7.

The number 17,834 represents the total number of predicted bounding boxes generated by the YOLO (You Only Look Once) object detection model. This large number of boxes results from the model's grid-based prediction system, where the input image is divided into a grid of cells, and each cell is responsible for predicting multiple bounding boxes. The specific number of boxes depends on the grid size, the number of anchors used for each cell, and the down sampling factor of the network. In this case, the shape of (17,834) likely results from the grid size and the number of anchor boxes. The model evaluates each potential box against the actual objects in the image, which is why so many are generated.

8.

The maximum number of boxes is determined by the grid size multiplied by the number of anchor boxes per grid cell. If you increase the grid size or use more anchors, the number of predicted boxes will increase. Conversely, the number will decrease if you decrease the grid size or use fewer anchors. For instance, if the grid size is 13x13 and each cell predicts 5 anchor boxes, the number of predictions would be 13 \* 13 \* 5 = 845 boxes. However, YOLO typically predicts across multiple scales, which could increase this number significantly. The minimum number of predicted boxes could theoretically be as low as 1, if the grid size is reduced to 1x1 and only one anchor is used.

A road with trees and grass

Description automatically generatedA road with trees and grass

Description automatically generated10.

Correctly detected objects – Car, Fire hydrant

Not detected objects – Cab, Traffic lights

A crosswalk with red lights

Description automatically generated

Correctly detected objects – Cab, Traffic Lights

Not detected objects – Bus, Traffic Lights