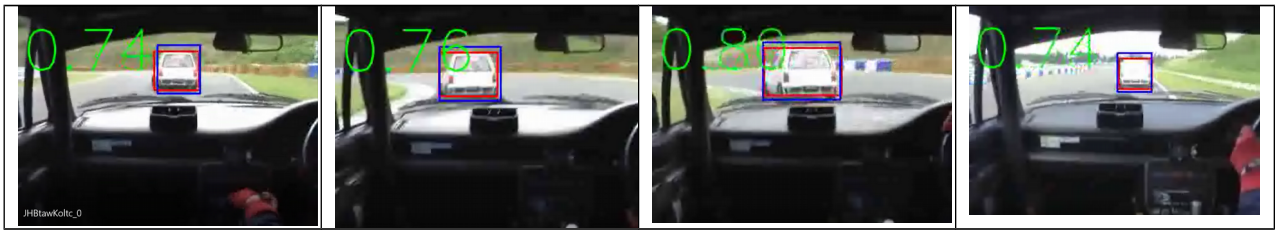
Homework Assignment #1

DUE: July 4th



You are to implement an object tracker. Object tracking is a CV problem that has been researched extensively for many years. While significant progress has been made (e.g., in tracking vehicles, pedestrians and bicyclists on the road), the general tracking problem is still very challenging (You can check out www.tracking-net.org for public-domain datasets and benchmarks).

The general tracking problem is formulated as follow: Given a video sequence and the bounding box (bbox) of the object of interest in the first n frames (often n is set to 1), track the object's locations in successive frames. For an ideal solution, not only you maintain accurate predictions of the bbox locations (regardless of self-occlusion and occlusion by others; change in scale, aspect ratio, perspective, and environmental lighting; deformation and articulation of the object, etc.), but you also generate such predictions in real-time.

Sample test images can be found under testimages/prog1 folder (or follow the local image archive link from the class web page). Sample frames from one such video are shown above. The red boxes are the ground-truth (GT) bboxes, the blue boxes are the predicted (PD) bboxes by "some" tracking algorithm. The numbers at the upper-left corner are IoU, or intersection-over-union, which is a widely-used metric in measuring tracking accuracy. IoU is defined as: (area of intersection of the GT and PD bboxes) divided by (area of union of the GT and PD bboxes). IoU ranges from 0 to 1, with larger numbers representing better correspondence.

It is highly recommended that you collect short videos yourself to test and debug your tracking program first before you try your hands on those sample videos provided. For your own test videos, you want to make sure that (1) the camera is stationary, (2) the scene is stationary except for the object of interest which is moving, (3) the trajectory of the object should be such that no occlusion and little self-occlusion occurs and the object's appearance does not change significantly in the video, (4) the object is rigid with no articulation and deformation, (5) the object has distinct surface textures to allow easy identification, so it should be evident what object you are tracking (e.g., a person, a vehicle, etc.). The simplified experimental scenario will serve as a reality check before you move onto more complicated tracking scenarios in testimages/prog1.

For grading, you should turn in your tracking program (Python, Matlab, C, C++, Java, etc.) and sample videos (roughly 10 to 20 seconds long). At least (2) sets of results are from your own videos and one (1) set from videos in testimages/prog1. The results must be an mp4 video. For your own videos, you need to put a bbox in every frame so the reviewer can easily judge the accuracy of your tracker qualitatively. For videos in testimages, you should display (a) GT bboxes in red, (b) PD bboxes in blue, and (c) IoU at upper-left corner – just like the results shown above for a more quantitative evaluation.