features are analyzed using trained models to detect any illegal modifications. If a violation is found, the system automatically generates alerts and records the event with supporting details. Administrators can manage users, configure system settings, and update models, while authorities can view real-time outputs and reports through a dashboard, ensuring efficient monitoring and enforcement of traffic regulations.

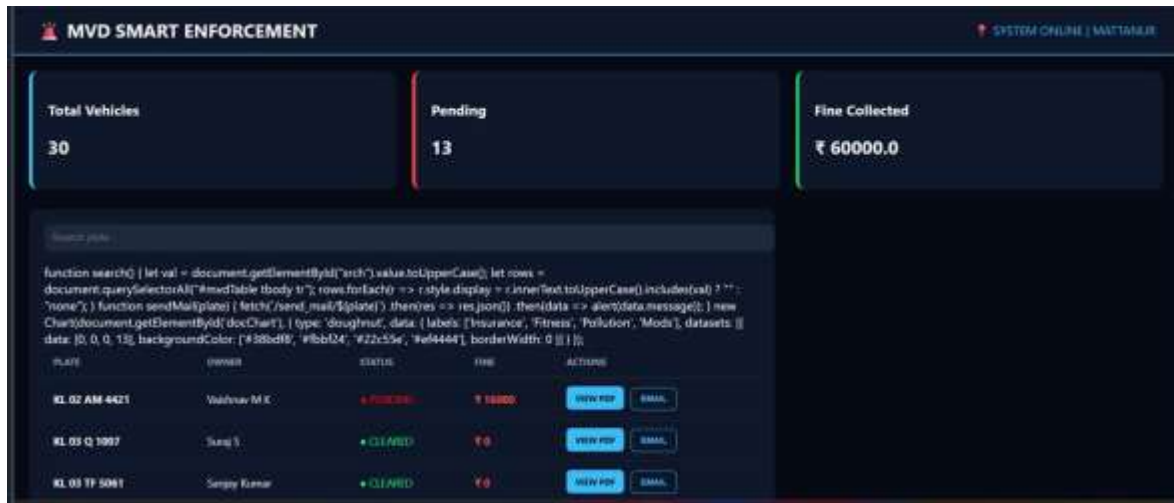


Fig 3. Dashboard

PLATE	OWNER	STATUS	FINE	ACTIONS
KL 02 AM 4421	Vaishnav M K	PENDING	₹ 16000	VIEW PDF, EMAIL
KL 03 Q 1007	Suraj S	CLEARED	₹ 0	VIEW PDF, EMAIL
KL 03 TF 5061	Sanjay Kumar	CLEARED	₹ 0	VIEW PDF, EMAIL
KL 04 V 6102	Shyam Lal	CLEARED	₹ 0	VIEW PDF, EMAIL
KL 05 BE 7349	Ashwin Babu	CLEARED	₹ 0	VIEW PDF, EMAIL
KL 05 RJ 0812	Joseph Varghese	CLEARED	₹ 0	VIEW PDF, EMAIL
KL 06 WM 8025	Aditya Nandan	PENDING	₹ 500	VIEW PDF, EMAIL
KL 07 AB 1234	Rahul	PENDING	₹ 2000	VIEW PDF, EMAIL
KL 07 AD 1234	Rahul Menon	PENDING	₹ 27500	VIEW PDF, EMAIL
KL 07 DA 8367	Akhil	PENDING	₹ 1000	VIEW PDF, EMAIL
KL 07 DE 3672	Thomas Mathew	PENDING	₹ 500	VIEW PDF, EMAIL

Fig 4 : Dashboard



fig 5:Modification Detection



trafficvision.ai@gmail.com

to rohankp1608 ▾


Greetings Rahul,

A traffic violation was detected for your vehicle KL07AD1234.

Violations: Tinted\_Glass, Spoiler, Body\_kit, Exhaust, Exhaust, Exhaust

Fine: ₹6500

\*\*\*

One attachment • Scanned by Gmail ⓘ  Add to Drive



← Reply

→ Forward



fig6:Generated Report

## VII. CONCLUSION

The proposed AI-based vehicle modification detection system provides an efficient and automated solution to identify unauthorized vehicle alterations in real time. By leveraging computer vision and deep learning techniques, the system minimizes the need for manual inspection and improves the accuracy and speed of detection. It is capable of continuously monitoring vehicles, detecting violations, and generating instant alerts with relevant information for authorities. The integration of a user-friendly dashboard and data storage further enhances monitoring and decision-making. Overall, the system contributes to improved road safety, effective law enforcement, and the development of smarter traffic management systems.

## ACKNOWLEDGMENT

We would like to sincerely thank St. Thomas College of Engineering and Technology for their support. The resources provided through the Research and Development Lab were essential to the success of this project and gave us the environment we needed to turn this research into a working reality.

## REFERENCES

- [1] L.Figueiredo,I.Jesus,J.T.Machado,J.R.Ferreira,andJ.M.DeCarvalho, "Towards the development of intelligent transportation systems," in Proc. IEEE Intell. Transp. Syst. (ITSC), Aug. 2001, pp. 1206–1211.
- [2] J.Azimjonov and A. Ozmen, "A real-time vehicle detection and a novel vehicle tracking systems for estimating and monitoring traffic flow on highways," Adv. Eng. Informat., vol. 50, Oct. 2021, Art. no. 101393.
- [3] J. Dai, H. Qi, Y. Xiong, Y. Li, G. Zhang, H. Hu, and Y. Wei, "Deformable convolutional networks," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Oct. 2017, pp. 764–773.
- [4] X. Yin, Z. Yu, Z. Fei, W. Lv, and X. Gao, "Pe-YOLO: Pyramid enhancement network for dark object detection," in Proc. Int. Conf. Artif. Neural Netw. Cham, Switzerland: Springer, 2023, pp. 163–174.
- [5] B. Zhang, J. Suo, and Q. Dai, "A complementary dual-backbone transformer extracting and fusing weak cues for object detection in extremely dark videos," Inf. Fusion, vol. 97, Sep. 2023, Art. no. 101822.
- [6] K. He, X. Zhang, S. Ren, and J. Sun, "Deep residual learning for image recognition," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2016, pp. 770–778.
- [7] L. Wen, D. Du, Z. Cai, Z. Lei, M.-C. Chang, H. Qi, J. Lim, M.-H. Yang, and S. Lyu, "UA-DETRAC: A new benchmark and protocol for multi-object detection and tracking," Comput. Vis. Image Understand., vol. 193, Apr. 2020, Art. no. 102907.
- [8] F. Yu, H. Chen, X. Wang, W. Xian, Y. Chen, F. Liu, V. Madhavan, and T. Darrell, "BDD100K: A diverse driving dataset for heterogeneous multitask learning," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit. (CVPR), Jun. 2020, pp. 2633–2642.
- [9] R. Girshick, "Fast R-CNN," in Proc. IEEE Int. Conf. Comput. Vis. (ICCV), Dec. 2015, pp. 1440–1448.
- [10] T.-Y. Lin, P. Dollar, R. Girshick, K. He, B. Hariharan, and S. Belongie, "Feature pyramid networks for object detection," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit. (CVPR), Jul. 2017, pp. 936–944. K. Li, G. Wan, G. Cheng, L. Meng, and J. Han, "Object detection in optical remote sensing images: A survey and a new benchmark," ISPRS J. Photogrammetry Remote Sens., vol. 159, pp. 296–307, 2020.
- [11] G. Cheng and J. Han, "A survey on object detection in optical remote sensing images," ISPRS J. Photogrammetry Remote Sens., vol. 117, pp. 11–28, 2016.
- [12] S. A. Ahmadi, A. Ghorbanian, and A. Mohammadzadeh, "Moving vehicle detection, tracking and traffic parameter estimation from a satellite video: A perspective on a smarter city," Int. J. Remote Sens., vol. 40, no. 22, pp. 8379–8394, 2019.

- [13] Q. Song and X. Chen, “Vehicle detection method for remote sensing images based on feature anti-interference and adaptive residual attention,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 7988–7996, 2022.
- [14] P. Gao, T. Tian, T. Zhao, L. Li, N. Zhang, and J. Tian, “Double FCOS: A two-stage model utilizing FCOS for vehicle detection in various remote sensing scenes,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 15, pp. 4730–4743, 2022.
- [15] X. Ma, K. Ji, B. Xiong, L. Zhang, S. Feng, and G. Kuang, “Light-YOLOv4: An edge-device oriented target detection method for remote sensing images,” *IEEE J. Sel. Topics Appl. Earth Observ. Remote Sens.*, vol. 14, pp. 10808–10820, 2021.

#### Copyright & License:



© Authors retain the copyright of this article. This work is published under the Creative Commons Attribution 4.0 International License (CC BY 4.0), permitting unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.