

Object Detection Summary Engine - Task 1

Submitted by: Raj Chovatiya

Submission Date: 08 June 2025

1. Introduction

This task involves building an object detection summary engine using a pretrained model (YOLOv5). The model processes every 5th frame of a 15-20 second video and generates structured outputs, visualizations, and metadata describing object classes, bounding boxes, confidence scores, and summary statistics.

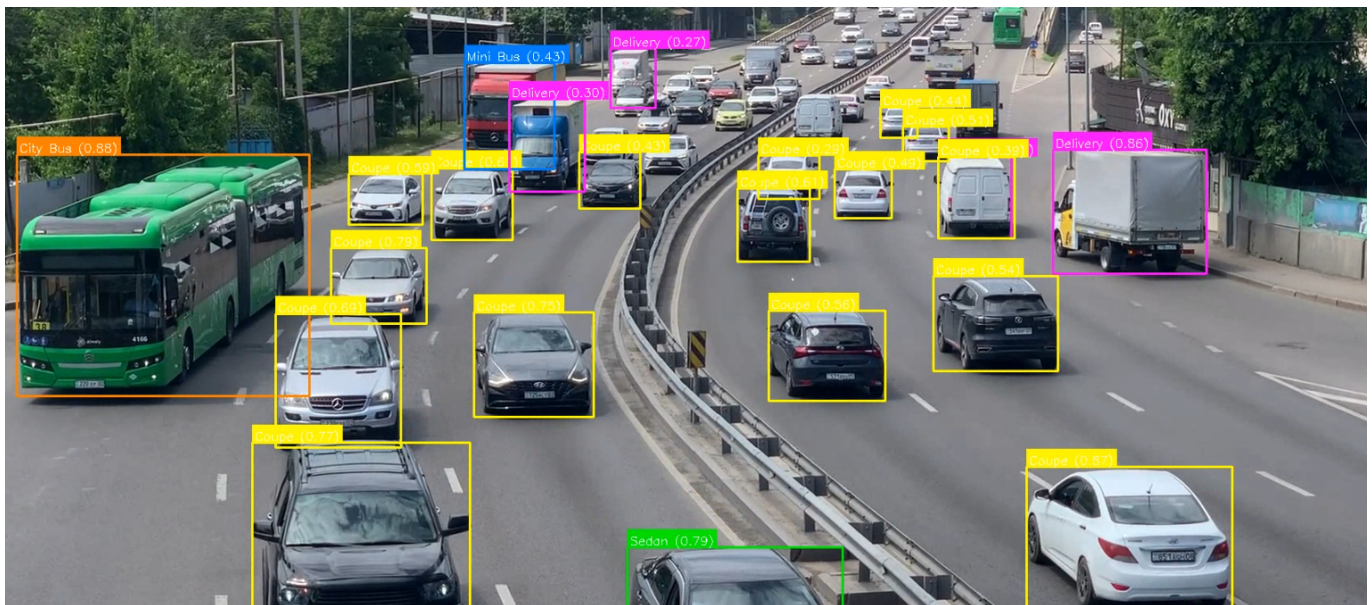
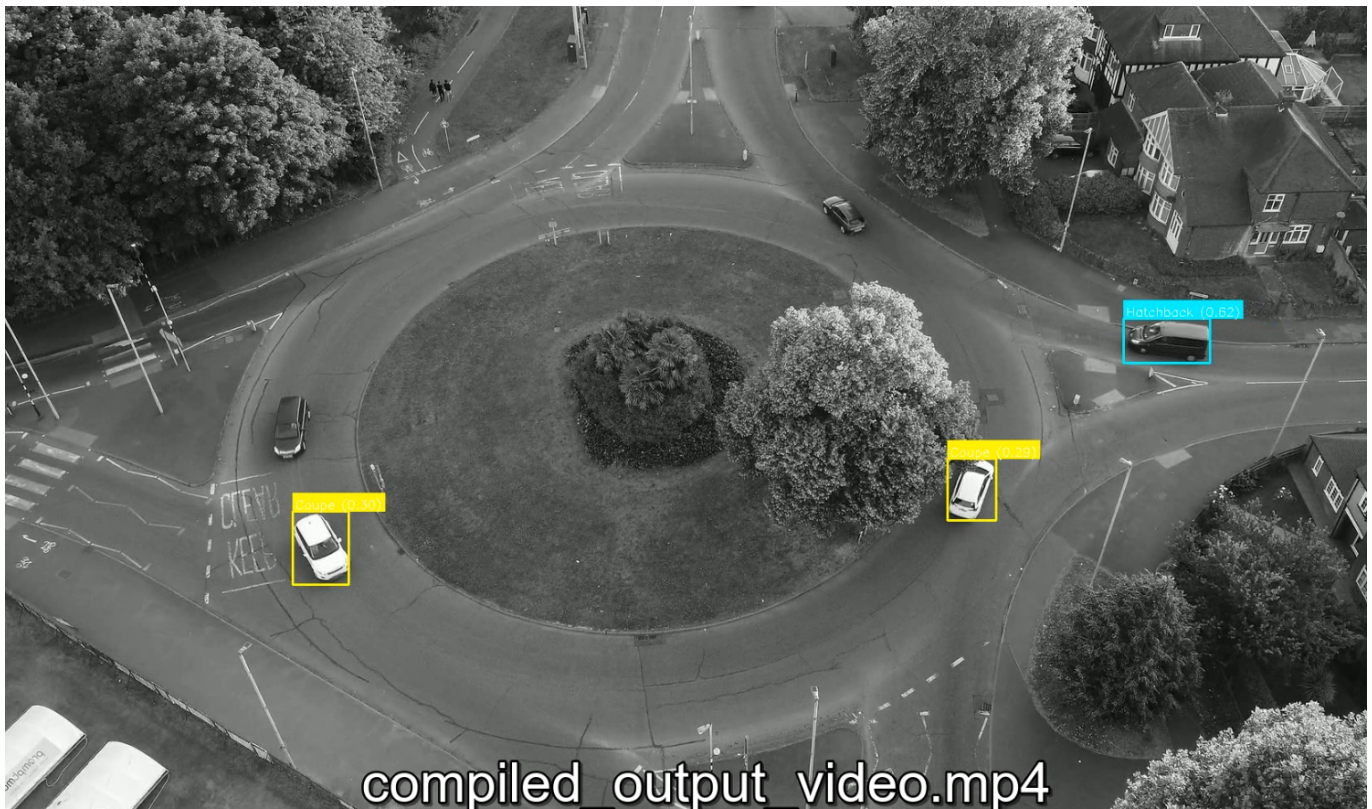
2. Tools and Libraries Used

- Python
- OpenCV
- YOLOv5 (PyTorch implementation)
- Matplotlib
- JSON

3. Processing Logic

The system loads a YOLOv5 model and processes every 5th frame from the input video. For each frame, it detects objects, stores class labels, bounding box coordinates, and confidence scores in JSON format. It calculates total object counts per class and identifies the frame with maximum class diversity.

4. Visual Results



5. Object Count & Frame Diversity

- Object counts were tallied per class.
- Frame with highest class diversity identified.
- Results visualized with bar chart (available in notebook).
- Optional video output includes frame annotations.

6. Conclusion

This task demonstrates a compact yet effective computer vision pipeline using pretrained object detection models. It highlights capabilities in model inference, visualization, statistical analysis, and documentation.