Computer Vision 2023
(CSE344/ CSE544/ ECE344/ ECE544)
Assignment-4 [BONUS Assignment]
Max Marks (UG/PG): 40 / 40 **Due Date**: May 5, 2023, 11:59 PM

Instructions

- Keep collaborations at high-level discussions. Copying/plagiarism will be dealt with strictly.
- Your submission should be a single zip file **Roll_Number_HW[n].zip**. Include only the **relevant files** arranged with proper names. A single **.pdf report** explaining your codes with relevant graphs, visualization and solution to theory questions.
- Remember to **turn in** after uploading on Google Classroom. No justifications would be taken regarding this after the deadline.
- Start the assignment early. Resolve all your doubts from TAs during their office hours **two days before the deadline.**
- Kindly **document** your code. Don't forget to include all the necessary plots in your report.
- All [**PG**] questions, if any, are **optional for UG** students but are **mandatory for PG** students. UG students will get BONUS marks for solving that question.
- All [**BONUS**] questions, if any, are optional for all the students. As the name suggests, BONUS marks will be awarded to all the students who solve these questions.
- Your submission **must include a single python (*.py*) file for each question**. You can submit *.ipynb* along with the *.py* files. Failing to follow the naming convention or not submitting the python files will incur a **penalty**.

---

1. (20 points) **Multi-Object Tracking with Detection**
   Multi-Object Tracking (MOT) is the task of predicting the positions of objects throughout a video using their spatial as well as temporal features. Tracking can either be done by object detection or without object detection. In this assignment you are required to perform tracking by detection using a YOLOv5 object detection model and the Simple Online Realtime Tracking (SORT) algorithm. See this tutorial for a detailed conceptual discussion.

   1. Download the required data and code containing the following:
      (a) a benchmark dataset containing a sequence of images from a video (KITTI_17_images.zip).

(b) A Python implementation (by Alex Bewley) of the SORT algorithm.

(c) Code for an evaluation kit to assess the performance of multiple object tracking algorithms.

2. Import the dataset and the SORT implementation code from part 1 into a copy of this Colab Notebook and fill in the code for section 6 in order to perform multi-object tracking only for the class of pedestrians (10 points) on the given video sequence. Your code must produce an output consisting of a video sequence (5 points) that resembles the example output video sequence (outputVideoFile.mp4, see Figure 1) and a text file (5 points) containing bounding boxes in the same format as the ground truth (gt.txt).

Figure 1: Example video output frame

Note: Please make your own copy of the Colab Notebook.
Note: A pre-trained YOLOv5 model will be loaded for you.

2. (20 points) **Analysis**

1. Look at the README file for the evaluation kit for MOT algorithms that is an official part of the MOT15 benchmark. The ground-truth bounding boxes are provided in a file called "gt.txt" that is part of the provided dataset.

(a) Using your output text file and the ground truth, use the evaluation kit to report the MOTA and MOTP accuracy metrics. Fill in your code in section 7. (10 points)

(b) Comment on the values of the metrics obtained. See this discussion of metrics and their significance. (10 points)