Computer Vision - Project 2 Approach

Sasu Alexandru-Cristian, Group 407

June 25, 2024

1 General details

As a means of detecting and tracking cars, only a pre-trained variant of the YOLOv9 model has been used, namely "yolov9e.pt" from the Ultralytics website.

2 Tasks 1 and 2

In order to detect whether a car occupies one of the 10 parking spaces, the YOLO model was used.

For task 2, the interest was only in the last frame of the video.

Before performing inference, approximate coordinates were retained for each parking space, thus delimiting each space as a polygon. Afterwards, because the model had trouble detecting cars that were situated in parking spaces 5-10 (as they were farther away), a rectangular cropping of the last 5 parking spaces was realized. Then, having both the whole image and the cropping, these two images were given to the model in order to perform object detection. After obtaining the detections, only their bounding boxes were retained, and in order to verify if a parking space is occupied by a car, the following check was performed: if the pixel halfway from the center of a car's bounding box to the middle of the bottom line of the car's bounding box is situated inside the polygon associated to that parking space, then that parking space is occupied.

3 Task 3

In order to track the desired car, object detection was performed by the YOLO model at each frame. Thus, in order to be aware of the bounding box of the car in a certain frame, the IoU score was calculated between its previous bounding box that the algorithm was aware of, and all of the bounding boxes of the detections from the current frame. The chosen bounding box for the current frame was decided by picking the one that yielded the highest IoU score. But if the highest IoU score was less than 0.45, then it was decided unfeasible to locate the bounding box of the tracked car in the current frame, and its previously aware-of bounding box would not be updated.

4 Task 4

In order to detect how many cars queue to the traffic light, the YOLO model was once again used.

The interest was only in the last frame of the video.

Firstly, three consecutive rectangular regions starting from below the traffic light and going along the right lane, downwards, were cropped (similarly to the cropping of the last 5 parking spaces from tasks 1 and 2; as once again, the model had difficulty detecting farther away cars). Using these cropped regions, and the whole image, object detection was then performed by the model on these four images and only the bounding boxes of the detections were kept.

Then, by approximately delimiting the entirety of the right lane by using four coordinates and forming a polygon, there were kept only the bounding boxes whose center pixel would fall inside this polygon. But due to the previous three croppings, there may have been generated more than one bounding box for each car, so, in order to have exactly one bounding box for every car, the duplicate bounding boxes were removed by computing how much each bounding box intersects with other different bounding boxes, and so there were removed the boxes situated the most in other boxes.

Having a single bounding box for each car in the right lane, it was then checked how many cars queued to the traffic light. First it was verified if the center pixel of the first car in the right lane from top to bottom was situated inside the polygon determined by four coordinates that was denoted as the start of the queuing area. If there wasn't such a car, the algorithm would determine that the number of queued cars is equal to zero and it would end. But if there was such a car, the algorithm would then check if there was a car whose y coordinate of the top left corner of its bounding box was situated above the y coordinate of the bottom right corner of the first car's bounding box, and in such a case there would be counted two cars queuing to the traffic light. The algorithm would then go on to do the same for the coordinates of the bounding boxes of the third, fourth, and so on cars, until there were no more cars in the right lane, or until the check would fail.