# Question 7

YOLO Model Box Filtering Explanation

In this section, we delve into the output processing of a YOLO (You Only Look Once) model, with a specific focus on how bounding boxes are filtered based on confidence scores and class probabilities. The provided code snippet demonstrates the use of TensorFlow in filtering these bounding boxes, which is a crucial step in a typical object detection pipeline.

The YOLO model divides an input image into a grid, with each cell in the grid predicting multiple bounding boxes, each associated with an object confidence score and class probabilities. In the code example, the image is divided into a 19x19 grid, and each cell predicts 5 bounding boxes. This results in a total of 19 x 19 x 5 = 1805 bounding boxes. The yolo\_filter\_boxes function is then used to filter these boxes by calculating a combined score, which is the product of the box confidence and the highest class probability for each box. Only the boxes with a combined score greater than a threshold of 0.5 are retained. In this instance, 1783 boxes passed this threshold, indicating that 22 boxes had scores below 0.5 and were filtered out.

**Maximum Number:** The maximum possible number of boxes that could pass the filtering process is 1805, which represents all the initial boxes generated. This would occur if all boxes had a combined score above the 0.5 threshold.

**Minimum Number:** The minimum number of boxes that could pass the filtering process is zero. This would happen if all boxes had combined scores below the 0.5 threshold, resulting in the removal of all boxes.

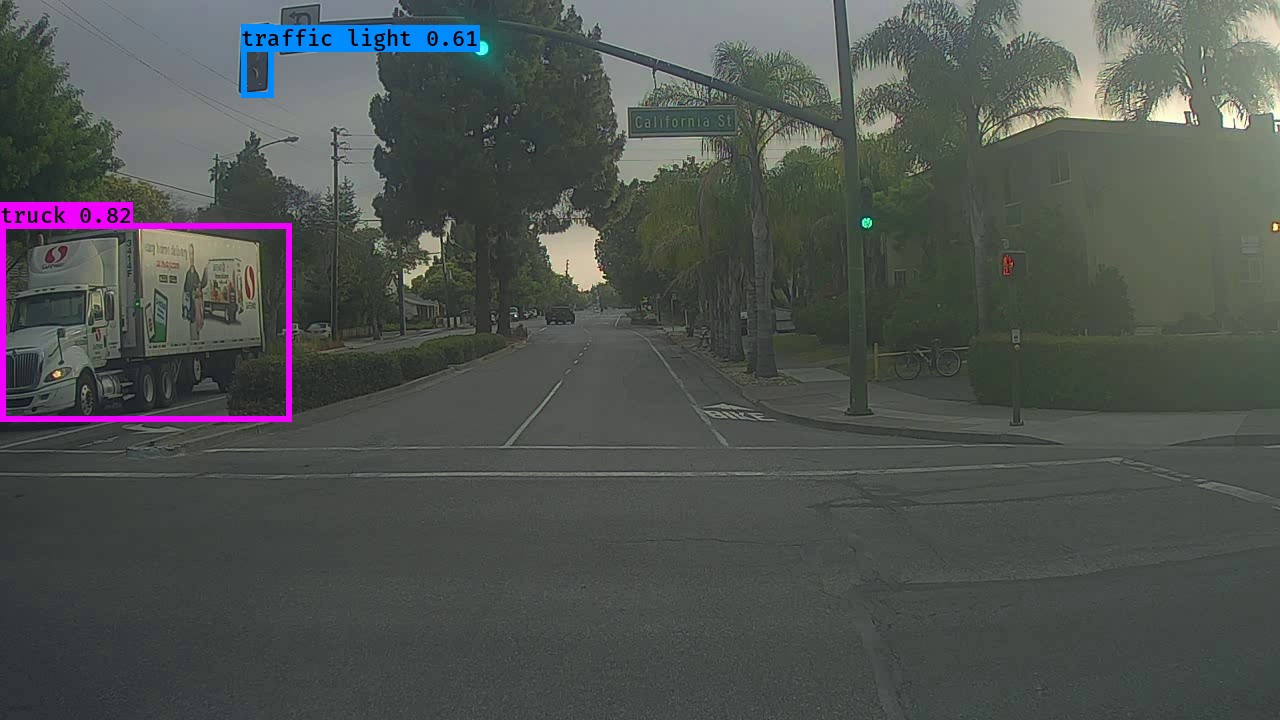
# Question 8

Understanding Anchor Boxes in YOLO

Anchor boxes enable the detection of objects with varying shapes and sizes by predefining potential bounding boxes with different scales and aspect ratios. This approach enhances detection accuracy, particularly in images containing multiple objects of different sizes. By using anchor boxes, the model doesn't need to independently learn the size and aspect ratio for each object, thereby reducing the complexity and improving the efficiency of the network.

The sizes of anchor boxes are typically determined using a clustering method on the training dataset, such as k-means clustering. This method involves grouping the ground truth bounding boxes based on their width and height to minimize the Intersection over Union (IoU) loss, rather than using traditional Euclidean distance. The number of clusters (k) is a tunable hyperparameter, chosen based on the diversity of object sizes within the dataset.

# Question 10

Image 0104 Observations

* Correctly Detected Objects: Traffic Light, Truck
* Incorrectly Detected Objects: there is no incorrect detection
* Undetected Objects: other traffic lights, long-distance car and bicycle
* Incorrect Bounding Boxes: there is no incorrect bounding boxes

Image 0116 Observations

* Correctly Detected Objects: car, traffic light
* Incorrectly Detected Objects: there is no incorrect detection
* Undetected Objects: : other traffic lights, one truck and other long-distance cars
* Incorrect Bounding Boxes there is no incorrect bounding boxes

# Question 11

When the max\_boxes parameter was adjusted in the YOLO model's evaluation settings, no changes were observed in the output for images from the autonomous driving dataset. This suggests that the original max\_boxes setting was likely adequate and did not limit the detection results, indicating that the parameter was not a bottleneck in the model’s performance.

To improve object detection accuracy, the score\_threshold was lowered from 0.6 to 0.3, while keeping the original values for max\_boxes and iou\_threshold. This adjustment resulted in a noticeable enhancement in detection accuracy, as the model was able to identify additional object classes that were previously missed with the higher threshold. This indicates that the original threshold may have been too restrictive, leading to the omission of valid detections. The newly detected objects with the lower threshold demonstrate the model’s improved ability to recognize and classify a broader range of objects within the autonomous driving dataset.