

Supplementary Material: Unsupervised Hard Example Mining from Videos for Improved Object Detection

SouYoung Jin*, Aruni RoyChowdhury*, Huaizu Jiang, Ashish Singh, Aditya Prasad, Deep Chakraborty, and Erik Learned-Miller

College of Information and Computer Sciences, University of Massachusetts, Amherst
{souyoungjin, arunirc, hzjiang, ashishsingh, aprasad, dchakraborty, elm}@cs.umass.edu

In the attached video, `suppVideo.mp4`, we have visualized how a large number of hard negatives can be obtained *automatically* by analyzing the output of a trained detector on video sequences. **Red boxes** denote the detections that are *isolated in time* due to “detector flicker” (*i.e. hard negatives*), among the high confidence detections (**green boxes**). Additionally, some of the hard negatives from the video are visualized in this document in Fig. 1-8 for faces and Fig. 9-11 for pedestrians. We also demonstrate the presence of “hard positives” (shown near the end of the video from 2:01–2:20), which are positive examples that are surrounded in time by correct detections.

* Authors contributed equally

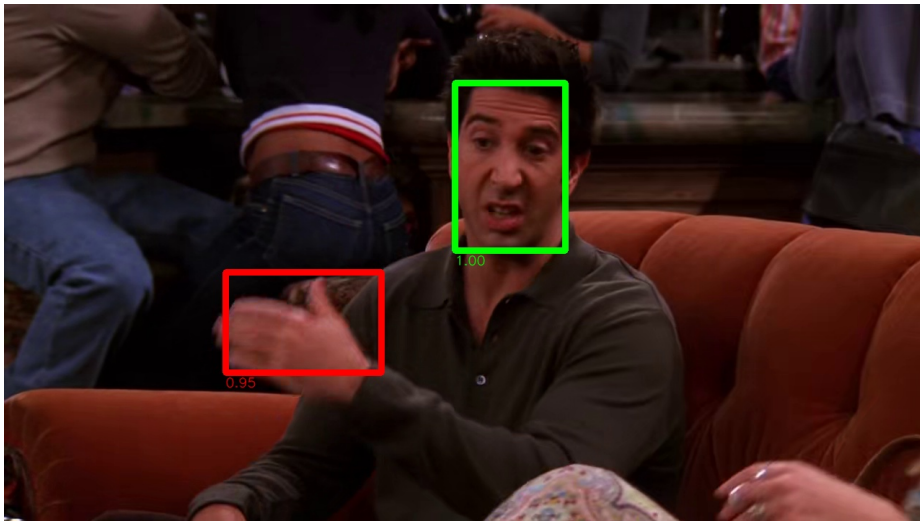


Fig. 1

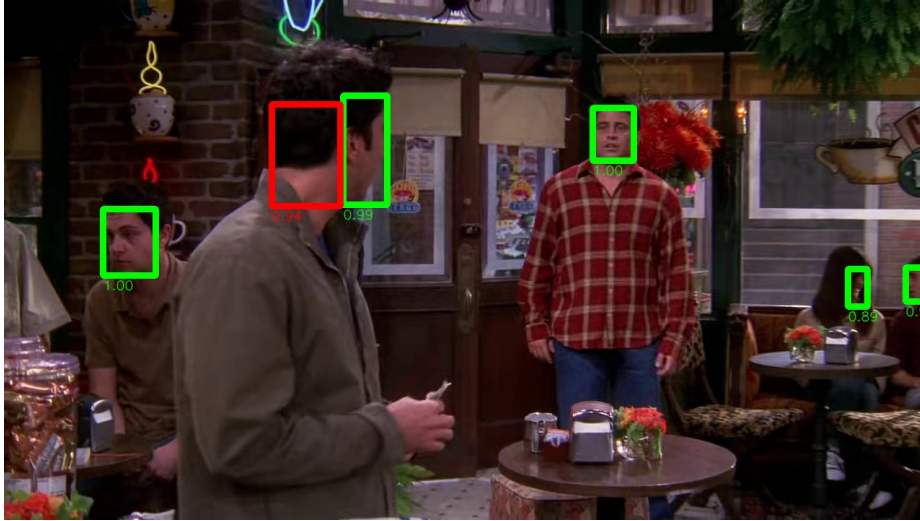


Fig. 2

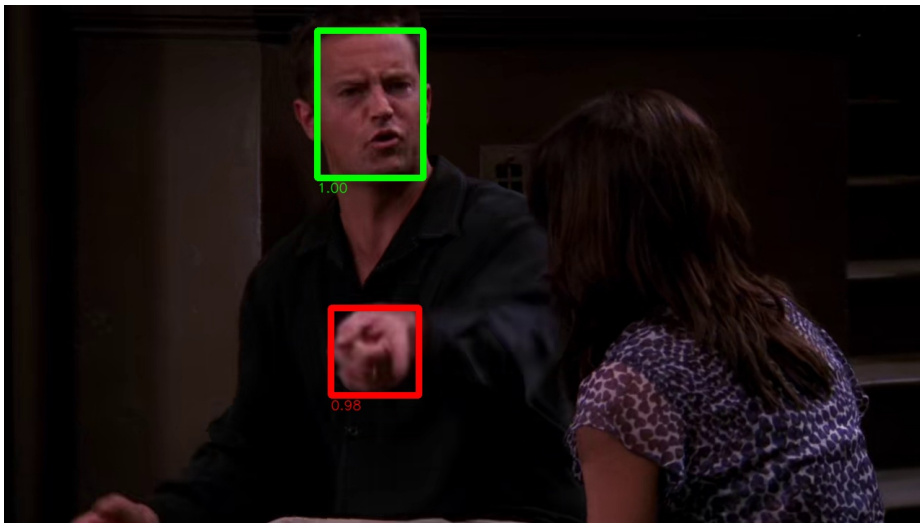


Fig. 3

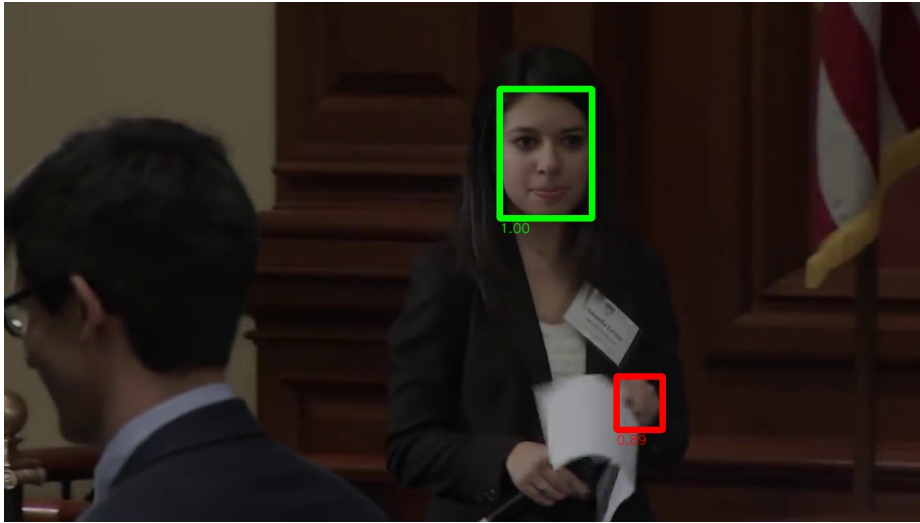


Fig. 4



Fig. 5

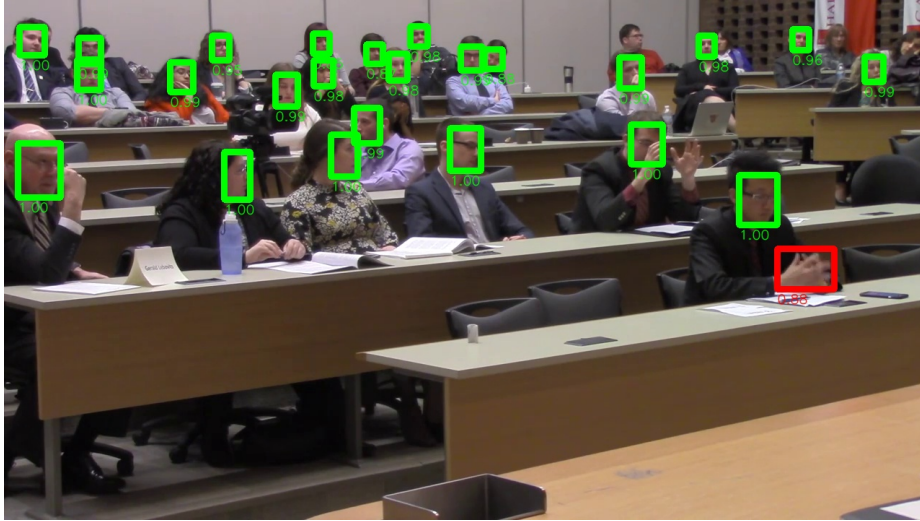


Fig. 6



Fig. 7



Fig. 8



Fig. 9



Fig. 10

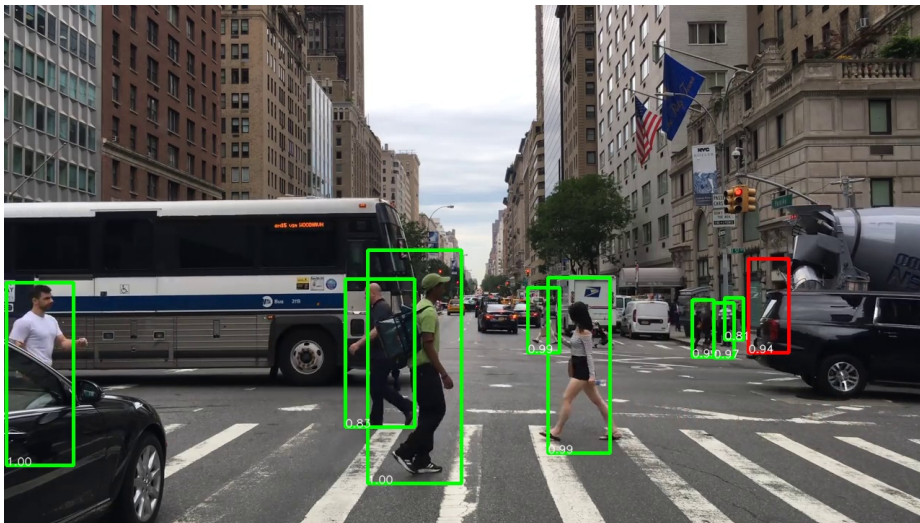


Fig. 11