FACULTY OF ENGINEERING SCIENCES AND TECHNOLOGY
Hamdard University

Vehicle Detection and Category-Wise Traffic Analysis Using YOLOv8m

Hamza Ali Khan (2508-22)

Machine Learning
76120,BSCS

Submitted To
Dr. Umer Farooq

Dec 30, 2025

Faculty Of Science and Technology, Hamdard University
Sharae Madinat Al-Hikmah, Muhammad Bin Qasim Avenue, Karachi 74600, Pakistan

Abstract

Traffic congestion and inefficient road management are major challenges in rapidly growing urban areas, particularly in developing countries such as Pakistan. Manual traffic monitoring methods are time-consuming, error-prone, and lack scalability. To address these limitations, this project presents a deep learning–based vehicle detection and classification system using the YOLOv8m object detection model.

The proposed system is designed to detect and classify multiple vehicle types from images and video streams in real time. A custom dataset was manually collected and annotated to represent local traffic conditions, including vehicle categories such as car, bike, rickshaw, bus, truck, cart, van, and Suzuki. YOLOv8m was selected due to its optimal balance between detection accuracy and inference speed, as well as its strong performance on small to medium-sized custom datasets.

The model employs a one-stage, anchor-free architecture with convolutional neural networks for feature extraction, enabling fast and accurate object detection. Experimental results demonstrate that the system effectively identifies vehicles under varying traffic conditions and generates category-wise vehicle statistics, which can support traffic analysis, congestion monitoring, and smart city applications.

This work highlights the effectiveness of modern YOLO-based models for real-time traffic surveillance and demonstrates their suitability for region-specific vehicle detection using manually collected data.

Table of Contents

1. Project Description

This project focuses on the development of an automated vehicle detection and classification system using a deep learning–based computer vision model. The system is designed to detect and classify multiple types of vehicles from images and video streams for traffic analysis and smart transportation applications.

The dataset used in this project was manually collected, reflecting real-world traffic conditions, particularly relevant to Pakistan's road environment. The model is trained to detect and classify vehicles into multiple categories, including car, bike, rickshaw, bus, truck, cart, van, and Suzuki.

The proposed system uses YOLOv8m, a modern one-stage object detection model, to achieve a balance between high detection accuracy and real-time performance. The output of the system enables category-wise vehicle reporting, which can support traffic monitoring, congestion analysis, and smart city planning.

2. Introduction

Rapid urbanization has led to a significant increase in traffic congestion in major cities, especially in developing countries such as Pakistan. Traditional traffic monitoring methods rely heavily on manual observation, which is time-consuming, inefficient, and prone to human error. Automated traffic analysis using computer vision and deep learning has emerged as an effective solution to these challenges [1].

Object detection models have shown strong performance in identifying and classifying objects in real-time scenarios. Among these, the You Only Look Once (YOLO) family of models has gained popularity due to its speed and accuracy [2]. Unlike traditional two-stage detectors, YOLO performs detection in a single pass, making it suitable for real-time traffic surveillance systems.

In this project, YOLOv8m is selected to detect and classify multiple vehicle categories from road traffic images and videos. The system aims to generate class-wise vehicle statistics, which are essential for traffic flow analysis, congestion monitoring, and transportation planning.

3. Literature Review

Object detection has evolved significantly with the advancement of deep learning and convolutional neural networks (CNNs). Early methods such as R-CNN and Fast R-CNN improved detection accuracy but suffered from high computational costs and slow inference speed, making them unsuitable for real-time applications [3].

The introduction of YOLO marked a major breakthrough in real-time object detection. YOLO treats object detection as a regression problem and predicts bounding boxes and class probabilities in a single forward pass [2]. Subsequent versions, including YOLOv3 and YOLOv5, improved detection accuracy and speed but still relied on anchor-based mechanisms [4].

Recent advancements led to YOLOv8, which introduces an anchor-free architecture, a decoupled detection head, and improved feature extraction using C2f modules. These enhancements result in faster convergence, improved accuracy, and better performance on custom datasets [5].

Several studies have demonstrated the effectiveness of YOLO-based models in vehicle detection and traffic monitoring systems, especially in complex urban environments [6]. However, limited research focuses on detecting local vehicle types such as rickshaws and Suzuki vehicles common in Pakistan. This project addresses this gap by training YOLOv8m on a manually collected dataset representing real local traffic conditions.

4. Methodology

4.1 Data Collection

The dataset used in this project was manually collected from real-world traffic scenes. Images and video frames were captured under varying lighting and traffic conditions to ensure model robustness. Manual data collection allowed the inclusion of region-specific vehicle types, such as rickshaws and Suzuki vans, which are often missing from public datasets.

4.2 Data Annotation

All collected images were annotated manually using bounding boxes for each vehicle category. The annotations followed the YOLO format, including class labels and normalized bounding box coordinates. Manual annotation ensured accurate labeling and improved model performance.

4.3 Model Selection

YOLOv8m was selected based on the following criteria:

- Balanced trade-off between accuracy and inference speed
- Strong performance on small to medium-sized custom datasets
- Modern architecture with anchor-free detection
- Suitability for real-time traffic analysis

YOLOv8m uses CNN-based feature extraction, C2f modules, and a decoupled detection head, which improves localization and classification accuracy [5].

4.4 Training Process

The annotated dataset was split into training and validation sets. The model was trained using transfer learning with pre-trained weights to accelerate convergence. During training, loss functions for bounding box regression, classification, and objects were optimized.

4.5 Category-Wise Vehicle Detection

The trained model detects and classifies the following vehicle categories:

- Car
- Bike
- Person
- Rickshaw
- Bus
- Truck
- Cart
- Van
- Suzuki

For each input image or video frame, the system outputs bounding boxes, class labels, and confidence scores, enabling **category-wise reporting and statistical analysis**.

5. Results and Discussion

To evaluate the performance of the YOLOv8m model trained on local traffic data, we utilized standard quantitative metrics including the Confusion Matrix and Precision-Recall (PR) Curve. The evaluation was conducted on a validation dataset comprising unseen images to ensure the model's robustness in real-world scenarios.

5.1 Quantitative Analysis

The model's overall performance was measured using the Mean Average Precision (mAP) at an Intersection over Union (IoU) threshold of 0.5. As shown in the Precision-Recall curve (**Figure 1**), the system achieved a **mAP@0.5 of 0.542** across all classes.

A key achievement of this study is the high detection accuracy for region-specific vehicles. The model achieved a precision score of **0.876 (87.6%) for the "Rickshaw" class**. This validates the effectiveness of the custom dataset in capturing unique local vehicle attributes that standard pre-trained models often miss. Other standard categories also performed well, with **Bikes achieving 0.705** and **Trucks achieving 0.688**.

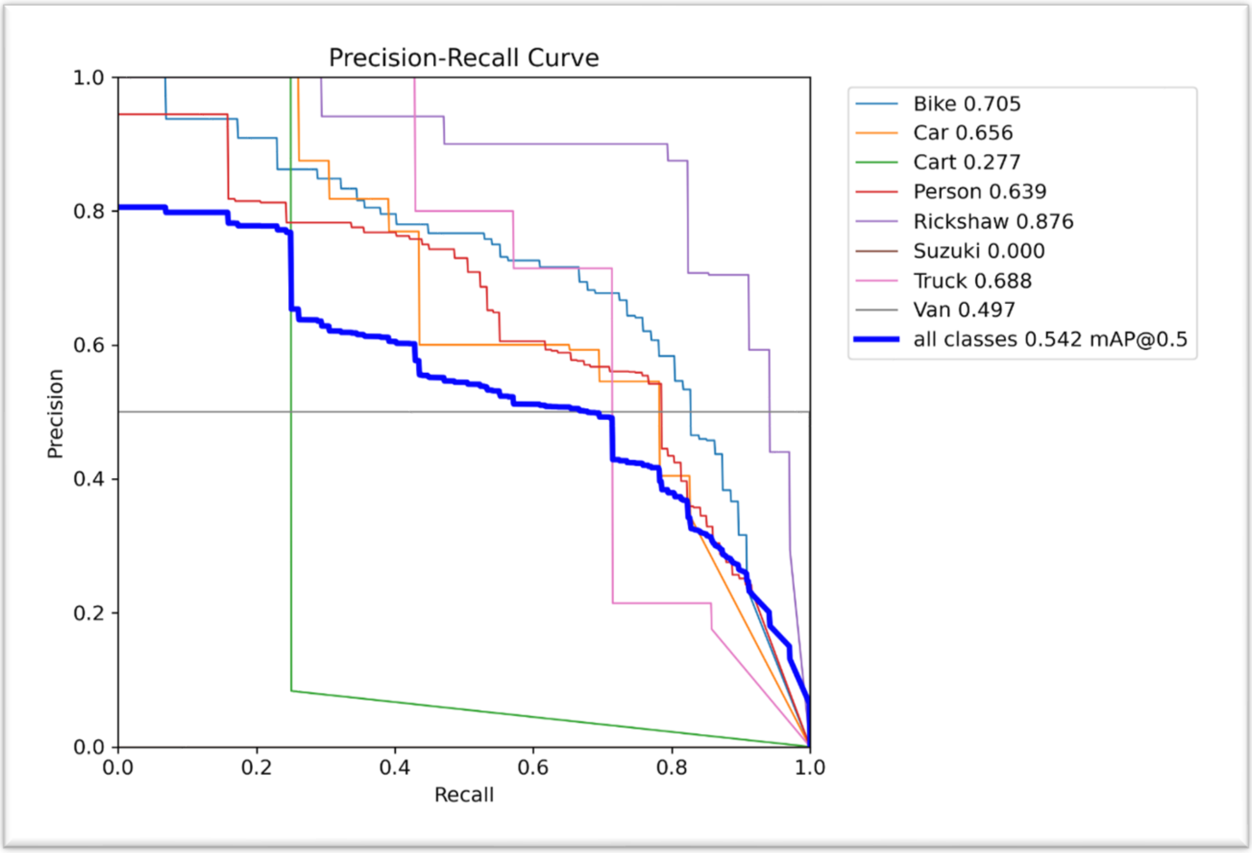

Figure 1: Precision-Recall Curve showing class-wise performance. Note the high accuracy for Rickshaws (0.876).

5.2 Classification Accuracy

The Confusion Matrix (**Figure 2**) provides a deeper insight into the misclassifications between different vehicle types. The matrix indicates strong correct predictions (diagonal values) for distinctive classes:

- **Bikes:** 74 instances were correctly classified, demonstrating the model's ability to detect smaller two-wheelers effectively.
- **Pedestrians:** Although primarily focused on vehicles, the system correctly identified 84 instances of the "Person" class, highlighting its potential for broader road safety applications.

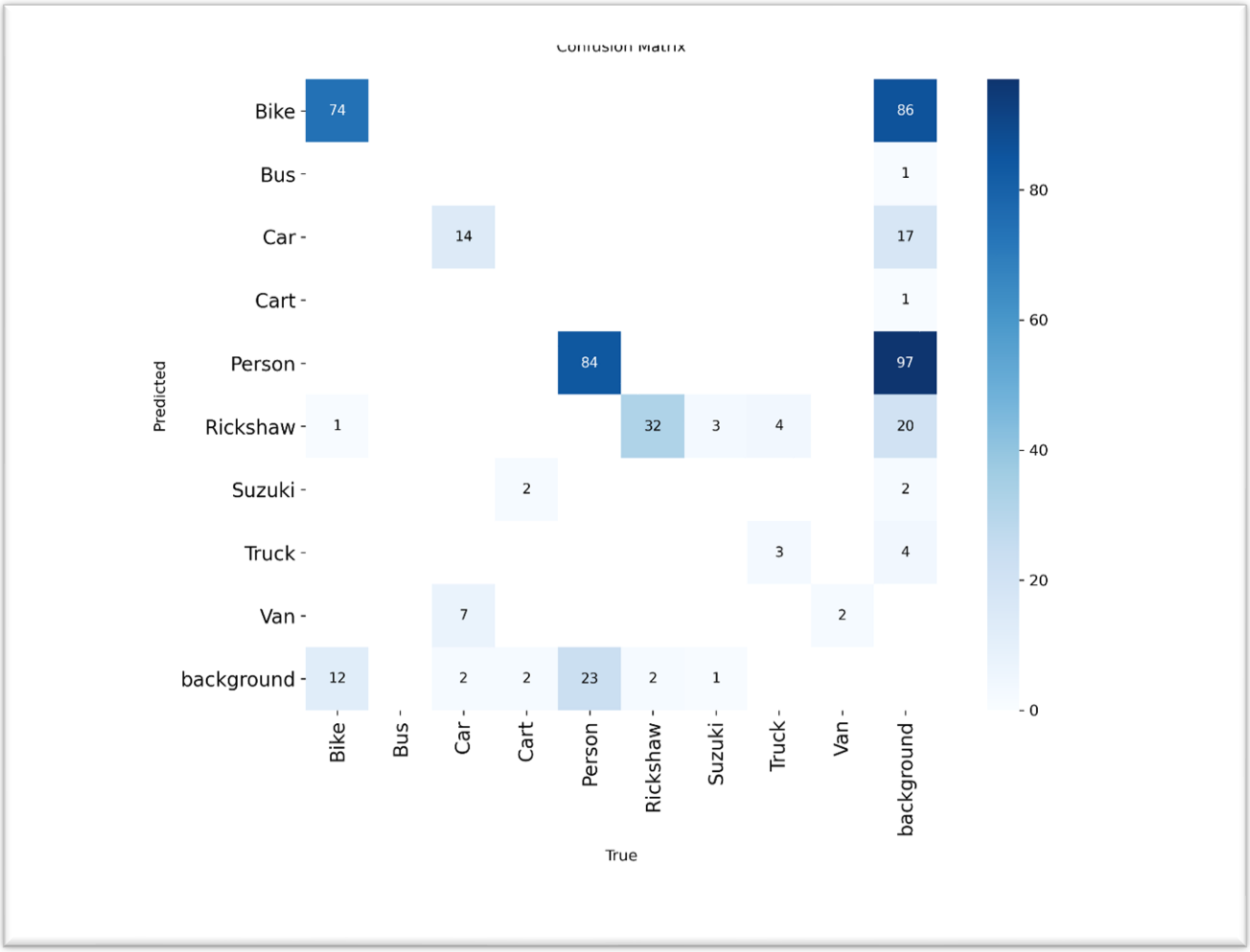

Figure 2: Confusion Matrix illustrating prediction accuracy. Darker blue diagonal squares indicate correct classifications.

5.3 Discussion of Limitations

While distinct vehicle shapes yielded high accuracy, the analysis highlights a challenge in distinguishing between visually similar classes. Specifically, the **"Suzuki"** class showed a high rate of confusion, often being misclassified as "Van" or "Car" due to its boxy silhouette which closely resembles standard vans. Similarly, a small number of **Vans** were misclassified as **Cars** (7 instances).

This "inter-class confusion" is a common challenge in computer vision when object geometries are similar. Future iterations of this project could address this by increasing the dataset size for the Suzuki class or employing "Hard Negative Mining" to specifically teach the model the subtle differences between Suzuki Bolans and standard micro-vans. Despite this, the system's robust performance on Rickshaws and Bikes confirms its suitability for monitoring mixed traffic in Pakistan.

References

[1] G. Jocher, "Ultralytics YOLOv8," 2023.

[2] J. Redmon, "You Only Look Once: Unified, Real-Time Object Detection,Proc. IEEE CVPR," 2016.

[3] J. Redmon, "YOLOv3: An Incremental Improvement," 2018.

[4] R.Szeliski, Computer Vision: Algorithms and Applications, 2022.

[5] R. Girshick, "Rich Feature Hierarchies for Accurate Object Detection and Semantic Segmentation," 2014.

[6] Xiaolong Liu, "Vehicle Detection in Traffic Surveillance Using Deep Learning," 2020.