

The threshold for detection dramatically affects the performance here. As shown above, a bird is incorrectly detected as human when the threshold is set too low. On the other hand, YOLO couldn't handle occlusions well. We could switch to density map in crowd settings. In regards to that, traditional method based on HOG and SVM Classifier could be utilized here. The density map method normalized the contrast in pixels to get better accuracy; HOG and SVM will separate human from the environment using the constraint margin, which is invariant to geometric and photometric transformations.