Report

Part 1)

Dataset:

1. a gt.txt file
2. KITTI_17_images.zip
3. outputVideFIle.mp4
4. sort-master.zip
5. TrackEval-master.zip

Task: Multi-object tracking with detection(MOT) to predict the positions of peoples/pedestrians with a given sequence of images provided in the KITTI_17_images.zip file (frames of the video) using SORT and a model Yolov5 to do object tracking and detection respectively.

Approach:

Firstly, Iterates over a sequence of frames of the video with help of a function **`enumerate`** and then each frame is converted from PIL format image to OpenCV format image using **`cv2.cvtColor()`** function. Now, the yolov5 model is used to detect pedestrians using the **`yolov5_detector`.** Object detections are filtered to keep it for the people/pedestrians only and object tracker is updated with the filtered detections using the **` obj_tracker.update()`** function and then tracked objects are drawn on the frame with help of bounding boxes and object_ids using the **`cv2.rectangle()`** and **cv2.putText()** functions respectively. Also tracking information of each frame is written to a predict.txt file using the **`fptr_out.write`** function. A video with tracked pedestrians is saved after combining its tracked frames together using **`final_out.write()`**. Finally, a tracked_OutputVideo.mp4 releases using the function **`final_out.release()`**.