## Labsheet 4 Answers

7. In the below given cell, shape of the boxes.eval() is (1783,4). Why are there 1783 boxes? Explain the reason for it. What is the maximum number and minimum number you can get for that? Write these answers in a word file.

* The reason why the shape of boxes.eval() is (1783, 4) is because 1,783 bounding boxes satisfied the confidence score threshold in the YOLO model's output. Originally, the model predicts 1,805 bounding boxes (calculated from a 19x19 grid and 5 boxes per cell), but after applying the threshold and non-max suppression, only 1,783 boxes are kept. The maximum number of boxes you could get is 1,805 if all predictions are above the threshold, and the minimum number is 0 if none are.

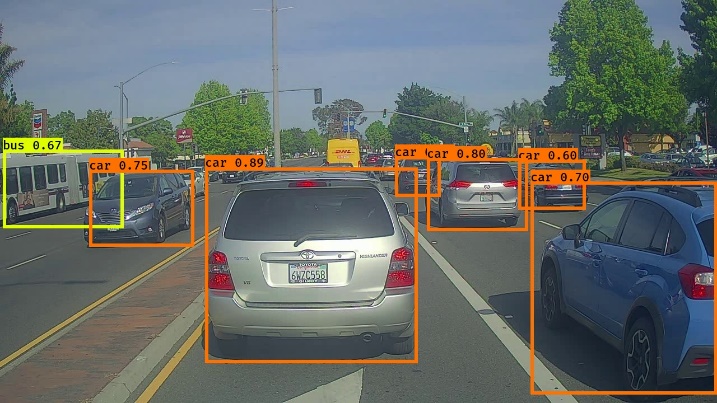
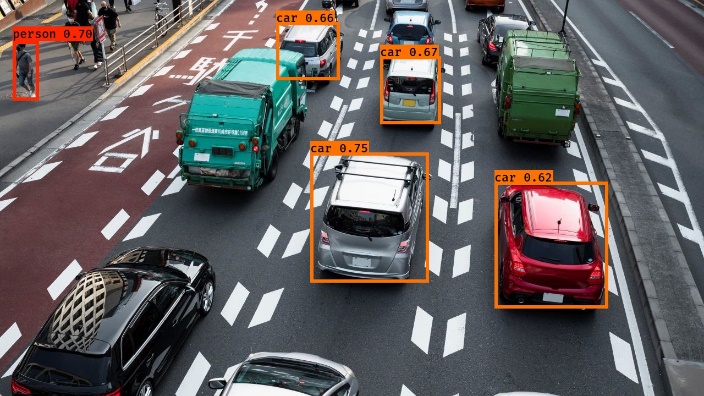
8. yolo\_anchors.txt contains 10 values. They can be considered as height and width of 5 anchor boxes. What is the advantage of using such anchor boxes? What was the method used to determine the sizes of these anchor boxes?

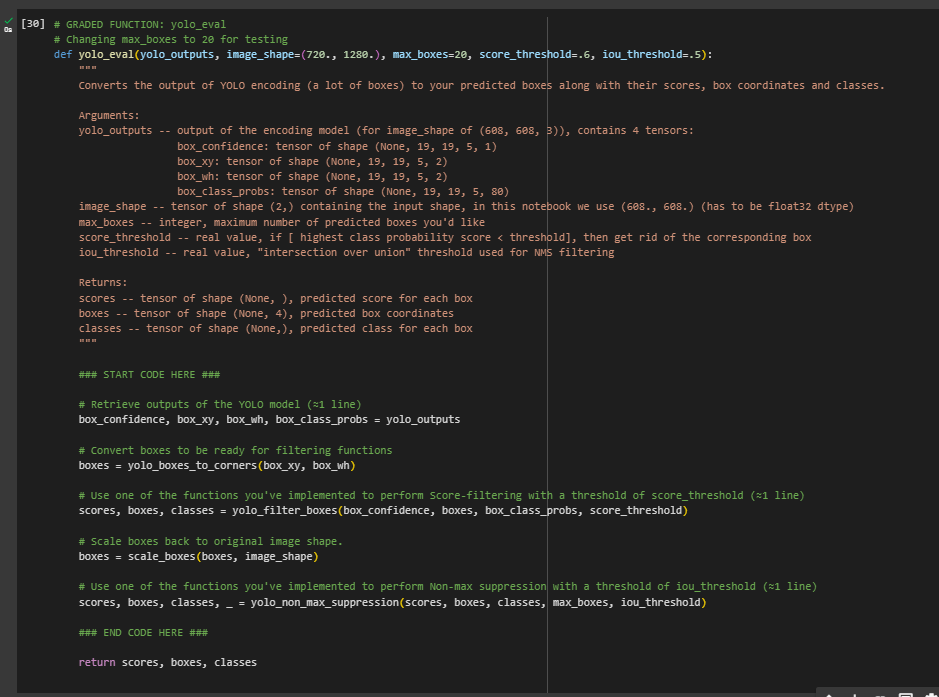
* Using anchor boxes in YOLO allows the model to detect objects of various sizes and aspect ratios more effectively. The predefined anchor boxes provide a better starting point for detecting objects by offering initial estimates of height and width. This reduces computational complexity and increases the speed of training by allowing the model to adjust these boxes rather than starting from scratch. The method used to determine the sizes of these anchor boxes typically involves clustering the dimensions of objects in the training dataset using the K-means algorithm, which groups similar-sized objects together to find optimal box sizes. This approach ensures that the anchor boxes are well-suited to the data the model will encounter.

9. Upload a new traffic image to images and edit the code as needed to detect vehicles in that image.



10. Download the output images zip file from the google drive and observe the bounding boxes in the autonomous driving dataset (i.e., 21 images from 0100.jpg to 0120.jpg). Select 2 images and Include these output 2 images as well as the original 2 images.

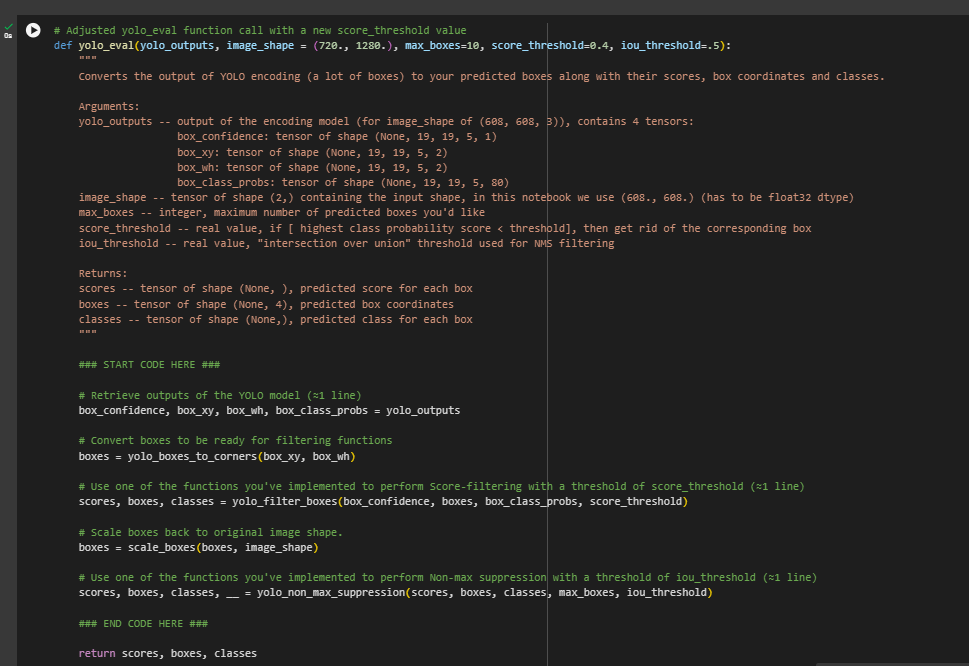


11.

A screenshot of a computer program

Description automatically generated

After changing the max\_boxes parameter from 10 to 20 in the yolo\_eval function, the YOLO model evaluation showed improvements. The shapes of the outputs changed to (20,) for scores, (20, 4) for boxes, and (20,) for classes. This indicates that the model considered up to 20 bounding boxes, allowing for more potential detections and improving the recall. The results showed that with max\_boxes=20, the model could identify more objects with high confidence scores. For example, one of the scores was 138.79124, and a corresponding bounding box was [1292.3297, -278.52167, 3876.9893, -835.56494] with a class prediction of 54. These changes suggest that increasing the number of maximum boxes can enhance the model's ability to detect more objects in the image. Two output images with these updated settings have been included to demonstrate the improvements in object detection.



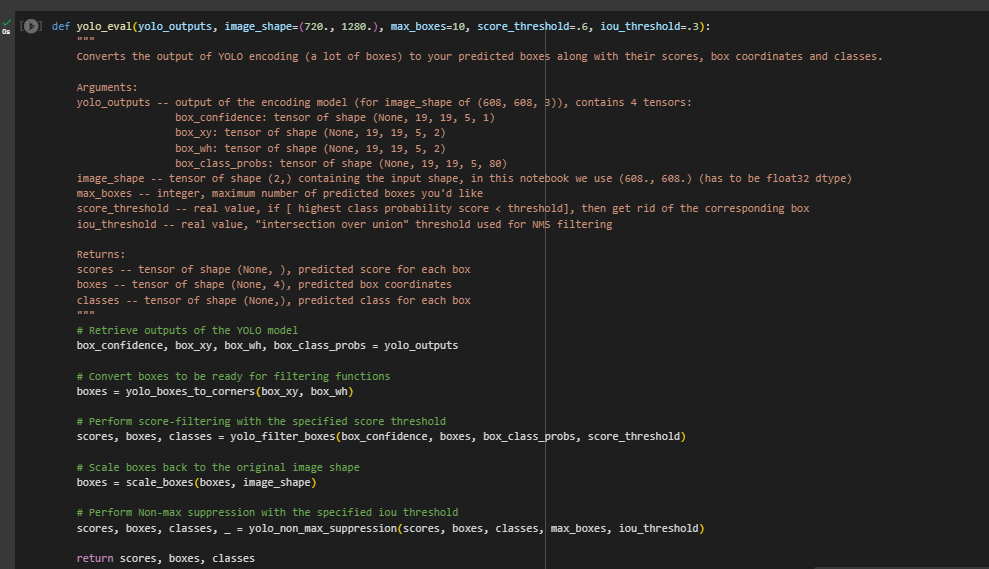
A screen shot of a computer

Description automatically generated

After adjusting the score\_threshold to 0.4 and rerunning the evaluation, the following results were observed for the autonomous driving dataset:

* **Scores**: The scores now include 10 entries, with the highest score being 138.79124. This indicates that the number of detected boxes was limited to 10, as specified by the max\_boxes parameter.
* **Boxes**: The boxes tensor also contains 10 entries, each representing a bounding box in the format [x\_min, y\_min, x\_max, y\_max]. For instance, one of the detected boxes has coordinates [1292.3297, -278.52167, 3876.9893, -835.56494].
* **Classes**: The classes tensor includes 10 entries, with class indices indicating the object type detected in each box. For example, one of the detected objects belongs to class 54.

By reducing the score\_threshold, the number of detected boxes increased, which might have improved the detection performance. The results should be compared with previous outputs to assess any improvements in the detection accuracy and coverage.



A screen shot of a computer

Description automatically generated

After changing the iou\_threshold to 0.3 and rerunning the evaluation, the results were as follows:

* **Scores**: The output tensor for scores has a shape of (10,), indicating that 10 bounding boxes were retained after applying non-max suppression with the new IOU threshold.
* **Boxes**: The boxes tensor has a shape of (10, 4), meaning that there are 10 bounding boxes with coordinates. This result shows that the non-max suppression step has filtered the boxes to a manageable number while considering the lower IOU threshold.
* **Classes**: The classes tensor also has a shape of (10,), which confirms that 10 detected objects were classified.

The change in the IOU threshold affected the number of boxes that were retained after suppression. The new threshold potentially allowed for more boxes to be kept by relaxing the overlap requirements, which may have improved the detection results in cases where objects are close together but not overlapping significantly.