Fatma Moalla
fatma.moalla@student.ecp.fr

VIC 2020
March 10, 2020

Introduction to Visual Computing: assignment 2

1 Algorithm motivation and description

The objective of this assignment is to detect pedestrians from video-surveillance. The input contains 685 frames extracted from the video and the output is a list of bounding boxes for each image. We are also provided with the ground truth bounding boxes with the following format = `[frame,bb_number,x,y,dx,dy]`. `For this assignment, we will use two different detectors and also use non−maximum suppression for both. For this assignment, we treated frames separately and used the package opencv for the detection and imutils library for non−maximum−supression task`

1.1 First detector:

The first one, is a pretrained Histogram of Gradients descriptor with a Linear SVM classification model applied to opencv defaultPeople detector. We also added non-maximum suppression with an overlap threshold to this detector in order to filter out the extra-bounding boxes. We used this *biased* detector as a first approach to compare our results with the existing pedestrian detetcors. The function `pedestrian_detect1` enables to calculate the bounding boxes for a given image and also can activate or not non-maximum suppression through `nms`.

1.2 Second detector:

The second detector, is also a pretrained detector called *CascadeClassifier*[1]. The `haarcascade_fullbody.xml` represents the pretrained classifier on several fullbody images that will be used to detect pedestrians. This detector uses a Haar feature-based cascade classifiers.In fact, Haar features can represent edges, lines, corners and using the casscade classifiers they can be grouped into different stages of classifiers and applied one at a time. Similarly, we use the function `pedestrian_detect2` for this detector.

2 Results

	default frame n° 300	max IoU	average IoU	median IoU
Detector 1 (HoG + SVM)	**43.3**	**68.6%**	11.4%	4.0%
Detector 2 (Haar + cascade)	23.9%	51.7%	10.6 %	**5.5%**

As shown in the table above, the best detector is **the first one** as it uses a default pedestrian detector and that's why it has outstanding results. The second detector presents a maximum score that exceeds 50%. The mean and average score for both detectors are low because we replaced all 'nan' values by zeros resulting and due to the fact that all frames does not contains pedestrian. Also, the comparison here, did include non-maximum supression as the results are improved.

3 Discussion

In this part, we will suggest that in the next steps, we should take advantage of the succession of the image and use the spatio-temporal information and take advantage of optical flow techniques to detect the motion of the pedestrian and therefore predict the exact position of bounding boxes in the previous and following frames.

[1] 'https://docs.opencv.org/3.4/db/d28/tutorial$_c ascade_c lassifier.html'