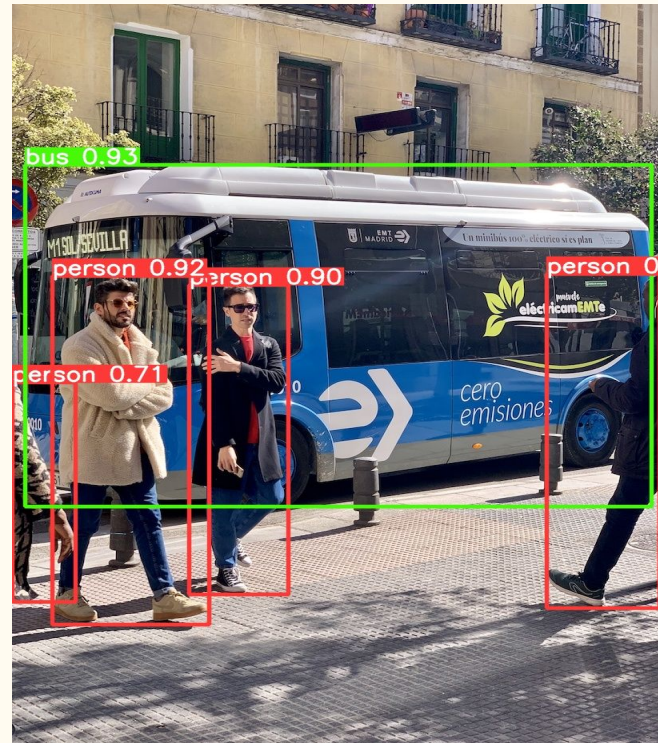


Object Detection on COCO Data

By - Ashwani Rajan
Chahak Sethi

Object Detection

- Identify and locate objects within an image/video
- Project Goals
 - Perform object detection on Common Objects in Context (COCO) Dataset.
 - Fine-tuning and comparing state-of-the-art object detection models for a custom dataset.



Dataset - COCO

- Used Common objects in context (COCO)-2017 validation set
 - 5k+ Images and annotations for 80 common object categories.
 - Filtered Labels: [car, truck, bus, bicycle] (732 images)

Train	Test
500 images	232 images

- Trained different architectures to identify objects:
 - Faster RCNN
 - Pretrained YOLO architecture
 - YOLO with further trained weights

Faster RCNN

- General Workflow



- ResNet-50-FPN backbone
- Hyperparameters:
 - Learning rate = 0.005
 - Optimizer = SGD
 - Trained for 5 epochs
- Imbalance in Car and Bus class.

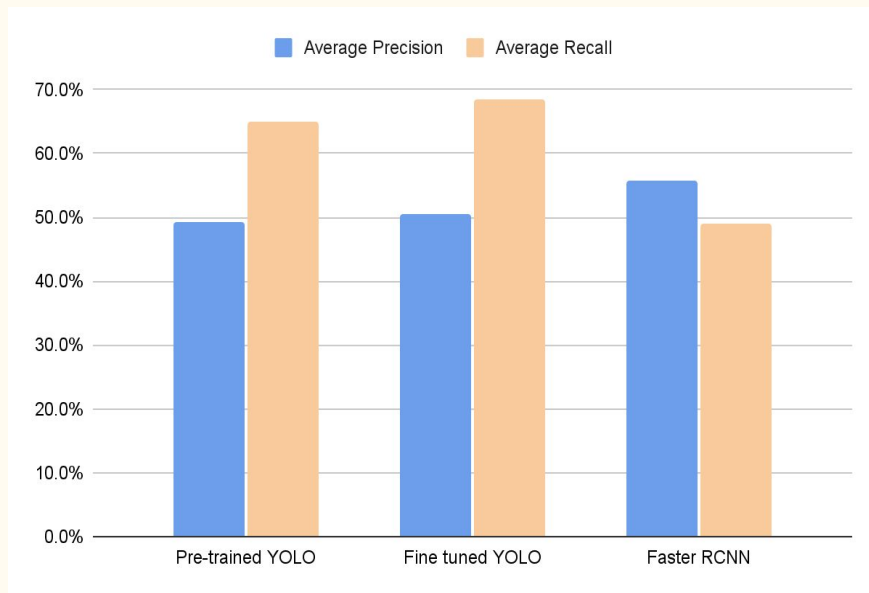
Classes	Precision	Recall	f1-score
Car	0.51	0.70	0.59
Bus	0.53	0.72	0.61
Truck	0.38	0.32	0.35
Bicycle	0.35	0.60	0.44
Wgt. Avg	0.48	0.64	0.54

YOLO vs Faster RCNN

YOLO	Faster RCNN
Very Fast (45 Frames per second - better than real-time)	Slow - cannot be used for real-time prediction
1-step architecture - Does Classification and bounding box regression at the same time	2-step architecture - first classifies and then finds in the object
Difficulty predicting objects that are small and close to each other	Easily detect small objects

Performance Comparison

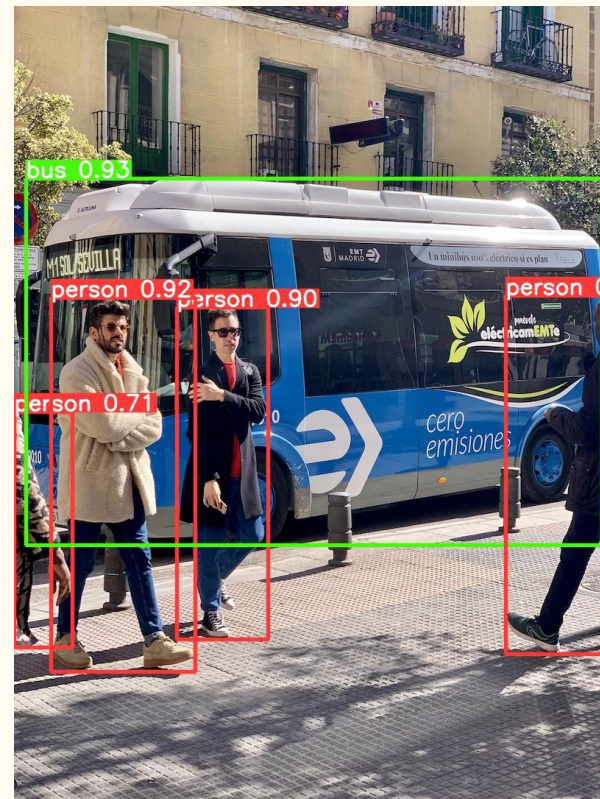
- Faster RCNN has higher precision and lower recall than YOLO
- Higher precision means that an algorithm returns more relevant results than irrelevant ones and high recall means that an algorithm returns most of the relevant results



Results



Faster-RCNN



YOLO-v5

What went well/didn't go well?

- We were able to explore faster RCNN more than the YOLO architecture.
- Detailed tutorials available to deal with COCO and Faster-RCNN (Fiftyone Library)
- Faster RCNN and YOLO are very complex architectures, training took a lot of time and resources
- Trained for only 4 classes out of a total of 80

Conclusion

- Faster RCNN has a better average precision whereas YOLO gives a better average recall. YOLO trains faster compared to Faster RCNN.

THANK YOU