

**SE4050**

**Deep Learning**

**4th Year, 1st Semester**

**Lab 04**Submitted to

Sri Lanka Institute of Information Technology

**IT21166488**

In partial fulfillment of the requirements for the

Bachelor of Science Special Honors Degree in Information Technology

30/08/2024

1. 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.

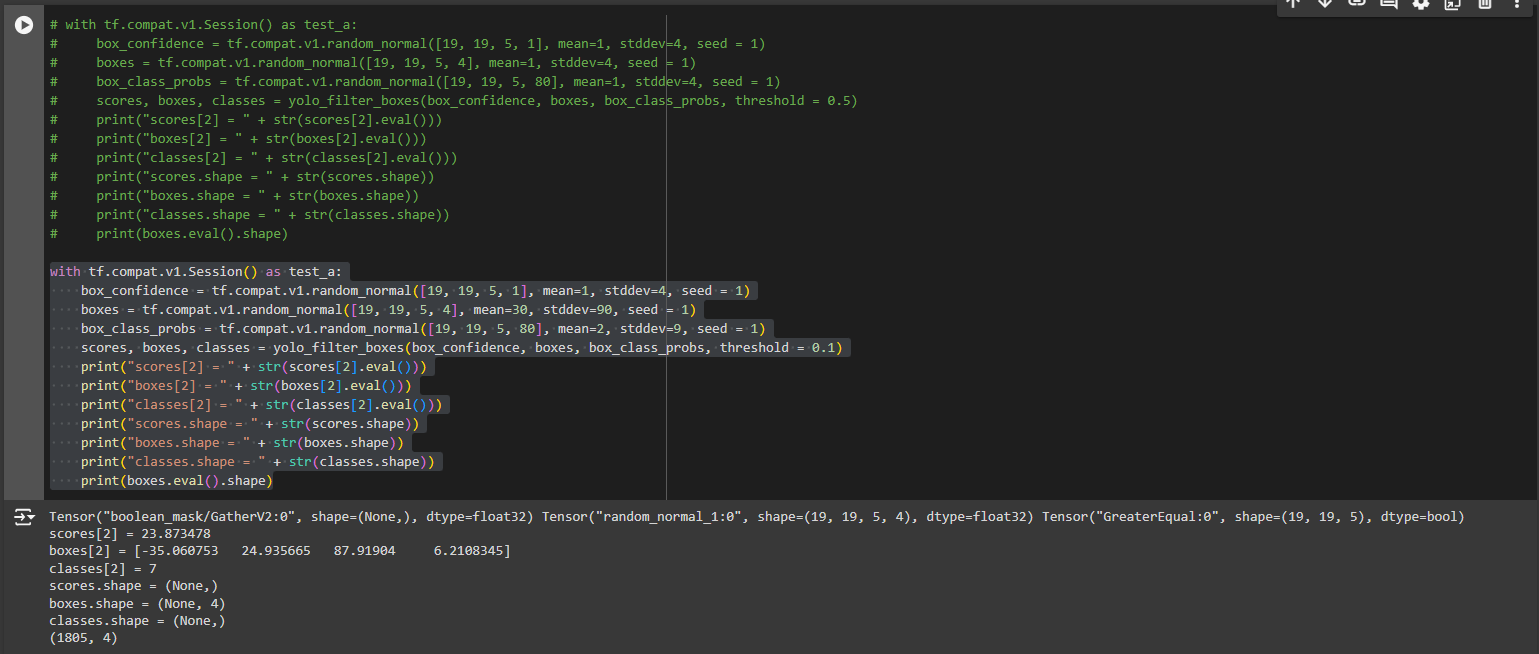
The 1783 boxes are the result of filtering out the original set of 1805 predicted boxes. The YOLO model divides the image into a grid of 19x19 cells, with each cell predicting 5 boxes. This gives a total of 1805 boxes. After applying a confidence threshold, which removes boxes with low confidence, 1783 boxes remain.

What is the maximum number of boxes?

The maximum number of boxes you can get is 1805. This number comes from multiplying the 19x19 grid cells by 5 anchor boxes per cell.

What is the minimum number of boxes?

The minimum number of boxes you can get is 4. This could happen if only one box is detected with high confidence.



1. 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?

Anchor boxes allow the YOLO model to detect objects of various shapes and sizes more effectively. By using predefined anchor boxes with different heights and widths, the model can better predict bounding boxes that match the actual objects in an image. This improves the accuracy of object detection, especially when objects have different aspect ratios.

1. What was the method used to determine the sizes of these anchor boxes?

The sizes of the anchor boxes were determined using a method called **k-means clustering**. This technique groups the bounding boxes from a dataset into clusters based on their width and height. The center points of these clusters are chosen as the sizes for the anchor boxes, making them better suited to the objects in the images.

1. Write what you observe regarding correctly detected objects, incorrectly detected objects, undetected objects and incorrect bounding boxes.

|  |  |
| --- | --- |
| Original Image | Output Image |
|  | Cars on a road with a mountain in the background  Description automatically generated |
| A street with a green light  Description automatically generated | A street with a green light  Description automatically generated |
| A traffic jam on a street  Description automatically generated | A traffic jam on a street  Description automatically generated |
| A traffic jam on a highway  Description automatically generated | A traffic jam on a highway  Description automatically generated |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| Aspect | Image 1 (Rainy Road) | Image 2 (Urban Street) | Image 3 (Busy Intersection) | Image 3 (Highway Traffic) |
| Correctly detected objects | 3 cars | 1 bus | 6 cars | 10 cars |
| Incorrectly detected objects | Nonvisible | Nonvisible | Nonvisible | Nonvisible |
| Undetected objects | Road signs, guardrails | Cars, traffic lights, pedestrians | Many cars, motorcycles, auto-rickshaws, people | Many cars |
| Incorrect bounding boxes | Slightly large, but accurate | Accurate for the bus | Mostly accurate for detected cars | Generally accurate |
| Overall performance | Good in simple scene | Poor in urban scene | Poor in complex scene | Moderate in dense traffic |

1. Adjusting parameters like max\_boxes, score\_threshold, and iou\_threshold of the yolo\_eval function can potentially address the limitations.

Change the max\_boxes [integer value]

A screen shot of a computer

Description automatically generated

Figure 1: with max box 10

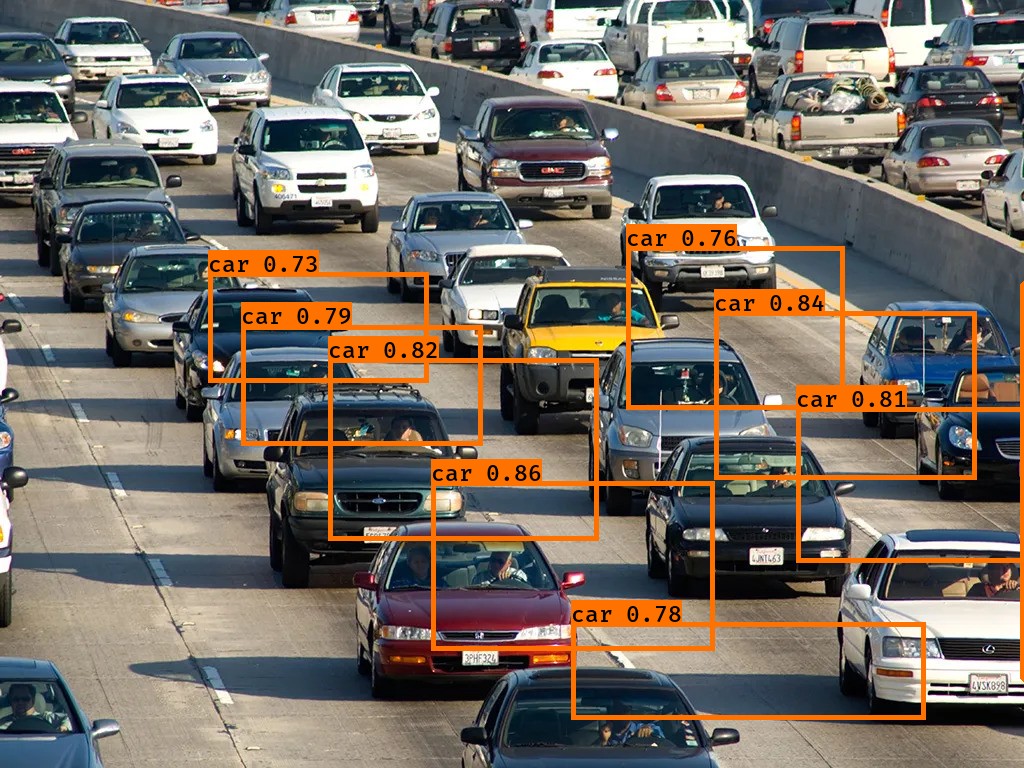


Figure 2: with max box 10

A screenshot of a computer program

Description automatically generated

Figure 3: with max box 100

A traffic jam on a highway

Description automatically generated

Figure 4: with max box 100

Overall: Not Much Effective

Also after changing score threshold, and iou\_threshold values, it is same as before result.