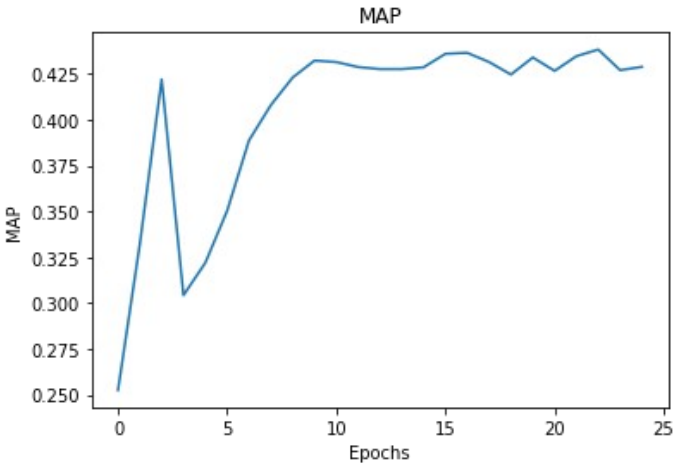
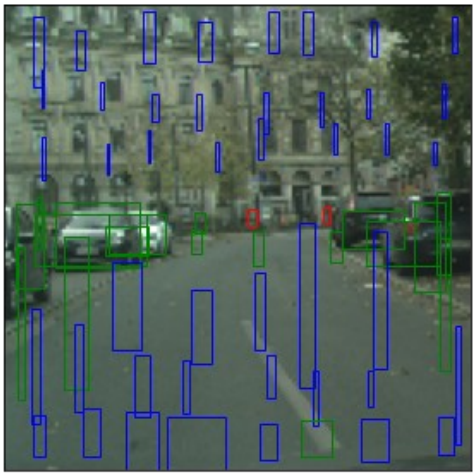
Assignment 2b

Nagarakshith Makam Sreenivasulu

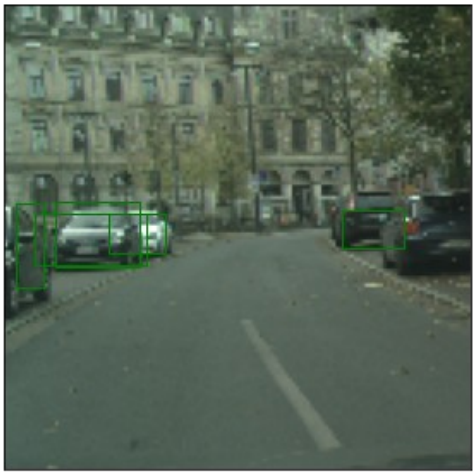
7.5)



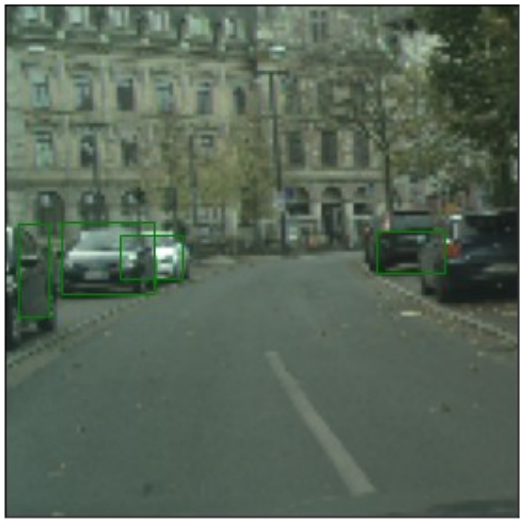
7.6)



Network prediction

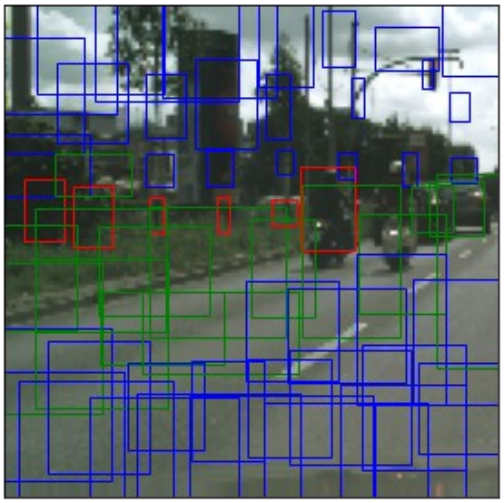


After thresholding with 0.3 confidence score

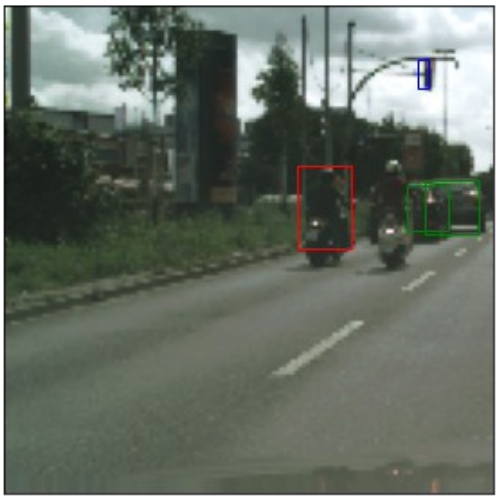


After Non Max Supression

7.7) The inference is done in the last epoch during training

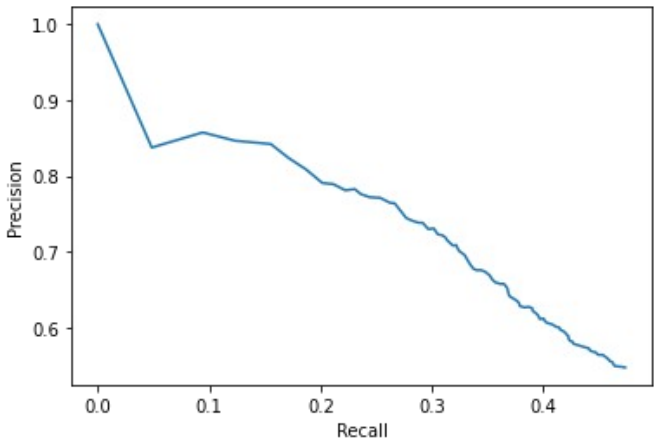


Network prediction

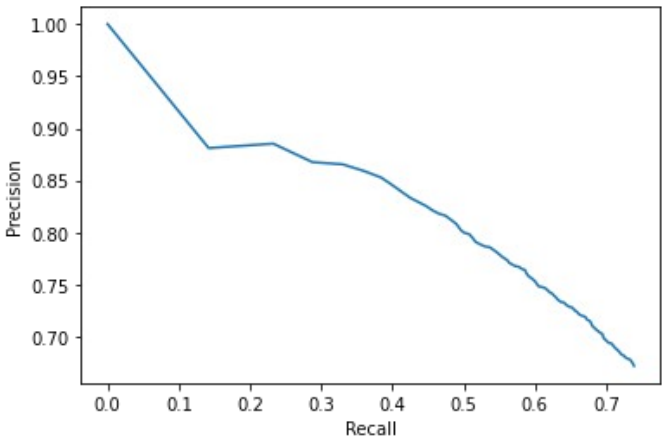


After thresholding with 0.3 confidence score

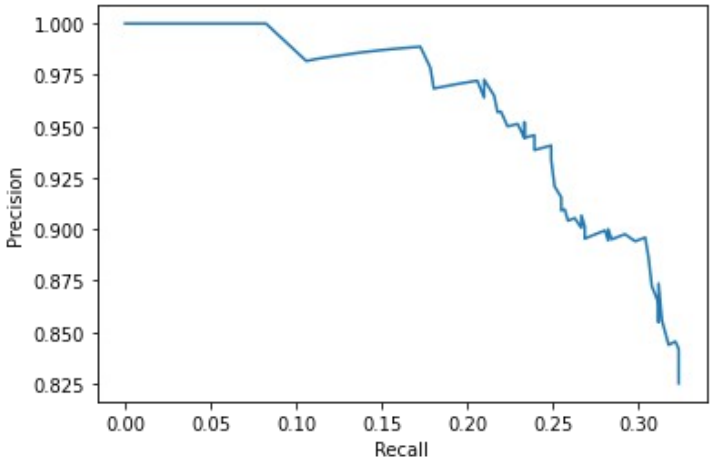
After Non Max Supression



Pedestrian Class



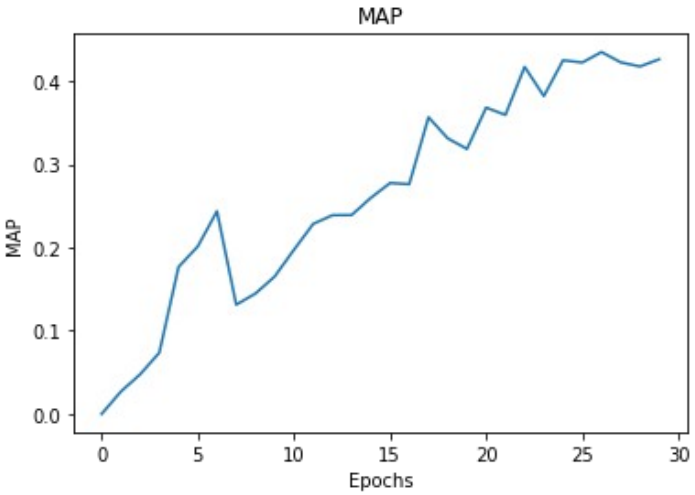
Car Class



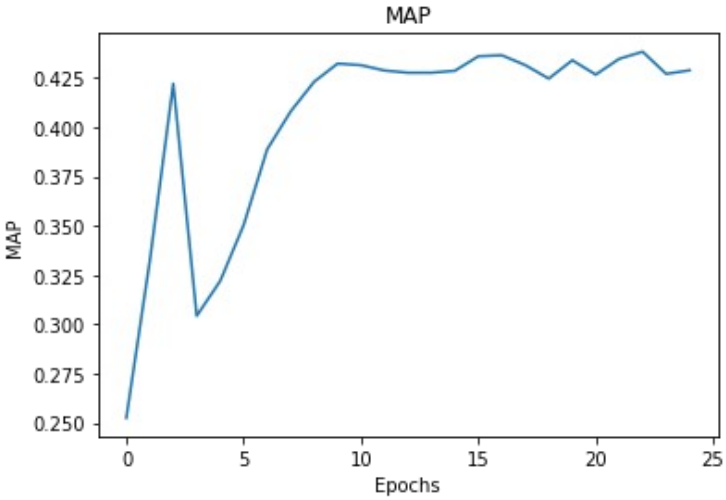
Traffic Light class

MAP achieved is : **0.429**

7.8) Detecting the traffic lights was a huge challenge. When the loss function as described was used no traffic lights were detected. Hence the object detection was decoupled from the localization in the loss function without considering the IOU. This did not affect the MAP much but slightly induced a localization error. I also removed the weight decay.



MAP with the given Loss function

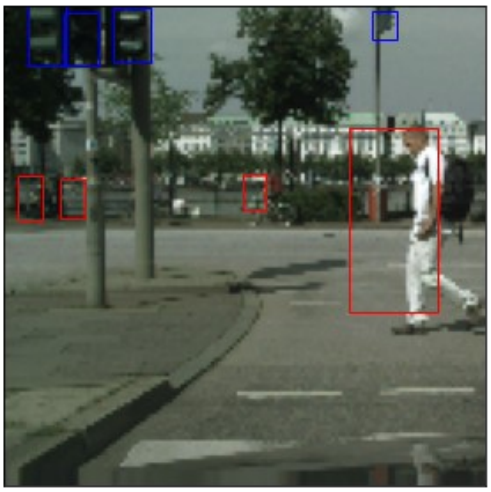
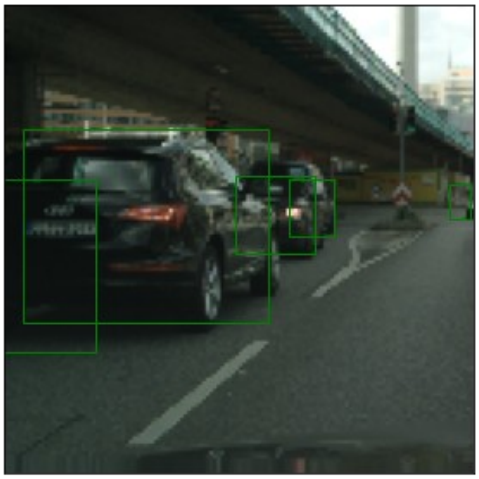


MAP with the modified Loss function

The reason for not much impact on MAP is because using the given loss function the AP for the pedestrian and car is higher whereas for traffic light class it was 0 but with the modified loss function the AP for car and pedestrian are affected due to localization errors but there was increase in the traffic light AP.

Class	Given Loss function	Modified Loss function
Pedestrian AP	0.596	0.357
Traffic Light AP	0	0.311
Car AP	0.426	0.618
MAP	0.429	0.428

Some illustartions

Detection with the given Loss	Detection with the modified Loss
	
	
	

To improve the performance of the network:

1) Since there is dataset imbalance we can try to penalize the loss for traffic light less compared to the car or pedestrian class.

2) Using multiple anchor boxes can help object assignment as in the last image in the table the person is classified as car and the aspect ratio is not quite right.

3) Using two or more bounding boxes per grid cell might also be hepful.