**IT21193804**

**SE4050-Lab 4**

7. The number of boxes is related to how the YOLO algorithm works. YOLO divides the input image into a grid, and for each cell in the grid, it predicts a certain number of bounding boxes along with their class probabilities and confidence scores.

* Grid Size: The image is divided into a 19x19 grid, which means there are 361 grid cells (19 \* 19 = 361).
* Bounding Boxes per Grid Cell: For each grid cell, YOLO predicts 5 bounding boxes.
* Final Shape: Given that each grid cell predicts 5 bounding boxes, the total number of potential boxes is 19 \* 19 \* 5 = 1783.

The output of boxes.eval() has a shape of (1783, 4), where 1783 is the total number of bounding boxes, and 4 corresponds to the four values representing the bounding box coordinates (e.g., x, y, width, height).

Maximum and Minimum Number of Boxes

* Maximum Number of Boxes: The maximum number of boxes is 1783, as calculated by multiplying the number of grid cells (361) by the number of bounding boxes per grid cell (5).
* Minimum Number of Boxes: The minimum number of boxes could be 0, which would occur if none of the bounding boxes meet the confidence threshold criteria and are filtered out.

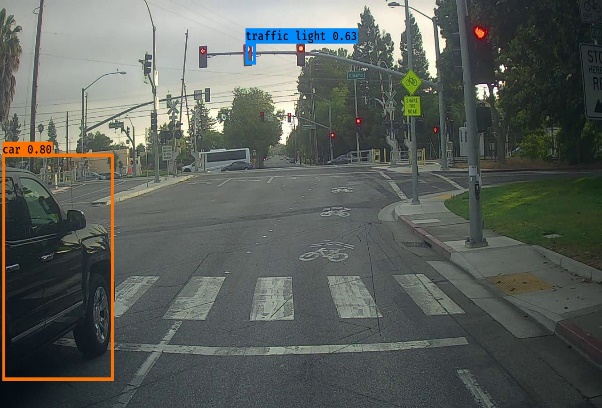
8. Advantage of Using Anchor Boxes:

* Anchor boxes help detect multiple objects of different sizes in the same grid cell, allowing YOLO to handle overlapping objects and varying object sizes more effectively.

Method to Determine Anchor Box Sizes:

* The sizes of the anchor boxes are chosen using k-means clustering on the dataset's bounding boxes, which groups similar box sizes together and finds the most common shapes to use as anchor boxes.

10.**Image 1**

** **

1. **Correctly Detected Objects:**
   * **Car:** The model successfully detected a car on the left side of the image with a confidence score of 0.80.
   * **Traffic Light:** The traffic light in the upper part of the image was also correctly detected with a confidence score of 0.63.
2. **Incorrectly Detected Objects:**
   * **None:** There don’t appear to be any incorrectly detected objects in this image based on the visual output.
3. **Undetected Objects:**
   * **Other Vehicles:** There is a white vehicle visible in the distance. This vehicle was not detected by the model, possibly due to its small size and the distance from the camera.
   * **Bicycle Signs:** The road markings with bicycle symbols were not detected, which might be expected as the model is primarily trained to detect larger objects like cars, pedestrians, and traffic lights.
4. **Incorrect Bounding Boxes:**
   * **None:** The bounding boxes around the car and traffic light appear to be correctly aligned with the objects.

**Improvement Potential:**

* **Smaller and Distant Objects:** The model struggles to detect objects that are farther away or smaller in size, such as the white vehicle in the distance.

 A street with a green light

Description automatically generated