# Second Laboratory Work: Detection and Tracking of pedestrians in public spaces

## I. Introduction

Detecting pedestrian in public spaces has become an important subject to be addressed in the computer vision and image processing communities. One of the reasons is that pedestrians are among the most important objects to be detected in images, and a large number of applications could benefit from this. For instance, in the fields of mobile robotics to inform Human-Robot Interaction systems, or automotive providing input to Advanced Driver Assistance Systems. Another important application is the tracking of pedestrians, with the goal of collecting the performed trajectories. The trajectories constitute a rich source of information, since it is possible to obtain the statistics of the activities being performed. As such, we can obtain typical trajectories that occur in a given scenario. An interesting consequence is that, it is possible to detect atypical or suspicious behaviour, that is, trajectories that deviate from the typical trajectories.

However, obtaining such algorithm for trajectory collection poses several challenges. For instance, the high variability that characterizes the pedestrians; the appearance of a pedestrian on the image that is affected by the person's pose; clothing; the atmospheric conditions that influence the illumination changes; the background clutter and occlusion. All he above issues play a role in making pedestrian detection a challenging problem to be solved.

The goal of this work is to develop an algorithm capable to provide the pedestrian's trajectories using conventional handcrafted features that has plateaued in recent years in the image processing community.

Also, it is expected that the output of the algorithm will be enriched using the visual information as much as possible.

Figure 1 shows an example of pedestrian in a public area (Fig. 1 left), and the respective detections represented by bounding boxes (Fig. 1 right).

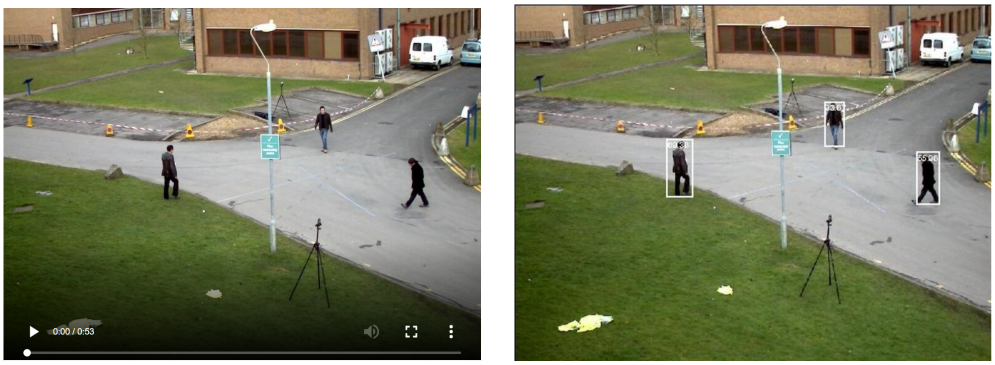

Figura 1. An image frame containing several pedestrians in a public space (left), and an image containing the ground truth of the using bounding boxes (right).

To start working please visit the following link:

[link]

Here you can find some information regarding the dataset. Important information comprises: $(i)$ samples $(ii)$ name of the sequence $(iii)$ frames per second (FPS) $(iv)$ image resolution $(v)$ video length $(vi)$ tracks $(vii)$ bounding boxes. From this link it is possible to download the images, as well as the labels of the ground truth. Also, there is an important paper that you should familiar with:

L. Leal-Taixe, A. Milan, I. Reid, S. Roth, and K. Schindler "MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking", arXiv 2015.

In this paper, it is described the context of the application (i.e. the pedestrian detection challenge). If you look at the Table 1 (page 4), you can see an overview of the training and testing sequences. Please note that in Section 3.4 (Data format) it is explained the format of the detections (see Table 2), that is, the ground truth.

Thus, as a first stage of the work, you are welcome to first plot the ground truth of the pedestrian's bounding boxes using the above information, throughout he sequence.

In the following, although I will suggest some lines of work, it is not mandatory to accomplish every single item to pass. Having said that, you are welcome to perform the following tasks:

1) Perform the tracking of pedestrians. A bounding box should be visible for each detection,
2) Plot the performed trajectories. To avoid a possible excess of the information visualisation, we can plot the trajectories dynamically,
3) If possible, try to assign a label for a given pedestrian, (or all if you want to),

4) You can also provide to the user the map of the trajectories performed in the video,
5) In the attempt to enrich the output of your algorithm, try to establish a heatmap where the color is assigned to the number of occurrences in a given position (region) of the image,
6) If possible, think about a way to provide optical flow of the pedestrian's motion.

## II. Evaluation metrics

Notice that, there is no perfect algorithm. This means that, no matter the approach is adopted, there is always some failures regarding the true location of the pedestrian. For instance, a merge or split in a given bounding box that can occur. Also, some misdetections may occur as well. Thus, one way to evaluate the algorithm is to use evaluation metrics. An evaluation strategy can be done as follows:

1) The first step is to built the ground truth as already mentioned (read paper [1]).
2) After this stage, the students are in conditions to show both the ground truth and the estimated bounding boxes provided by your algorithm.
3) Now, evaluation must be done. To accomplish this, the following metrics are suggested:
   a) Provide a precision-recall curve. The precision-recall (for a given frame in the sequence) are defined as follows:

$$precision = \frac{true\ positives}{true\ positives + false\ positives} \tag{1}$$

$$recall = \frac{true\ positives}{true\ positives + false\ negatives} \tag{2}$$

   This should be done integrated the above measures through the sequence, obtaining the corresponding curve.

   b) Provide the Intersection over union measure (IoU) or the *success* measure over the sequence. The IoU measure is defined as follows:

$$IoU = \frac{R_d \cap R_{gt}}{R_d \cup R_{gt}} \tag{3}$$

   where $R_d$ is the detected region estimated by the algorithm and $R_{gt}$ is the ground truth (manual labeled) region.

   Compute the IoU or the *success* plot. The IoU (or the success) plot shows the percentage of frames whose bounding box overlap ratio is higher than a given threshold. For threshold

values ranging from 0 to 1, with step of 0.05. The value of 1, means a perfect match, the value of 0 meas that the target is lost.

## III. READING MATERIAL

The students are welcome to read the following paper:

[1] L. Leal-Taixe, A. Milan, I. Reid, S. Roth, and K. Schindler "MOTChallenge 2015: Towards a Benchmark for Multi-Target Tracking", arXiv 2015.

[2] J. C. Nascimento and J. S. Marques. "Performance evaluation of object detection algorithms for video surveillance". IEEE Trans. on Multimedia, vol. 8, no. 4, pp. 761-774, Aug. 2006.

[3] Luis M. Fuentes and Sergio A. Velastin, "People tracking in surveillance applications", Proceedings 2nd IEEE Int. Workshop on PETS, Kauai, Hawaii, USA, Dec. 9 2001.