YOLOv5 Performance Evaluation on Traffic Video: Error Analysis for Vehicle Detection

Giovanni Mantovani and Andrea Palmieri

June 1, 2024

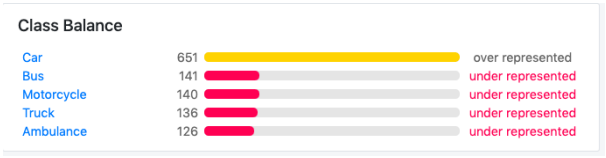
1 Introduction

The objective of this lab report is to evaluate the performance of the YOLOv5 object detection model on a recorded video containing various vehicle classes, namely ['Ambulance', 'Bus', 'Car', 'Motorcycle', 'Truck']. The evaluation will be conducted by analyzing the model's detections on the video frames, highlighting both correct and incorrect detections.

2 Methodology

2.1 Datasets

The Vehicles-OpenImages dataset, provided by Jacob Solawetz and released by Roboflow, consists of 627 images of various vehicle classes, including cars, buses, motorcycles, ambulances, and trucks. These images, resized to 416×416 pixels, are derived from the Open Images open source computer vision datasets and are labeled with bounding boxes to facilitate object detection tasks. The original data contained duplicate instances of each image and label file, which have been removed to ensure data integrity. It is important to note that the dataset has an overrepresentation of the car class compared to other vehicle classes. More information about the dataset can be found here.



2.2 Training

Training was performed on the Vehicles-OpenImages dataset using the YOLOv5(CSPDarknet53 backbone)training pipeline, by performing transfer learning on the Yolov5 weights from the original GitHub repository. A batch size of 8 and 25 epochs were used as hyperparameters.

3 Results

Figure 1: Image 1

Figure 2: Image 2

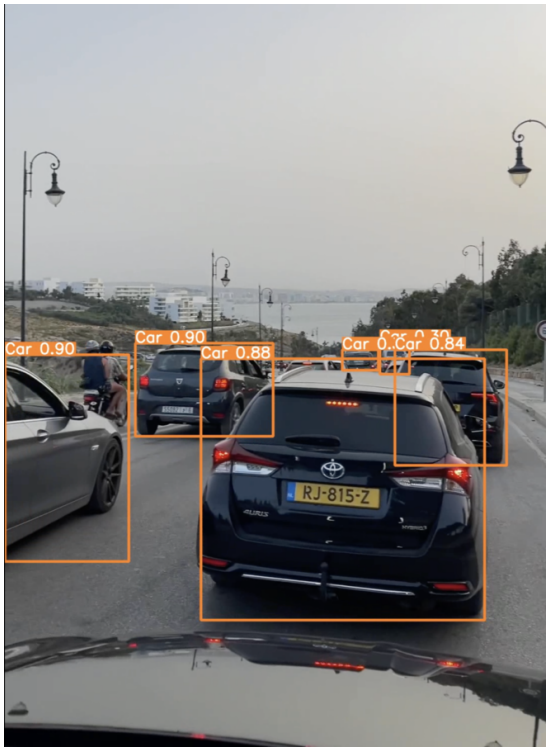


Figure 3: Image 3

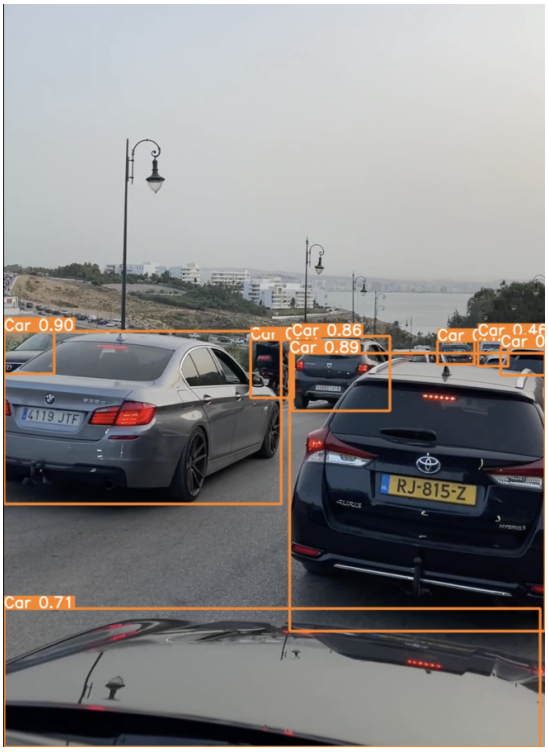


Figure 4: Image 4

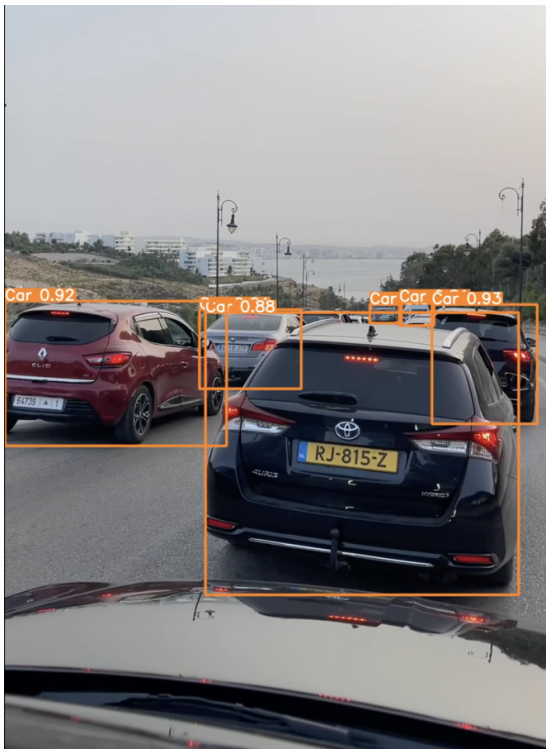


Figure 5: Image 5

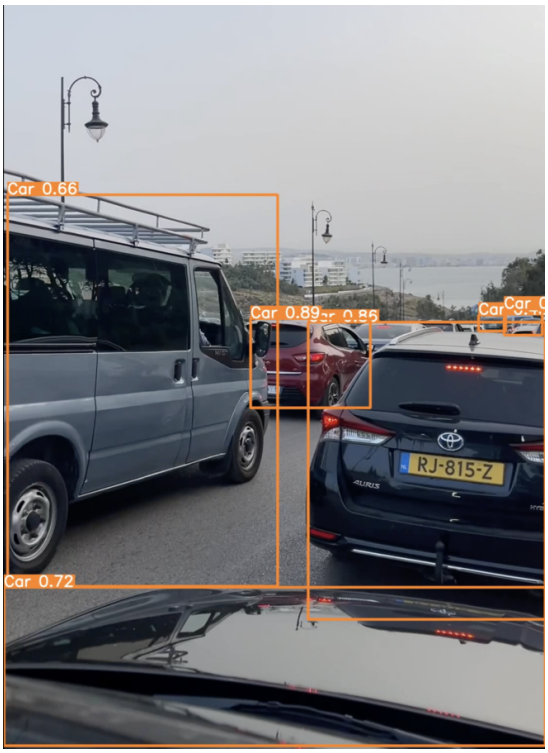


Figure 6: Image 6

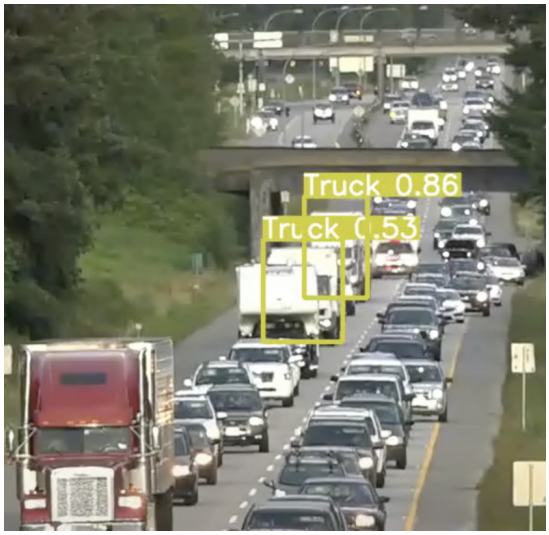


Figure 7: Image 7

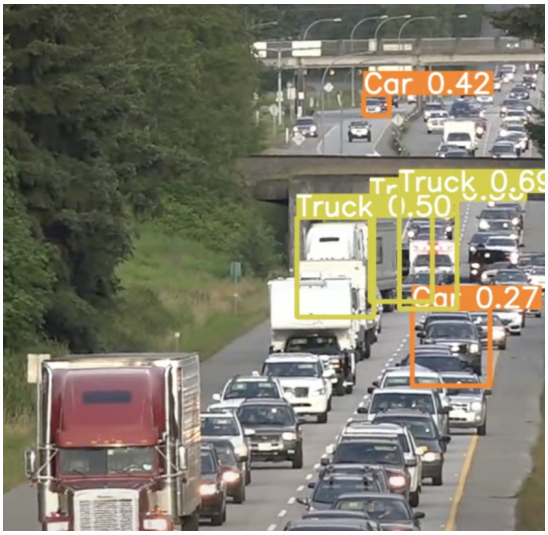


Figure 8: Image 8

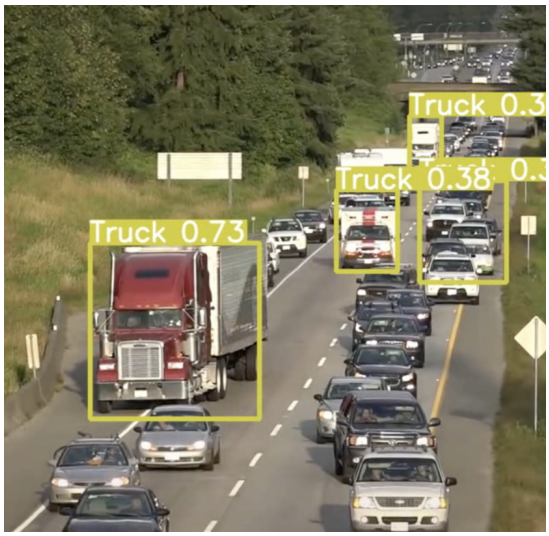


Figure 9: Image 9

Figure 10: Image 10

Figure 11: Image 11

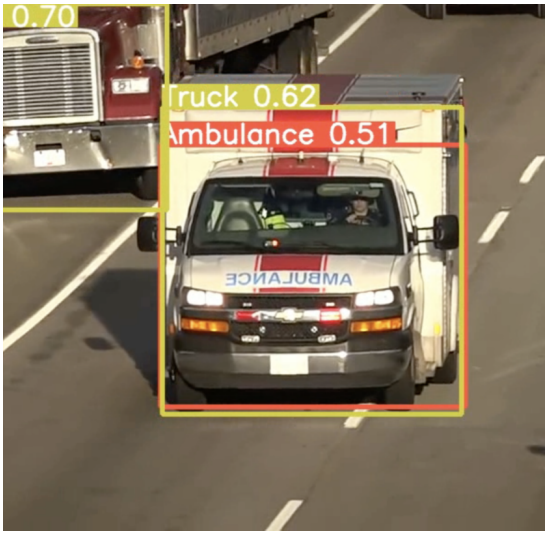


Figure 12: Image 12

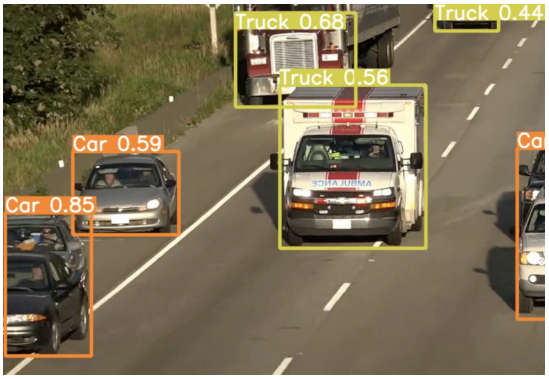


Figure 13: Image 13

4 Interpretation of Results

4.1 Aspects of Good Performance

The YOLOv5 model showed strong performance in the following areas:

- **High Confidence Detections:** The detection of cars was proportional to their size within the frame. The model had a harder time with trucks and performed better when cars were in the foreground. Larger vehicles, like trucks, were detected with higher confidence.
- **Consistency Across Frames:** The model consistently detected the same vehicles across consecutive frames, indicating reliable performance in tracking vehicles over time.

4.2 Areas of Poor Performance

The YOLOv5 model struggled in the following areas:

- **Small or Distant Objects:** The model had lower confidence scores and missed detections for smaller or more distant vehicles. It struggled to accurately detect cars that were partially occluded or appeared smaller in the frame.
- **Misclassification:** The model sometimes confused different vehicle types, such as mistaking cars for trucks, especially when visibility was poor.
- **Specific vehicle types Detection:** The model did not consistently detect specific vehicle types like ambulances. It often showed lower confidence scores for ambulances compared to other vehicles of similar size, as seen in frames where the detection confidence for ambulances was significantly lower (e.g., 0.51 and 0.41 in figure 11 12).

5 Conclusion

The results of our investigation showed that the model is strong at detecting vehicles with high confidence when they are clearly visible and larger in size, and it maintained consistency across consecutive frames. However, the model struggled with smaller or more distant vehicles, leading to lower confidence scores or missed detections. It also had difficulty accurately distinguishing between different vehicle types, particularly when they were partially occluded or appeared similar.