From my intuition, the most important criteria to select the model is to run in real time, since Deep Learning models tend to be resource intensive.

The field of action recognition emerged fairly recently. Among action recognition models, Openpose was the first open-source real time model published in 2018. It uses a bottom-up approach with action recognition, which is to form a skeleton by detecting key points. Alternatively, there is also the top-down approach, which is to detect the human first in a bounding box. The advantage of using the first approach is that it is faster, as proved by a shorter runtime in comparison with Mask R-CNN and Alpha-Pose. Besides the advantage, since it is a “classic” model as many people cited the paper, it is already been proven and tested, therefore in this an hour project it is optimal to choose this model.

From my testing with limited CPU ( 2.6 GHz 6-core Intel Core i7) , I was able to test with one frame. The key points are clear and accurate. However, one drawback of this model is that videos are broken into image frames.

**男人戴着耳机

低可信度描述已自动生成**