

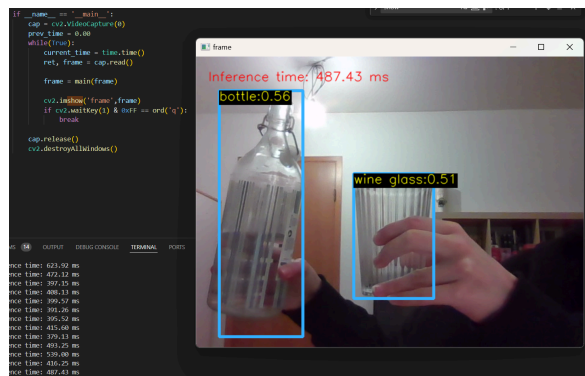
Assignment 3

Part I - pre trained models

I used the pre-trained YOLOv5 models:

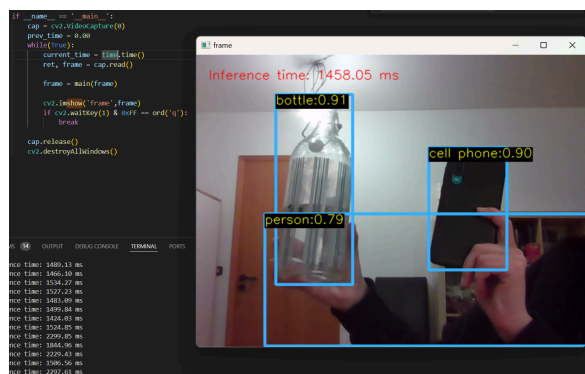
1. YOLOv5s (small):

- It achieved a significantly higher frame rate (inference around 400-700 ms), making it more suitable for real-time applications.
- The detection was quick but slightly less precise than the larger model. Some smaller objects were occasionally missed, especially those overlapping with other objects in the frame.
- Recall was high, with most objects being detected consistently, but often with lower confidence scores especially for distant or small objects.



2. YOLOv5l (large):

- Provided better precision due to its larger number of parameters, which allowed for more detailed feature extraction.
- The frame rate reduced significantly (inference around 1400-1700 ms).
- The model detected small or overlapping objects more accurately.



In conclusion, YOLOv5l outperformed YOLOv5s in terms of precision, correctly identifying objects with fewer false positives. Both models had similar recall, as they were both able to detect most objects in the video feed. However YOLOv5l was slightly better, for example, as we can see in the picture, detecting a person only through its hands.

Part 2 - fine tuned model

The model consisted of 157 layers and approximately 7 million parameters, requiring 15.8 GFLOPs for inference. The training was completed in 25 epochs over 7 minutes on Google Colab.

1. Create a confusion matrix:

While the training confusion matrix showed promising alignment, the test confusion matrix revealed high misclassification rates, the matrix revealed that misclassifications were frequent, particularly for smaller or less distinctive classes.

Actual \ Predicted	Ambulance	Bus	Car	Motorcycle	Truck
Ambulance	1	2	3	3	0
Bus	3	0	6	4	6
Car	5	5	22	31	12
Motorcycle	2	1	9	3	1
Truck	0	2	5	3	0

2. Compute the precision, recall, and the F1 score:

F1 Score: 0.11333333333333333

Precision: 0.12959595959595957

Recall: 0.11838888888888888

In conclusion, the test confusion matrix shows poor classification performance, especially for smaller classes like "Truck" and "Bus". The model struggles with both accurate predictions and minimizing false positives.