Assignment 3: Pedestrian Tracking
Deadline: 12/06/2024

The goal of the assignment is to develop an online Multi-Object Tracker based on the tracking-by-detection paradigm. The tracker will focus on pedestrian tracking.

(This image, from the Internet, is only to represent the idea, it is not the expected results)

You will need:
1. A pedestrian detector.
2. A track management system to handle new entries into the scene or to remove undetected pedestrians.
3. A method, based on *maximum bipartite matching*, to associate the detected bounding boxes with known identities considering relevant features (i.e., proximity, IoU, appearance...).

Here more details:
1. As pedestrian detector, you will use the DETR model - *End-to-End object detection with Transformers* (available at *https://github.com/facebookresearch/detr?tab=readme-ov-file*)
Unfortunately, DETR has been implemented in Torch. There are unofficial versions in tensorflow, but I suggest using the official ones.
You will need to:
- install torch and torch vision
- create the model by the following line (which will download for you the weights)
model = torch.hub.load('facebookresearch/detr:main', 'detr_resnet50', pretrained=True)

Use the code I will upload on the portale unipa to use the model.

2. You will have to decide on the proper strategies for starting new pedestrian tracks or for removing old tracks (people leaving the scene). Generally, this is done by counting how many frames the person is detected in the scene or how many frames the person is not detected respectively.

3. To solve the maximum bipartite matching, you need the method in scipy: *scipy.optimize.linear_sum_assignment*
Be aware that the above method solves a minimum problem!
To find the cost matrix, you are free to choose your own distance metrics and feature representation (*could you try recycling the outputs of the DETR transformer?*). If you want, you can also try using

a pre-trained network to obtain visual features or a Siamese Network to solve the re-identification problem (*the latter solution could take you time if you need to train the net from scratch, try it only if the rest of the assignment is complete!*)

Experiments:
You will use the MOT17 dataset (https://motchallenge.net/data/MOT17/) to validate and test your method.
You will have to report the *HOTA metrics* and the *CLEAR MOT metrics.* All of them can be computed with the *TrackEval framework* (code: https://github.com/JonathonLuiten/TrackEval).

You will need to introduce hyperparameters (like thresholds). Use the same hyperparameters for all videos (don't use different threshold values for each test video, this is wrong!). Use a subset of videos as a validation set to select the best hyperparameter values. Only when you are confident in the values you have selected, use the test set to calculate performance metrics.

You will present *tables* showing the average metrics on the validation set as the hyperparameter values vary.
Then, you will show in *another table* the results obtained on the test set.
You are free to split the dataset into test and validation sets as you like.

You will prepare a report (maximum 4-pages) to describe:
- the strategies you used,
- the materials (books, websites, papers) you have consulted to prepare the assignment,
- the experiments you have done (thus, tables of the results) and the discussion (why it works/fails)
- how you collaborated within the group.

Together with the report, you will have to submit:
- a zip file with your code implementation;
- the test video for which you achieve the best results;
- the test video for which you achieve the worst results.

Do not include the entire dataset **nor** the weights of the DETR model.

VERY IMPORTANT: *if you copy the solution from the web or present someone else's solution pretending to have implemented it yourself, the assignment will obviously not be considered valid. Since I would consider the incident particularly serious, I suggest not appearing for the exam if this is your intention. Please behave correctly and cite the websites you will consult to prepare the assignment.*