Virtual Training Assistant

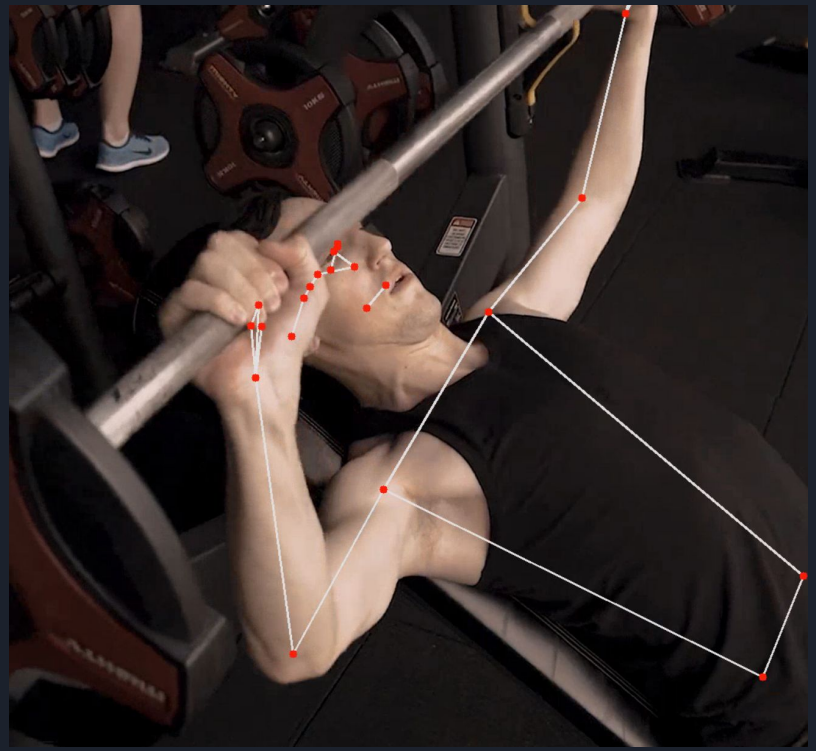
Computer Vision Course Project

Overview

```
You did 4 reps with an average rep duration of 2.20 seconds

FEEDBACK:
- On average, your reps were 78.52% slower then the reference exercise video.
```

Overview: Usage of Pose estimation



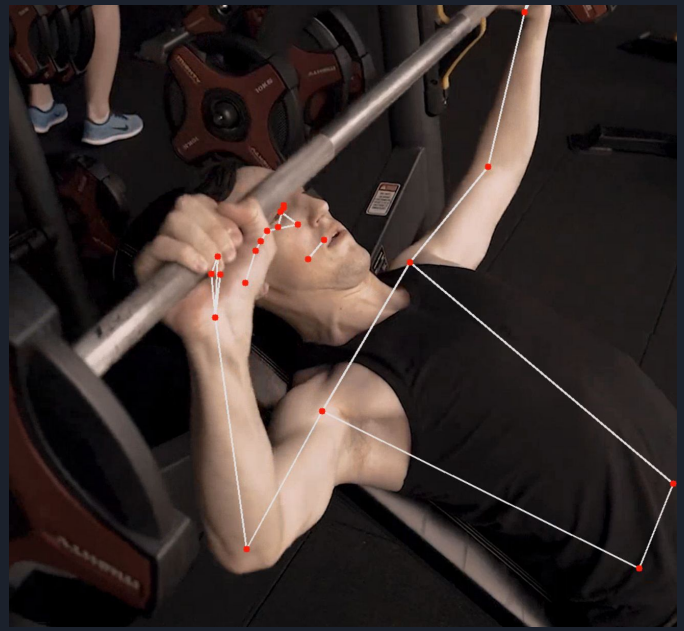
Our pipeline

01 Human Pose estimation using either BlazePose or YOLOv7

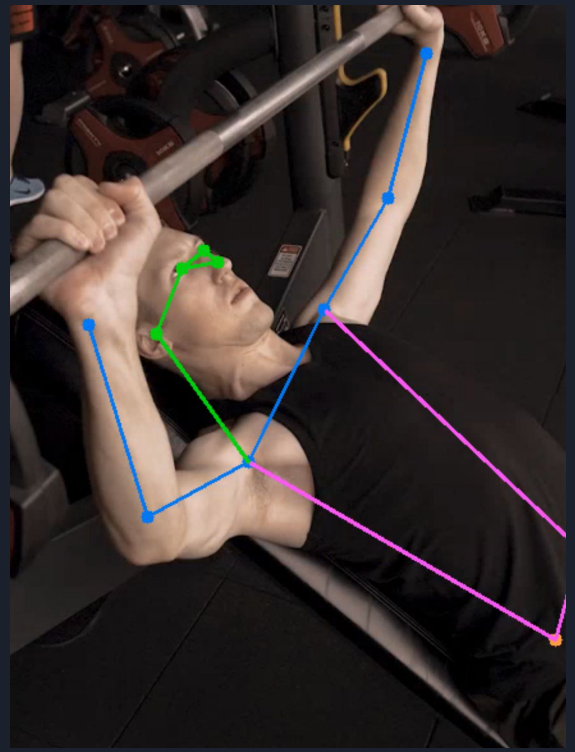
02 Repetition counting using timeseries analysis of the evolution of joint angles

03 Exercise feedback by comparing the repetition speed to the one from a reference video.

BlazePose vs YOLOv7 for Human Pose estimation



BlazePose

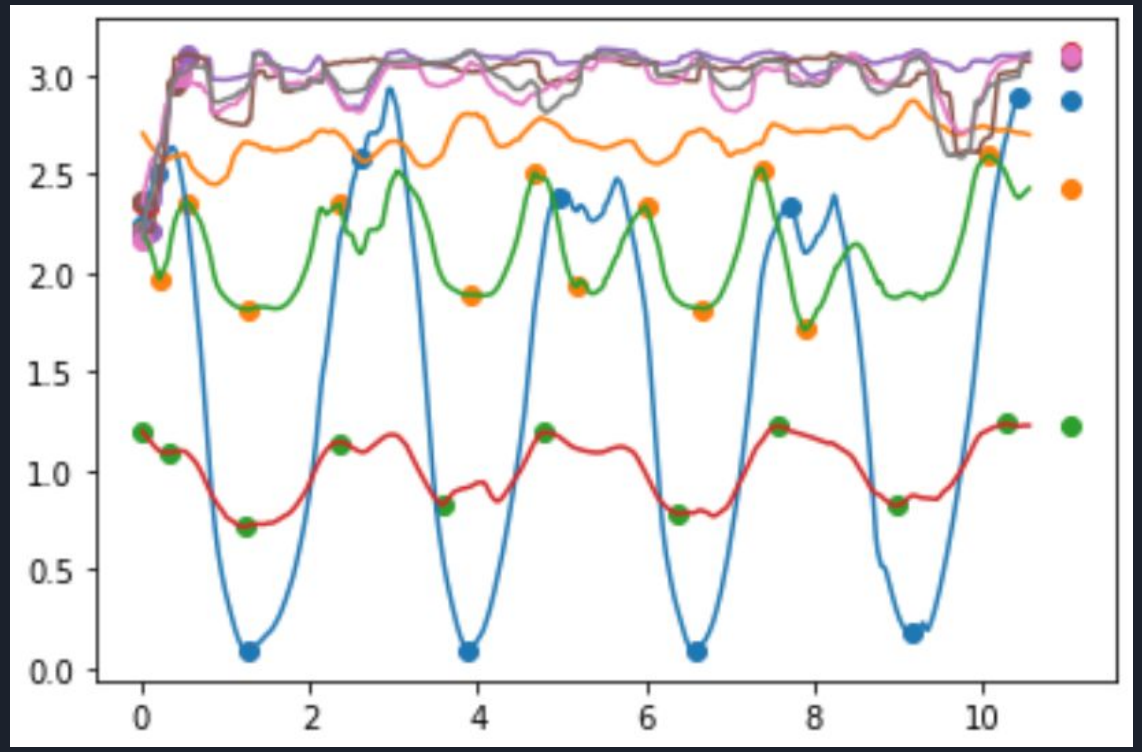


YOLOv7

Calculating reps from the Pose estimation

- Calculate 8 joint angles: Shoulders, elbows, hips and knees.
- Make a timeseries of each angle
- Apply a moving average filter
- Find the local extrema
- Count reps using the extrema: A rep is one we cross a given range threshold in both direction (back and forth).

Calculating reps from the Pose estimation



Exercise evaluation

We run a reference video through the same pipeline, then we compare the average rep duration across the videos.

Our solution is exercise agnostic: It can count reps for any exercise that is a repetitive movement pattern, and it can feedback for it given a reference video.

Experimented with exercise classification models

The classifier input is constructed from the keypoints of the pose estimation. We have experimented with 2 classifiers. Both gave very bad results because we did not have enough data:

- Single frame classifier: Uses a single frame of the video to predict the exercise. The inference is done by taking the majority prediction over all frames.
- Frame sequence classifier: LSTM network for the keypoints from every frame. It was overfitting to the training data and could not generalize to other videos.

Key observations

- Cannot perform in real-time
- Limited to the human body pose (does not contain the equipment)
- Noisy 3rd dimension in the pose estimation
- Manual tuning is still needed for the repetition counter

Thank you for your attention!