Workout analysis using real-time pose estimation

Sai Achyuth Vellineni

Abstract

The idea behind the project is to utilize pose estimation to effectively analyse workouts and provide visual cues to correct the posture during workouts to prevent long term injuries, muscular imbalances and also engage the muscles effectively for a proper workout. In this project I have explored the mediapipe platform from google and utilised its pose estimation framework to track a single workout and derive various insights.I have also explored various applications of pose estimation which I have summarized in the applications and future work section.

1.Introduction:

Human body pose estimation from real-time videos or images plays a central role in various applications such as health tracking, body language detection, and gestural control.It is a very challenging task and we have to deal with a number of variables such as the large number of poses ,different body sizes, numerous degrees of freedom, and part occlusions.

Pose estimation was originally done using pictorial structures which model the human body as a collection of rigid templates and a set of pairwise potentials which form a tree structure allowing for efficient inference at test time. Recently , with the advent of neural networks , Convolutional Neural Networks have shown remarkable and robust performance and high part localization accuracy which outperform prior methods by a large margin. A key feature of these approaches is that they integrate non-linear hierarchical feature extraction with regression and make use of the large datasets that are readily available. For human pose estimation , the CNN features are regressed in order to provide joint prediction of the body parts.

2.MediaPipe Framework

Mediapipe is a user-friendly and easy to use framework which provides a number of functionalities such as pose estimation , iris recognition , hand gesture recognition etc. The pose estimation functionality utilizes a two step detector-tracker pipeline which was proven to be effective in a number of other applications. The detector first locates the person region of interest (ROI) within the frame. The BlazeFace model is used as a proxy for a person detector and hence this model comes with an assumption that the person's face is visible in the frame which is the case for most of the mobile devices. The tracker then uses this pose alignment to subsequently predict the pose landmarks and keypoint coordinates, and redefines the region of interest for the current frame. When the person moves out of frame or the face is not visible then the detector network is re-run to estimate the pose alignment.

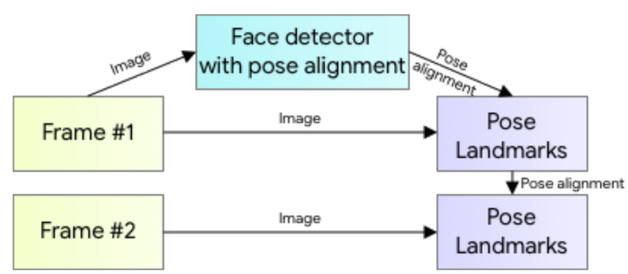


Fig1

The detector network finds two virtual keypoints based on Da Vinci's Virtuvian Man - the midpoint of the torso , radius of the circle inscribing the person and the incline angle of the line connecting the shoulder and shoulder and hip midpoints. This results in accurate tracking for a wide range of motions for all 3 degrees of freedom. The model predicts 33 keypoints and their coordinates for 3 degrees of freedom from the ROI.

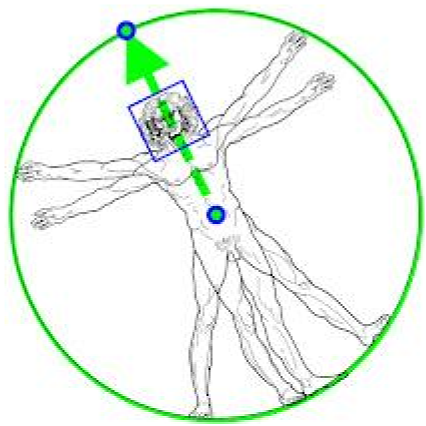


Fig. 2

3.BlazePose - Model and Architecture

The common approach for pose estimation involves producing heatmaps for all joints along with refining offsets over each coordinate.While this is a good and scalable technique , it is too heavy for single person pose estimation or to run on a mobile device. In contrast , regression based techniques are less computationally demanding and are also equally scalable. BlazePose uses a stacks hourglass architecture which gives a significant boost to the quality of prediction even with a small number of parameters. A encoder-decoder network architecture is used to predict the heatmap of all the joints , which is followed by another regressor which regresses directly on the coordinates of all the joints. The heatmap branch is used only in the training stage and it is discarded during the inference phase which results in a lightweight embedding and increase in performance. This is what makes the network suitable for mobile devices and CPUS. Skip connections are actively utilised between different layers of the network to achieve a balance between the high level and the low level features. Also the gradients from the enoder network are not propagated back to the heatmap trained features which increases the performance as well as regression accuracy to a large extent.

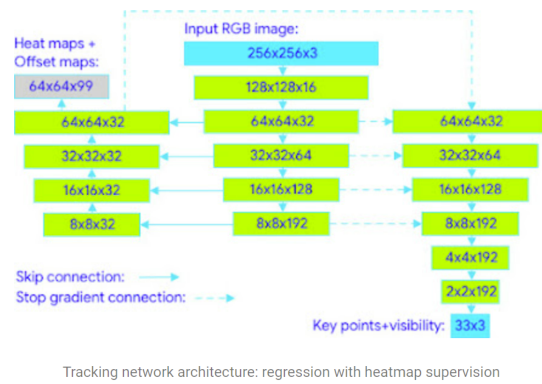


Fig 3

4.Experiments and Results

In this project I have used pose estimation to effectively track a workout as well as provide visual cues regarding the posture in real time. For this experiment I have taken bicep curls as a use case , where I track the workout using my laptop camera and provide visual cues on whether the posture is good or bad and also by how much one has to correct the posture. At the same time I also provide a rep count on the screen , which counts the number of repetitions of bicep curls being performed. For the pose correction , I calculated the angles between the wrist-shoulder line and the shoulder-waistline and set the threshold to a maximum of 30 degrees for good posture.

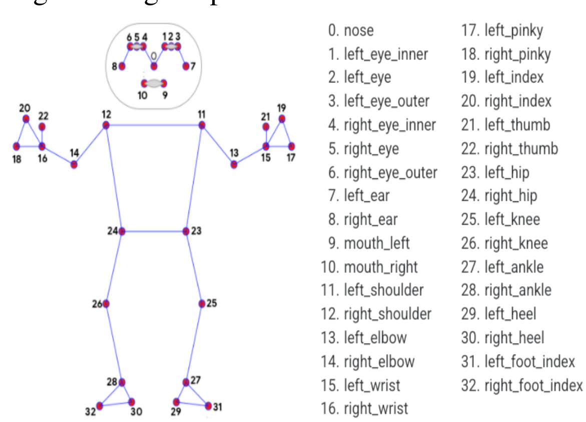


Fig 4. Keypoints of all the joints.

This comes from the fact that elbows have to be close to the waist while doing bicep curls to effectively engage the biceps. If we lift the elbows during the workout then we tend to involve the forearm as well.

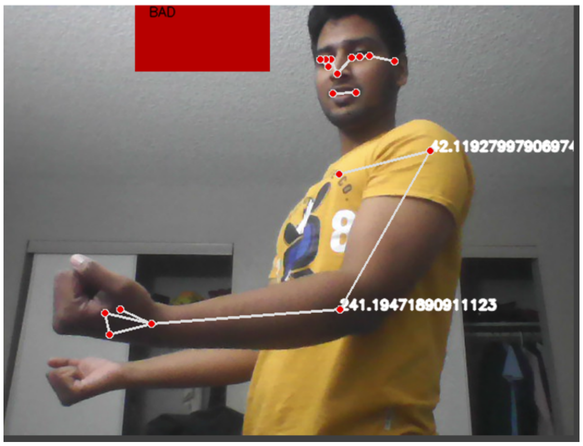


Fig.5 Bad posture where the angle between the shoulder and the hip is > 30 degrees,

5.Other Applications

Pose estimation has been used for a lot of other applications - mainly to prevent long term injuries in sports and athletics. Sports persons and athletes adapt their bodies to the movements they regularly perform which often result in muscle imbalances. Although this may not immediately result in injuries , it might hamper their posture or spinal positioning which can ultimately lead to issues in walking, sitting and even lying down, as time progresses. Since most of these motions are repetitive , we can investigate the pattern of muscle imbalances in these athletes and also evaluate evaluate those patterns based on the posture, balance, gait and movement variations and effectively reduce the chances of such injuries in future,

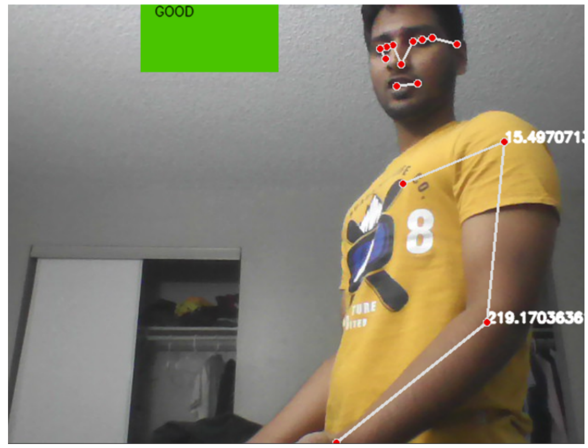


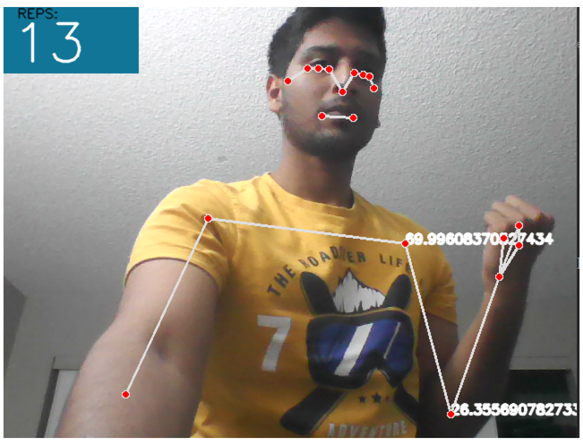
Fig.6 Good posture where the angle between the shoulder and the

6.Future Work

In the future I would like to develop an application , an online web service where people can upload their workouts and check their posture , as well as other key insights such as muscle imbalances between pairs of connected muscles , percentage of efficiency etc. In other words , a virtual trainer application which can also be used by gym trainers to keep track of progress and quality of workouts. Apart from this , pose estimation can also be used to monitor postures in work environments - work ergonomics , which can be used in conjunction with body language detection as well as face and emotion detectors , to effectively monitor the mood and stress levels of the employees and provide a better workplace.

7.Conclusion

Pose estimation techniques have become very efficient in recent years , so much that we can run a model in mobile phones and laptops and don't require any GPU or special hardware. Hence this can be effectively used to develop mobile applications as well as web services and apply it to solve problems in various industries from sports , fitness , health tracking , security , safety etc.



8.References:

[1] BlazePose: On-device Real-time Body Pose tracking:Valentin Bazarevsky, Ivan Grishchenko, Karthik Raveendran, Tyler Zhu, Fan Zhang, Matthias Grundmann

[2] Computer aid assessment of muscular imbalance for preventing overuse injuries in athletes:Maheshya Weerasinghe,Kapila Asanga Diaz,Anuja Dharmaratne,N.D. Kodikara

[3]Human pose estimation via Convolutional Part Heatmap Regression:Adrian Bulat and Georgios Tzimiropoulos

[4]BlazeFace: Sub-millisecond Neural Face Detection on Mobile GPUs:Valentin Bazarevsky, Yury Kartynnik, Andrey Vakunov, Karthik Raveendran, Matthias Grundmann

[5]Pramerdorfer, M. Kampel, J. Heering: "3D Upper-Body Pose Estimation and Classification for Detecting Unhealthy Sitting Postures at the Workplace