Gedeon Muhawenayo
gmuhawenayo@aimsammi.org
African Masters in Machine Intelligence
Computer Vision II

LAB3: Building an object tracker

Summary:
In this LAB, I got to work on detection in time series data. The lab was about tracking objects in video, so it was about comparing detected instances in two consecutive frames within a series of frames that make video. My approach was to use color coding and track the number of detected objects.

Part A: Detecting objects in frames

After loading frames, I used Detectron2 pre-trained model for detection and segmentation to run and store the predictions for all frames in the video clip. Then I visualized the predictions from a random set of frames to make sure things look correct.

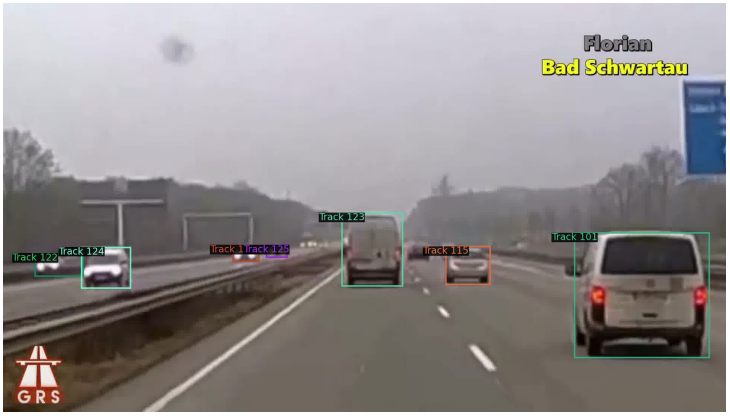
Part B: Tracking objects in pairs of frames

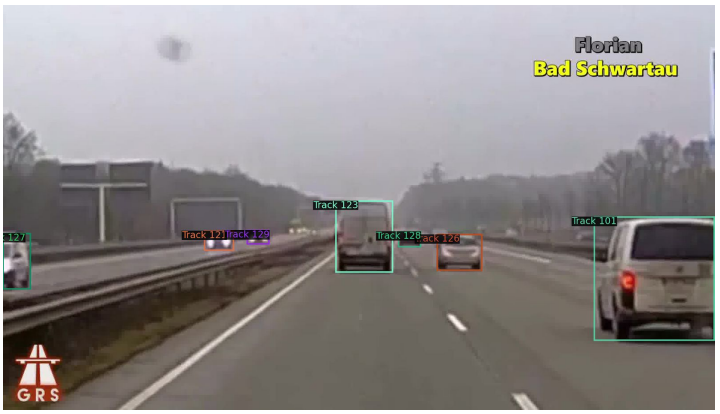
Using matching score and color-coding, I have tracked objects in pairs of frames. I implemented a pairwise tracker for consecutive pairs of frames.

Visualization of pairs of consecutive frames in the start of the given video

Visualization of pairs of consecutive frames in the middle of the given video

Visualization of pairs of consecutive frames in the end of the given video

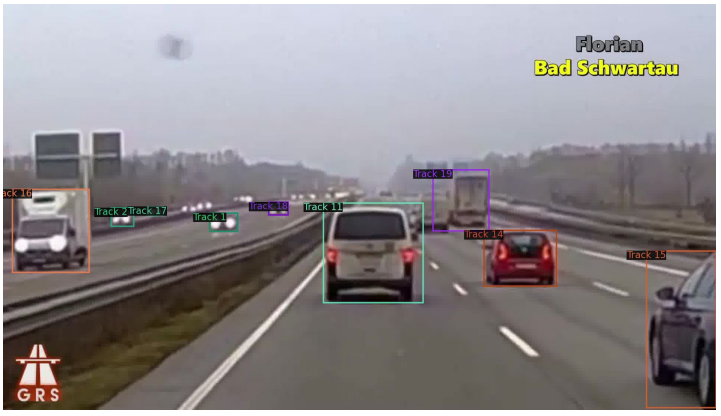


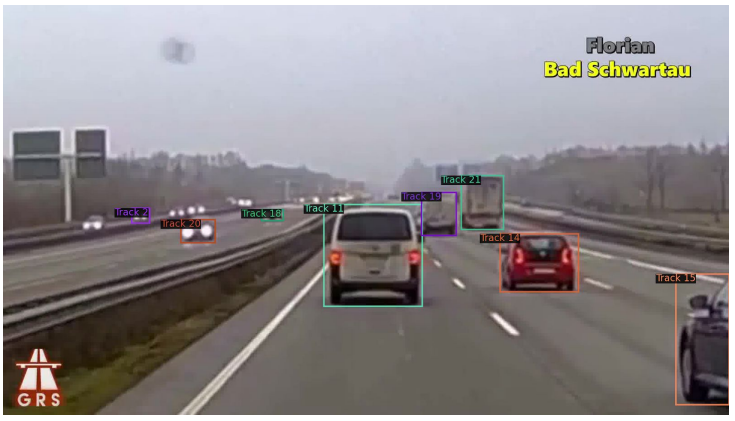


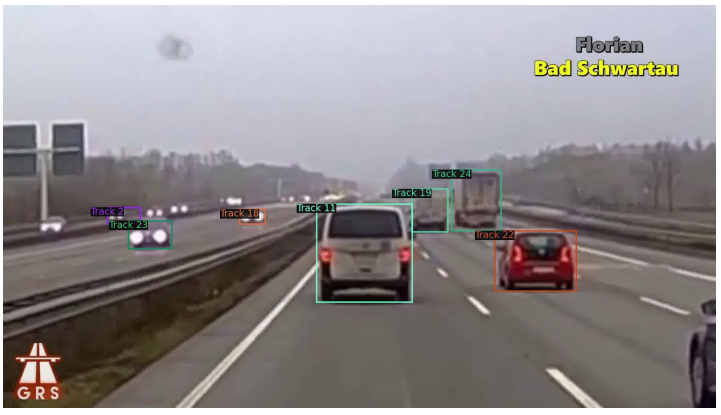
Observation: I have observed that the result of tracking in only two consecutive frames seems to be good! The problem arises when you want to relate instances in frame which are not in consecutive frames. So, I would say that for two frames the model would more accurately track objects from frame1 in frame2 as long as those two frames are consecutive.

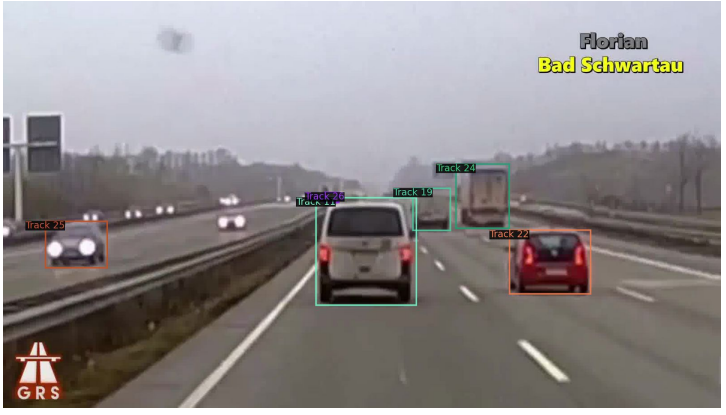
Part C: Tracking objects in videos

Using matching score and color-coding, I have tracked objects in a series of frames, I am using **track count** to show the same object. The following pictures show visualization of frame 2 up to frame 12.









Observation: Tracking in series of frames seems to give good results only if you observe the tracking in pairs, I think this is due to the fact that our matching score is based on comparing instances in two consecutive frames. However when you look at the general performance, this seems to give bad results. Because the same object in frame1 might be detected as another object in frame 3,4, etc.
I think we could compute the matching score by not only looking at the pair of frames, but also we should also look at the series of frames.