

Lab 05

Question 07

Why there are 1783 boxes? – ‘yolo_filter_boxes’ function filters the boxes using the confidence threshold to filter out the low-confidence boxes. This reduces the total number of boxes to 1783 which have confidence scores greater than or equal to 0.5. Also the mean and the stddev can effect the box_confidence , increasing the mean or decreasing the standard deviation will result in higher confidence scores. Also adjusting those will also affect class probability distribution. If we increase the threshold the number of retained boxes will be decreased as the confidence requirement is higher. I changed those values and observed the changes of the number of boxes.

Maximum number of boxes – 1805, maximum number of boxes is determined by the grid size and number of anchors per cell. For a grid with $19 * 19$ cels with 5 anchors each the number of boxes can be calculated by, $19 * 19 * 5 = 1805$.

Minimum number of boxes – 0, minimum number theoretically can be low as 0 if none of the boxes meet the confidence threshold.

Question 08

advantage of using such anchor boxes –

- Anchor boxes improve the efficiency of the YOLO model by reducing the complexity of predicting bounding boxes directly.
- The use of anchor boxes enables the detection of multiple objects in the same grid cell.
- Detect objects with various dimensions by multiple aspect ratios of anchor boxes.

Method used - K-Means clustering

Question 10

Original Image :



Output Image:



In here bus has been correctly identified but traffic lights , truck and jeep have not been correctly identified.

Original Image:



Output Image:

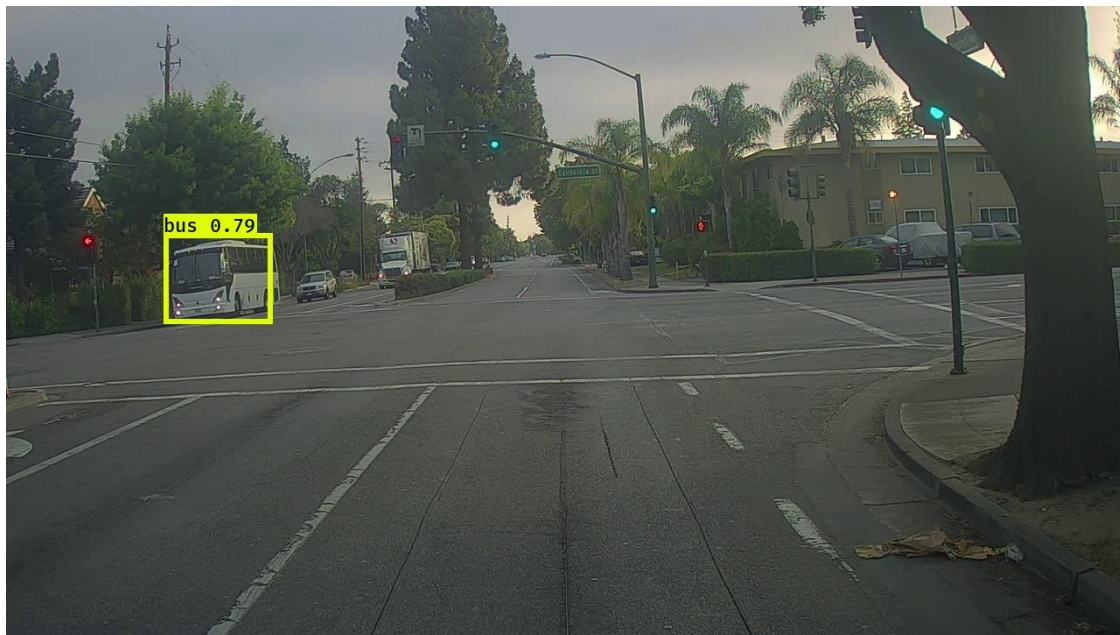


In here cars and other vehicles are not identified.

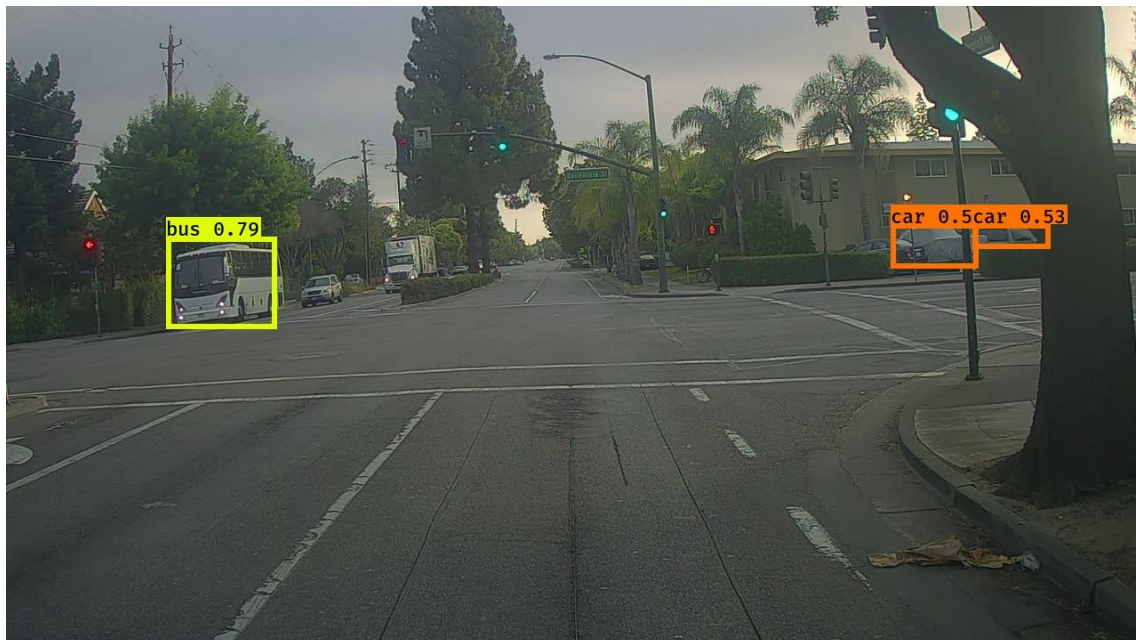
Question 11

Adjusting parameters like `max_boxes`, `score_threshold`, and `iou_threshold` of the `yolo_eval`

Before changing values:



After changing values:



After changing values



Before changing values:





```
# GRADED FUNCTION: yolo_eval
#score threshold default 0.6
def yolo_eval(yolo_outputs, image_shape = (720., 1280.), max_boxes=15, score_threshold=.5, iou_threshold=.4):
    """
    Converts the output of YOLO encoding (a lot of boxes) to your predicted boxes along with their scores, box coordinates and classes.

    Arguments:
    yolo_outputs -- output of the encoding model (for image_shape of (608, 608, 3)), contains 4 tensors:
        box_confidence: tensor of shape (None, 19, 19, 5, 1)
        box_xy: tensor of shape (None, 19, 19, 5, 2)
        box_wh: tensor of shape (None, 19, 19, 5, 2)
        box_class_probs: tensor of shape (None, 19, 19, 5, 80)
    image_shape -- tensor of shape (2,) containing the input shape, in this notebook we use (608., 608.) (has to be a tuple)
    max_boxes -- integer, maximum number of predicted boxes you'd like
    score_threshold -- real value, if [ highest class probability score < threshold], then get rid of the corresponding boxes
    iou_threshold -- real value, "intersection over union" threshold used for NMS filtering

    Returns:
    scores -- tensor of shape (None, ), predicted score for each box
    boxes -- tensor of shape (None, 4), predicted box coordinates
    classes -- tensor of shape (None, ), predicted class for each box
    """
```