# IT21479632 - Wijekoon W.M.K.G.S.S.B

# Deep Learning - SE4050

# Lab 04

## Question 07

The 1,783 boxes in the boxes.eval() output represent the number of bounding boxes that passed the thresholding criteria set in the YOLO filtering function. The input tensors box\_confidence, boxes, and box\_class\_probs have shapes that correspond to a 19x19 grid of cells, with 5 anchor boxes per grid cell. The yolo\_filter\_boxes function filters these boxes based on their confidence scores and class probabilities, retaining only those that exceed a certain threshold. In this case, the threshold is set to 0.5, meaning only boxes with confidence scores above this value are kept. The total number of possible boxes before filtering is 19 \* 19 \* 5 = 1,805. The resulting 1,783 boxes indicate that nearly all of the initial boxes met the confidence threshold. The maximum possible number of boxes that could be retained is 1,805 if all boxes exceed the threshold, while the minimum is 0 if none do. Changing the values of mean and stddev in the random tensor generation, as well as the threshold value, can affect the distribution of confidence scores, thereby influencing how many boxes pass the threshold.

**Maximum Number**: The maximum number of boxes can be the total number of boxes before filtering, which is 19 \* 19 \* 5 = 1,805. This would happen if all the boxes had confidence scores above the threshold.

**Minimum Number**: The minimum number of boxes would be 0 if all the boxes had confidence scores below the threshold.

**MEAN and STDDEV :**Increasing the mean and standard deviation of the **box confidence** causes more boxes to pass the filtering process, and lowering said values causes less boxes to pass.

Overall : Changing the **mean** and **stddev** values in the **tf.random.normal** function will affect the distribution of the generated confidence scores, box coordinates, and class probabilities. This, in turn, will influence how many boxes pass the threshold. A higher threshold will result in fewer boxes passing the filtering criteria, while a lower threshold will allow more boxes to pass.

## Question 08

**Advantages**

The primary advantage of using anchor boxes is that they allow the model to predict multiple objects of different sizes and aspect ratios in a single grid cell. By having a variety of anchor boxes, the model can better capture objects of varying shapes and sizes, which improves the accuracy of the detection. This approach also reduces the computational complexity since the model doesn't need to predict the exact shape and size of each object from scratch but rather adjusts the predefined anchor boxes to fit the objects.

**Method**

The sizes of the anchor boxes are typically determined using a clustering algorithm, such as K-means, on the bounding box dimensions of objects in the training dataset. The algorithm clusters the bounding box dimensions into a predefined number of groups (e.g., 5 for 5 anchor boxes) and computes the centroids of these clusters. These centroids represent the optimal height and width for the anchor boxes, which are then used during training and inference. This method ensures that the anchor boxes are well-suited to the specific dataset, leading to better performance in object detection.

## Question 10

**Image 0102:**

With the initial parameters given to the model, the image does not render the other vehicles, signs or traffic lights present in the image. The only thing that the model does identify is the bus present in it.

A street with a green light

Description automatically generated

A street with a traffic light

Description automatically generated

**Image 0116:**

Only 1 traffic light and 1 car were detected. Many of identifiable objects in the image were not detected.

A crosswalk with red lights

Description automatically generated

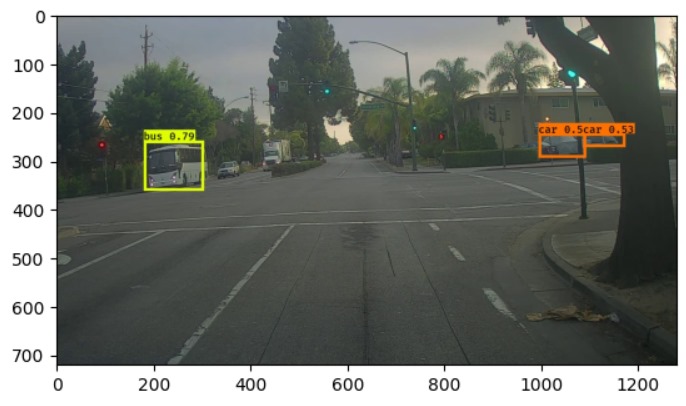
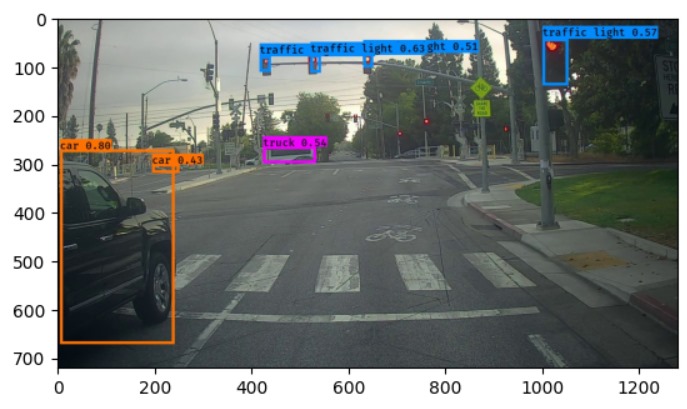
A crosswalk with traffic lights

Description automatically generated

## Question 11

* Increasing the max boxes for some images made them render more boxes after the filtering process. However, for most images with only a few items to identify. It didn’t affect them.  
    
  A street with a traffic light

  Description automatically generated  
    
  A crosswalk with traffic lights

  Description automatically generated
* Increasing the score\_threshold makes it such that the model predict/identify less things in the image. Reducing the threshold makes it identify more things however, they are most of the time not accurate.  
    
    
    
  
* Increasing the iou\_threshold makes the boxes that identify the images larger in size, decreasing the iou threshold makes the boes that identify the the images smaller in size.  
    
    
  A street with a traffic light

  Description automatically generated  
    
  A crosswalk with traffic lights

  Description automatically generated