Real-Time Vehicle Detection and Collision Risk Estimation Using YOLO and Depth Integration

Overview

This project presents a real-time vehicle detection system enhanced with depth-based distance estimation to improve spatial awareness in road scenes. By combining YOLO object detection with depth information from the KITTI dataset, the system moves beyond basic recognition to perception-driven analysis suitable for intelligent transportation and Advanced Driver Assistance Systems (ADAS).

Objective

The primary objective is to develop a computationally efficient system capable of:

• Detecting vehicles in real time

• Estimating their distance from the camera

• Evaluating collision risk dynamically

Dataset

The KITTI dataset, widely used in autonomous driving research, was utilized for training and evaluation.

Data leveraged:

• RGB images for object detection

• Ground-truth depth maps for distance estimation

This combination enables accurate real-world spatial measurements.

Methodology

Object Detection

YOLO was selected due to its single-stage architecture, enabling high-speed inference while maintaining strong detection accuracy.

Depth-Based Distance Estimation

For each detected vehicle:

1. Bounding box coordinates were extracted.

2. Corresponding depth pixels were retrieved.

3. Invalid values were removed.

4. Median depth was computed to ensure stable distance estimation.

Collision Risk Modeling

A dynamic risk scoring mechanism was introduced by combining estimated depth and bounding box area (used as a proxy for proximity).

Risk Score Formula:

Risk = (0.6 × Depth) + (0.4 × Bounding Box Area)

Objects were categorized into three risk levels:

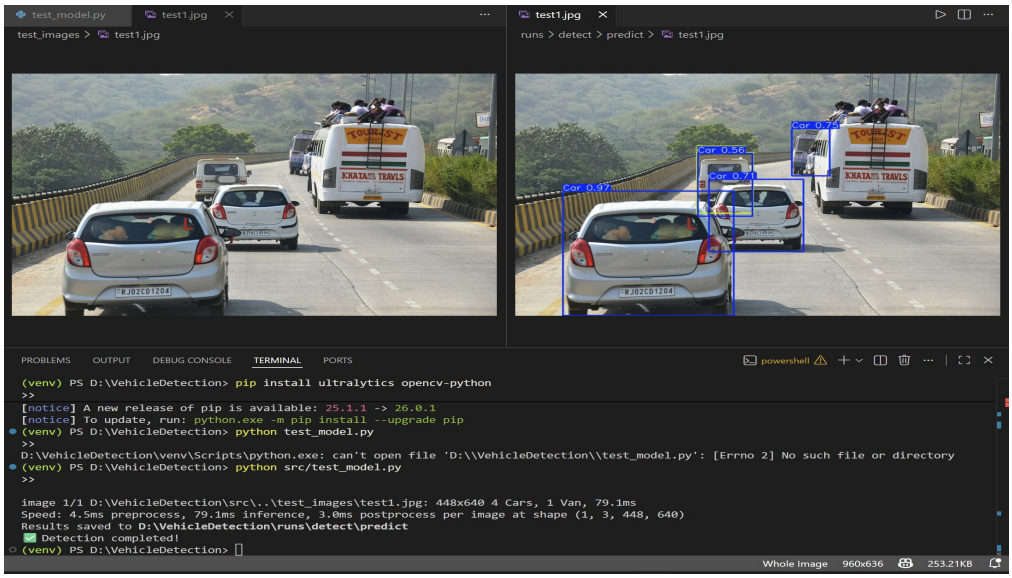
• High Risk

• Medium Risk

• Safe

This approach transforms the system from a detection model into a perception-oriented safety prototype.

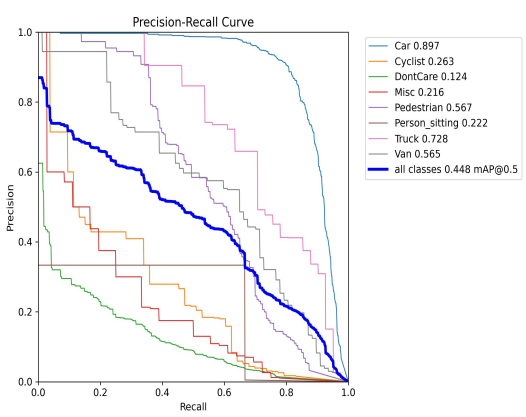
Detection Output :

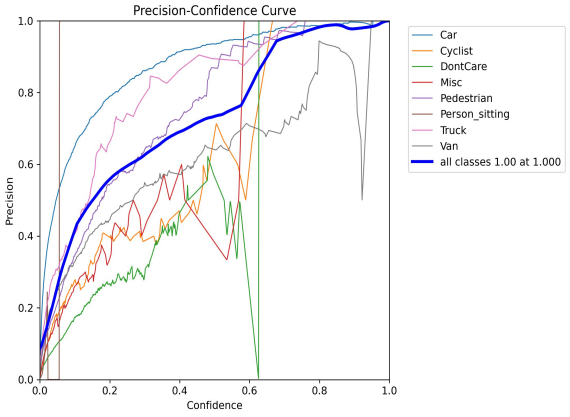


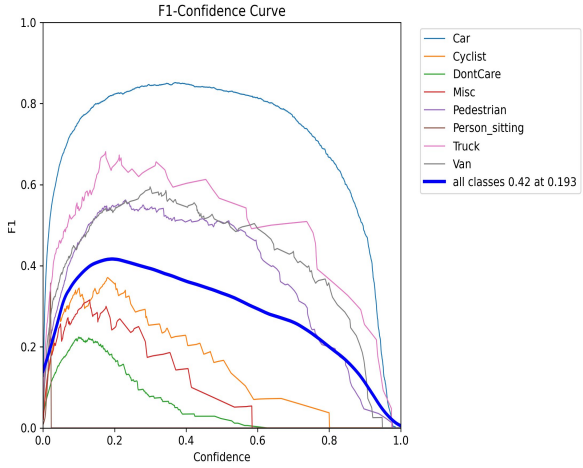
Performance Evaluation

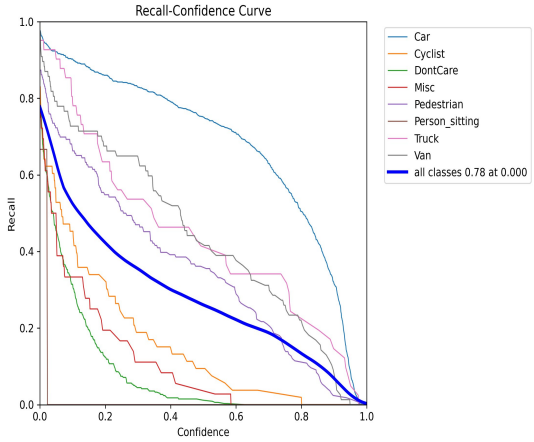
Metric	Value
Precision	**0.55**
Recall	**0.42**
mAP@50	**0.448**
mAP@50–95	**0.262**
Total Images (Validation)	225
Total Instances	1564
Inference Time	~2.3 ms per image
Post-processing Time	~4.7 ms per image

Detection Performance Curves and Confidence Analysis :

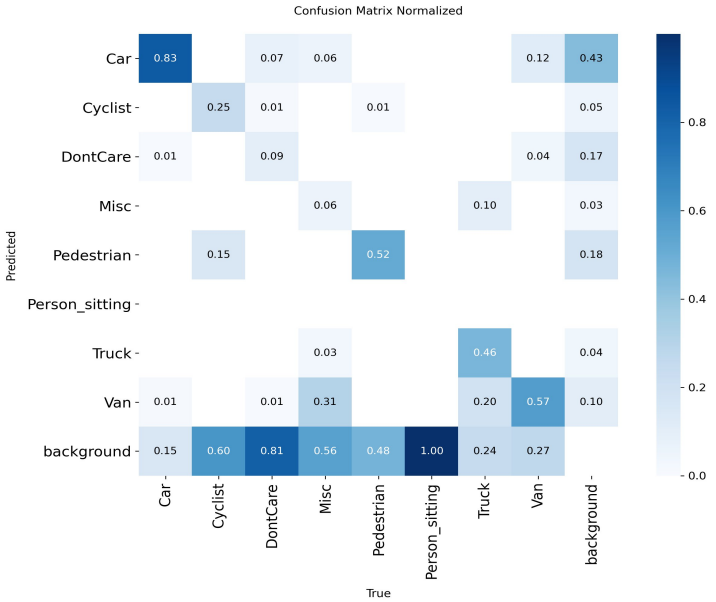


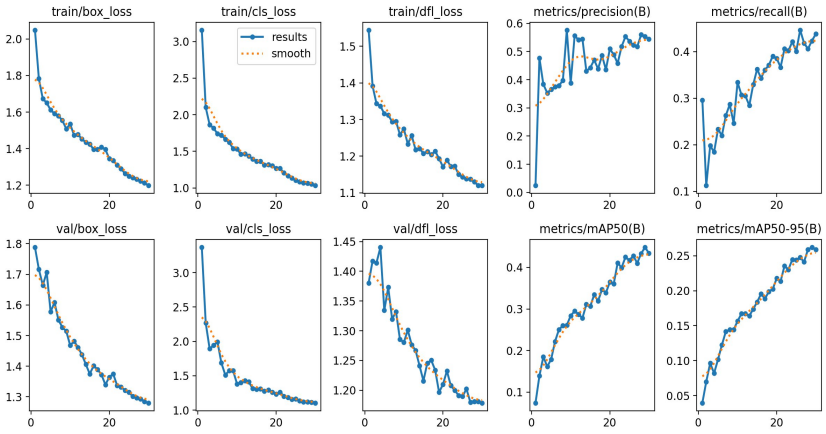




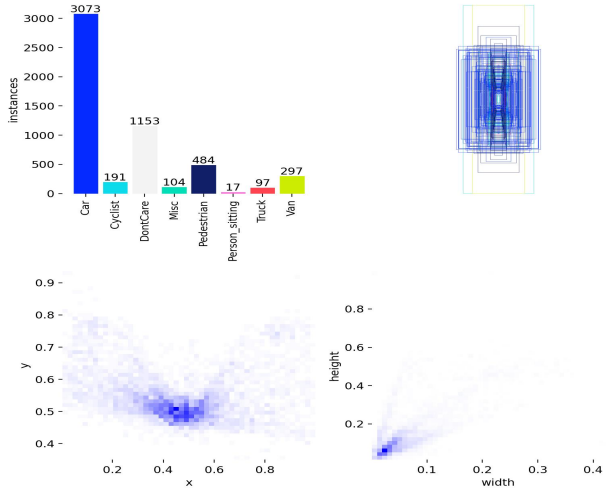


Training Dynamics and Confusion Matrix Analysis:





Bounding Box Spatial Analysis:



System Efficiency

The model achieves strong performance while maintaining low computational overhead due to YOLO's single-pass detection architecture. This balance between speed and accuracy makes the system suitable for real-world deployment scenarios.

Future Enhancements

This system can be extended with lane detection, object tracking, and audio alerts to function as a lightweight driver assistance module.

Conclusion

This project presents a real-time perception system that combines object detection with depth-based spatial reasoning. It enhances environmental awareness to support safer and more intelligent vision-driven transportation solutions.