

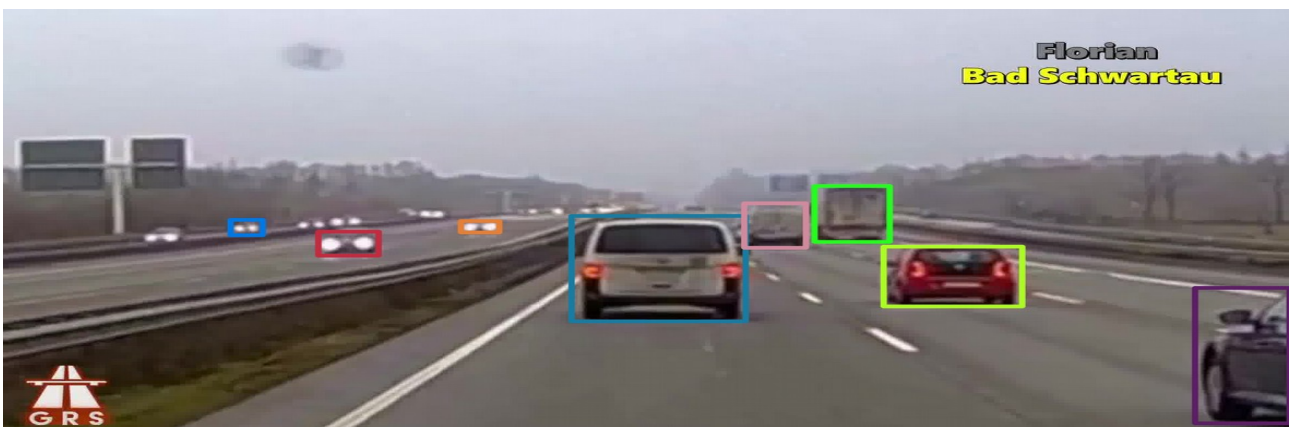
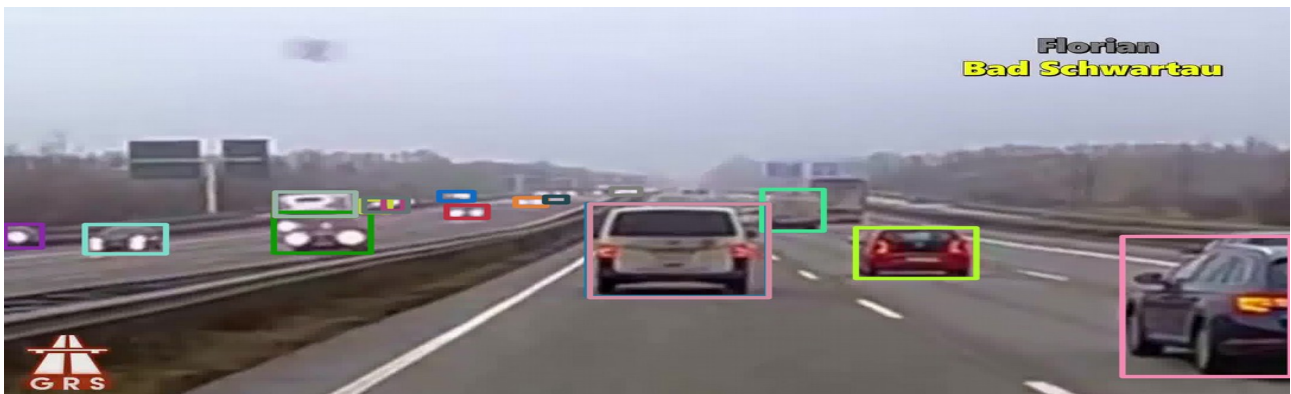
Building an Object Tracker

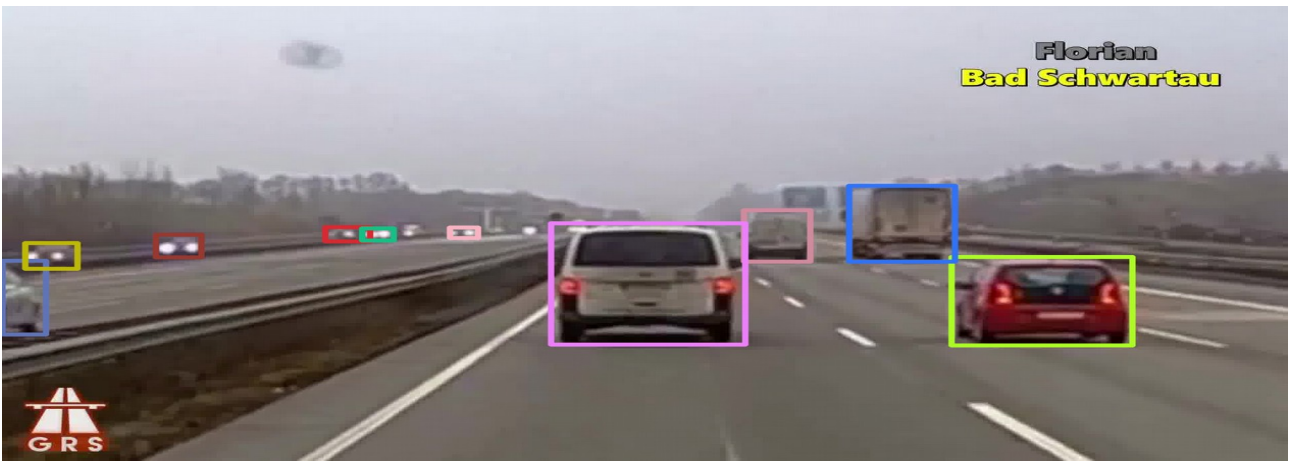
Introduction: This lab is concerned with building a simple object tracker that would detect objects on all frames of a video and then will link the predictions from one frame to the next. We would leverage on the predictions of a D2 model in Lab1.

Methodology: Having met all the necessary requirements in training detectron2 models, the first thing is to download a small video clip of 41 frames, run and store the predictions for all frames in the video clip, to do this, we iterate through the image directory and then append all images into a list. The second step is to track objects in pairs, based on the D2 predictions on each frame of the video done in step 1, we can track objects in time. let's assume we have two frames F_t and F_{t+1} , with their respective predictions P and Q , we would then compute the matching score between each prediction in P and that of Q , which links two predictions if they belong to the same object category AND/ or if their bounding box overlap is high. The third step is to extend what was done in step 2 to perform pairwise links for a time horizon of 10 frames with their different colours for each of the tracks.

Implementations Results and Interpretations.

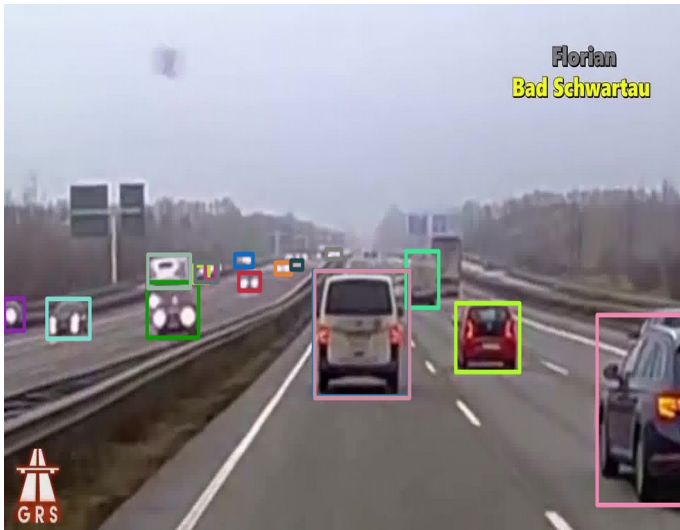
1. visualizations 2-frame tracks (Part B): Picking pairs of consecutive frames



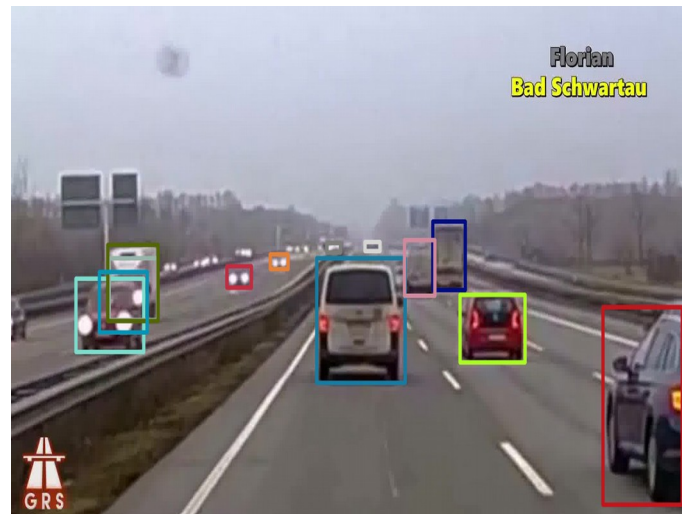


The visualization above represents the 1st frame, 4th frame and 9th frame consecutively, from the first frame using the color coding to indicate the two predictions that correspond to the same track, we observe that the white car had prediction of blue and pink colour, but on the 4th frame, we noticed a prediction of blue, however, on the 9th frame, the white car had a pink colour prediction. This might be because the pink IOU was greater than that of blue.

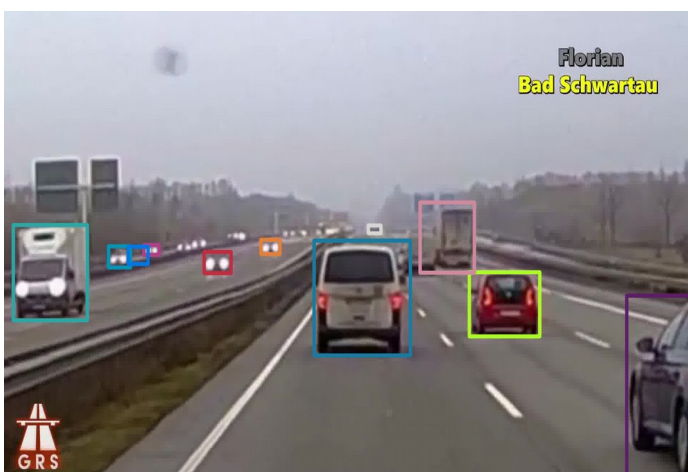
2. visualizations of 10-frame tracks (Part C)



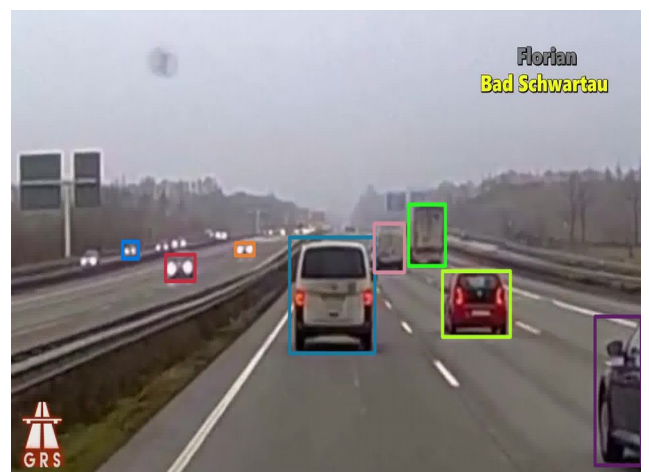
Frame 1



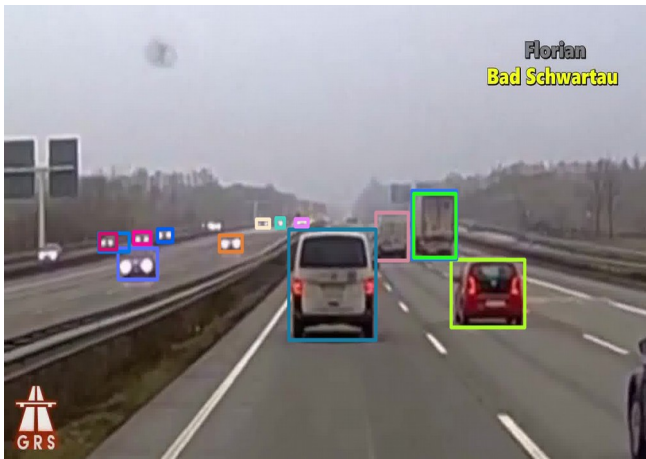
Frame 2



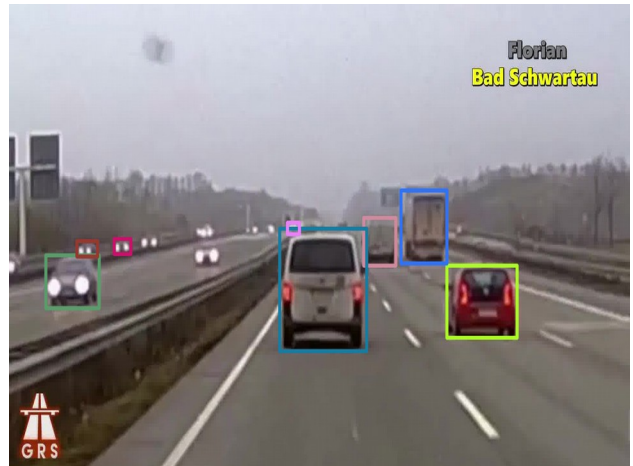
Frame 3



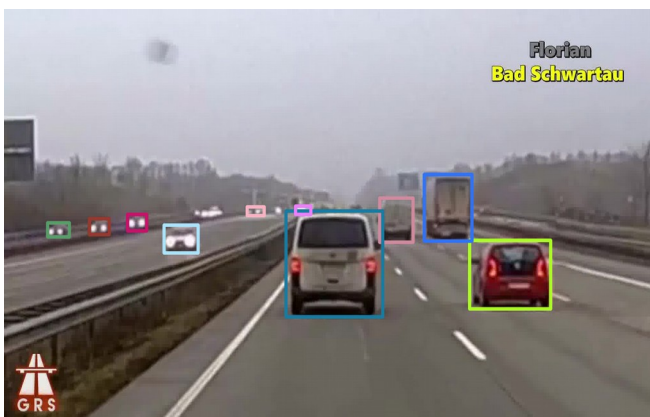
Frame 4



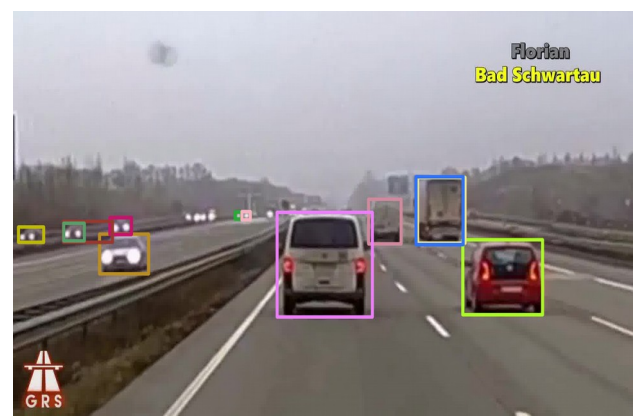
Frame 5



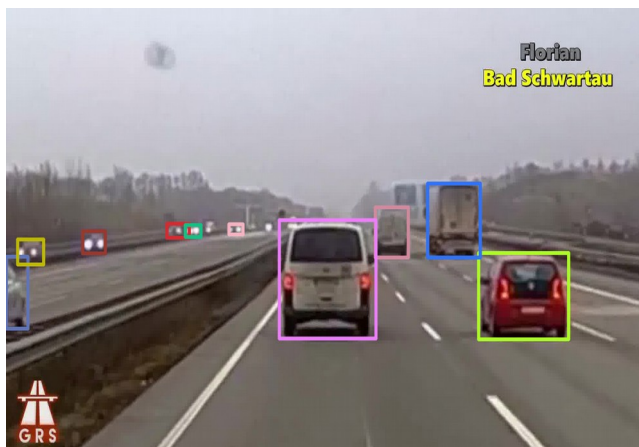
Frame 6



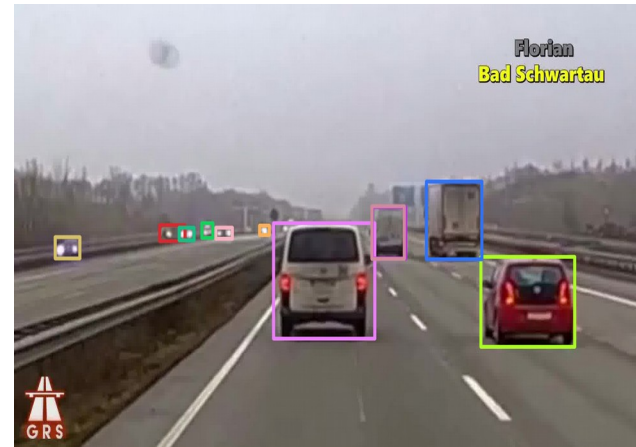
Frame 7



Frame 8



Frame 9



Frame 10

From the visualizations of the first 10 frames consecutively, using the colour coding, we can notice that the prediction belonging to the track of the red car is the lemon green prediction even from the 1st frame down to the 10th frame. However, for the white car, the predictions belonging to that track is light pink, blue and baby pink.

In conclusion, the video making was achieved by using the VideoWriter, VideoWriter_fourcc, imread and resize imported from CV2. We can get that by running the last 2 cells of the notebook.