I've tested *faster rcnn inception v2 coco*and generated a demo video with 30x speed. I also tested Haar and HOG as well.

I chose *faster rcnn inception v2 coco*because it balances accuracy and speed well.

Test Bench

Intel Core i5-5200U (2.20Gz), 4GB memory, without GPU turbo

I've edited the test video and selected six clips to test:

1.Normal working scene 1

2.Normal working scene 2

3.People gather together and overlap each other

4.Multiple people moving in the same scene

5.Individuals separate clearly

6.A scene without human

With these clips I want to know the performance of neural network-based models and their frequency of occurrence of missed detection, false detection and duplicate detection, unreliable detection boundary, and flickers in detection.

After seeing the demo video of *faster rcnn inception v2 coco*I find:

Missed detection and flickers in detection still occurs sometimes especially when people are **bend over** or some **part of body is outside of the camera**. Missed detection and flickers in detection occurs usually in old approaches.

Detection boundary is **quite** **accurate** and has significant progress compared to old approaches.

False detection **never occurs** while it occurs both in old approaches (especially in HOG).

Duplicate detection **only occurs in few frames** when people overlapping each other. I find that this model tends to think **two or more people overlapping as one** rather than make a duplicate detection.

There are some features of this model while detecting as well:

It can also detect people climbing a ladder (sometimes during the whole procedure) which can't be detected in old approaches.

People will be miss detected if the colour of their clothes are dark and can merge into other huge dark colour thing, like a truck. Whether similar bright colour can mislead the model is not confirmed yet.

Bending over can sometimes be detected if his surrounding is clear. Knowing the condition of detecting a bend over movement needs more related tests.

Detection on incomplete human figure (part of body is outside of the camera) is a lot better than old approaches.