IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 7.301

Volume 3, Issue 1, February 2023

BICEP Curl Tracker Using Mediapipe

Ms. S. Harishma[1] and Dr. R. K. Kavitha[2]
Department of Computer Applications[1,2]
Kumaraguru College of Technology, Coimbatore, India

Abstract: *There isn't "real" live tracking in the fitness system that is currently in place. People who exercise at the gym manually track their progress, which may occasionally be rather tedious. Recent advances in artificial intelligence's image processing techniques are used in this research. Traditionally, humans are simply seen as a bounding box in object detection (a square). Through a process called stance identification and pose tracking, computers can learn to interpret human body language just like humans can. The conventional techniques for tracking poses, on the other hand, are neither rapid enough nor resistant enough to occlusions to be used in a practical setting. The development of high-performance, real-time pose identification and tracking will be a driving force behind some of the most significant advancements in computer vision. By monitoring a person's location in real-time, for instance, computers will have the ability to develop a knowledge of human behavior that is both more fine-grained and more logical.*

Keywords: Object Detection, Computer Vision, Mediapipe, CNN

I. INTRODUCTION

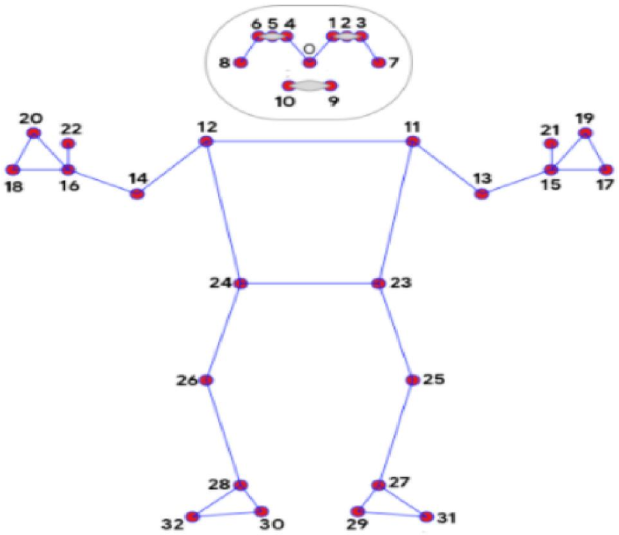
The "human pose estimation" computer vision job determines where several body landmarks on a person can be found within a given image or video by making educated guesses. By reviewing recordings that were taken while the patient was outside of a clinical setting, this technology may be able to perform virtual motion assessments. Human position estimate from video is crucial for a wide variety of applications, including the monitoring of physical activity, the comprehension of sign language, and the management of full-body motions. It is possible to use this as the basis for other forms of exercise, such as yoga, dancing, and fitness. Additionally, it can make it possible for augmented reality to layer digital information and material on top of the real area where it is being viewed.

II. REVIEW OF LITERATURE

There are numerous other studies based on this, such as Jatin [3] use of OpenPose to identify posture. For it, V Gupta [2] employed a deep learning model. Another well-known study by Chen [5] utilised MediaPipe on a gadget for in-the-moment hand monitoring. Robust articulated-ICP was utilised in A. Tagliasacchi's study [3] for real-time hand tracking. A.Toshev [4] completed a Deep Stance study in 2014 that utilised Deep Neural Networks to estimate human pose. In addition, a COCO Key Point Detection Task [6] was carried out in the year 2020. The major goal of the Pose Trainer project, which was carried out by S. Chen [5] using Stanford University, was to correct exercise posture. Another piece of work [1] has successfully finished a project for neural facial recognition on mobile devices. For the purpose of identifying human poses, composite fields were constructed. T Y Lin [7] published a study on common-based items in Springer in 2014. A work was [4] published which used a stacked glass network-based posture estimation study in Springer in 2016. On computer vision and pattern recognition, a researcher wrote and published a study on 3D posture estimates in single-depth pictures using single-view CNN to observe multiple CNN. The research mentioned above contains both positive and negative aspects [6]. The above-mentioned researches were expensive, and a team of workers was needed to carry them out. It also needed a lot of additional outside resources and maintaining them cost a lot of money [8]. Our study makes advantage of a device's webcam to record the numerous body coordinates needed to determine the angle and then provide the final count based on those values. It doesn't need any maintenance and doesn't require any more materials that the user would have to pay for. It is environmentally friendly and open to use by everybody.

III. PROPOSED WORK

The module or dependency utilized in this work is Mediapipe. As shown in figure 1, MediaPipe Posture is a machine learning (ML) method for high-fidelity body position tracking. It does this by inferring 33 3D landmarks and a background segmentation mask on the complete body from RGB video frames. The solution makes use of a detector-tracker machine learning pipeline that has two steps. The first step in the pipeline involves making use of a detector to locate the person-and-pose region of interest (ROI) in the frame. The ROI-cropped frame is subsequently utilized by the tracker as input for the purpose of forecasting the pose landmarks and segmentation mask included within the ROI. Note that the detector is only used in video use cases when it is absolutely required to do so, such as for the initial frame or when the tracker was unable to detect the presence of a body position in the frame before it. This is something that should be kept in mind. Other frames just have the ROI derived from stance markers from the preceding frame, as seen in fig. 2. The pipeline is constructed as a MediaPipe graph that renders using a particular posture renderer subgraph and uses a pose landmark subgraph from the pose landmark module. Both of these subgraphs come from the pose landmark module. The pose landmark subgraph makes use of an internal pose detection subgraph that is a part of the module that handles posture detection.



0. nose
1. left_eye_inner
2. left_eye
3. left_eye_outer
4. right_eye_inner
5. right_eye
6. right_eye_outer
7. left_ear
8. right_ear
9. mouth_left
10. mouth_right
11. left_shoulder
12. right_shoulder
13. left_elbow
14. right_elbow
15. left_wrist
16. right_wrist
17. left_pinky
18. right_pinky
19. left_index
20. right_index
21. left_thumb
22. right_thumb
23. left_hip
24. right_hip
25. left_knee
26. right_knee
27. left_ankle
28. right_ankle
29. left_heel
30. right_heel
31. left_foot_index
32. right_foot_index

Figure 1: Landmarks

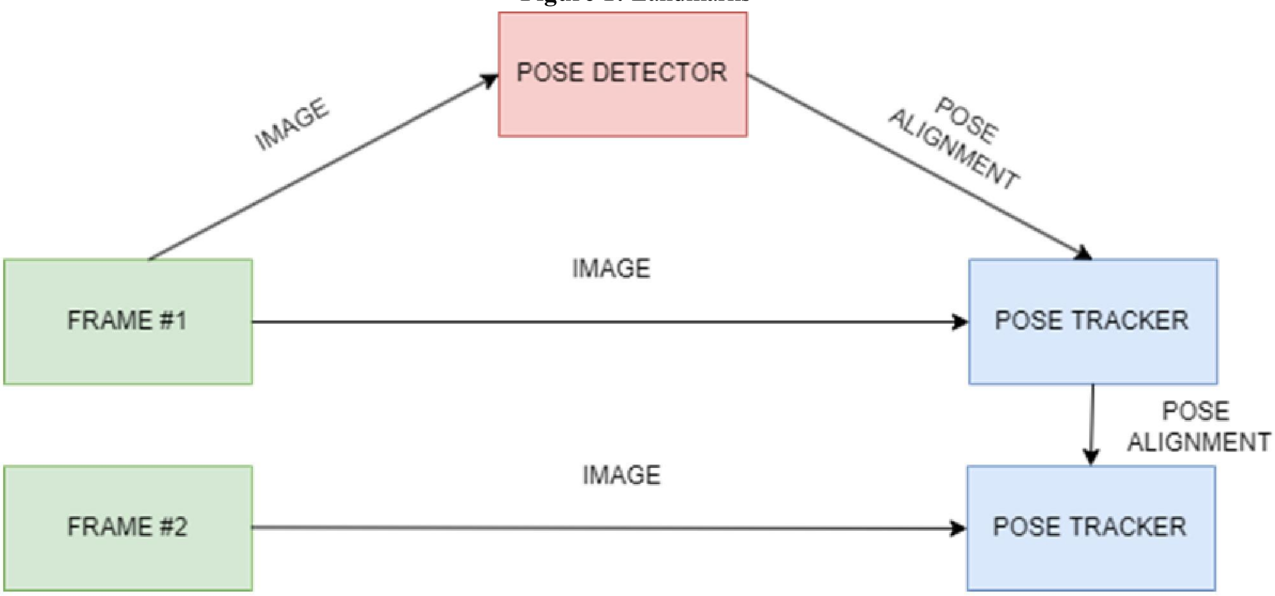


Figure 2: Pose Tracker

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 7.301

Volume 3, Issue 1, February 2023

3.1 The Internal Working of the Algorithm

A. Classification

To locate the posture samples that are closest to a target one, the k-nearest neighbor (k-NN) approach for classifying poses requires a feature vector representation of each sample as well as a metric to determine the distance between any two such feature vectors. In order to translate posture landmarks into a feature vector, we make use of pairwise distances, as shown in figure 3, between specified lists of pose joints. Some examples of these distances are the distance between the wrist and shoulder, the ankle and hip, and two wrists. Before the conversion takes place, each pose is normalized so that it has the same torso size and vertical torso orientation. This is done since the algorithm relies on distances.

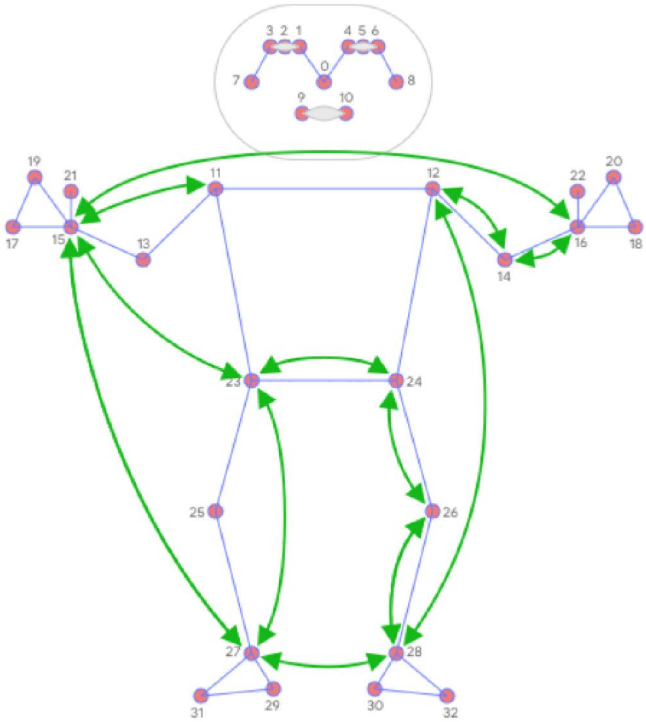


Figure 3: Main pairwise distance used for pose feature

Virtual motion assessments may benefit from using AI models that value body key points to characterize body position. Convolution Neural Networks (CNN) are commonly used in human pose estimation to make predictions about a person's position based on input photos or videos. Owing to the enormous amount of conceivable human poses, the enormous number of degrees of freedom, variations in appearance such as garments and illumination, changes in the environment, and occlusions make it a difficult task to determine the proper pixel coordinates of critical places on the body. In spite of these difficulties, a lot of reliable models have been developed that function quite well in areas like sign language, rehabilitation, and athletic training.

The BlazePose model, which was developed by Google, incorporates a detector-tracker inference pipeline with two steps. The tracker is used to follow the person in consecutive frames once the detector has finished running on the first frame or until a person is found. The encoder-decoder network architecture used by Bazarevsky et al. in their model "employs an encoder that regresses straight to the coordinates of all joints, followed by another encoder to forecast heatmaps for all joints." Because of its lightweight design and ability to do real-time inference, BlazePose is well-suited for use in mobile apps.

B. Repetition Counting

As seen in fig. 4 and fig. 5, the algorithm keeps track of the likelihood of a target pose class to count the repetitions.

- The algorithm indicates that the "down" posture class has been entered once the likelihood of the "down" pose class crosses a specific level during the initial time.

IJARSCT

ISSN (Online) 2581-9429

International Journal of Advanced Research in Science, Communication and Technology (IJARSCT)

Impact Factor: 7.301

Volume 3, Issue 1, February 2023

- The process indicates that the "down" posture class has been put off and raises the counter as the probability falls below the cutoff.

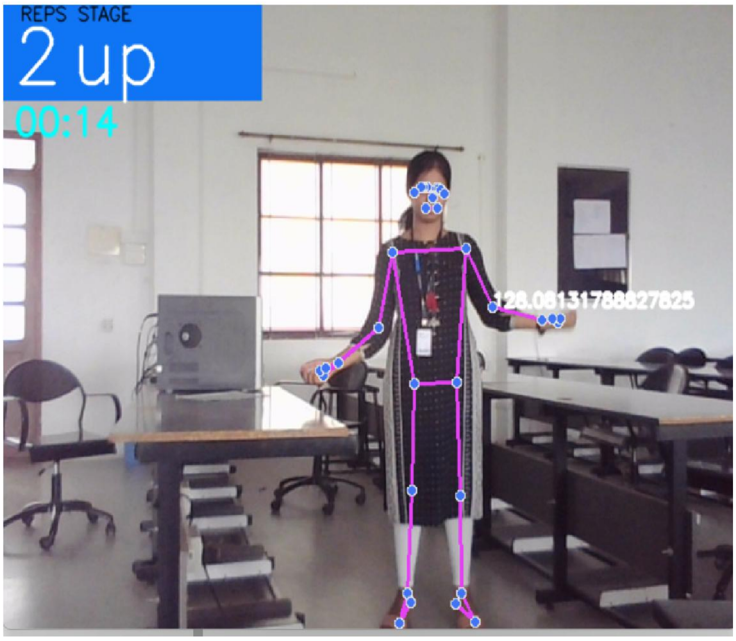


Figure 4: Curl counter -Up

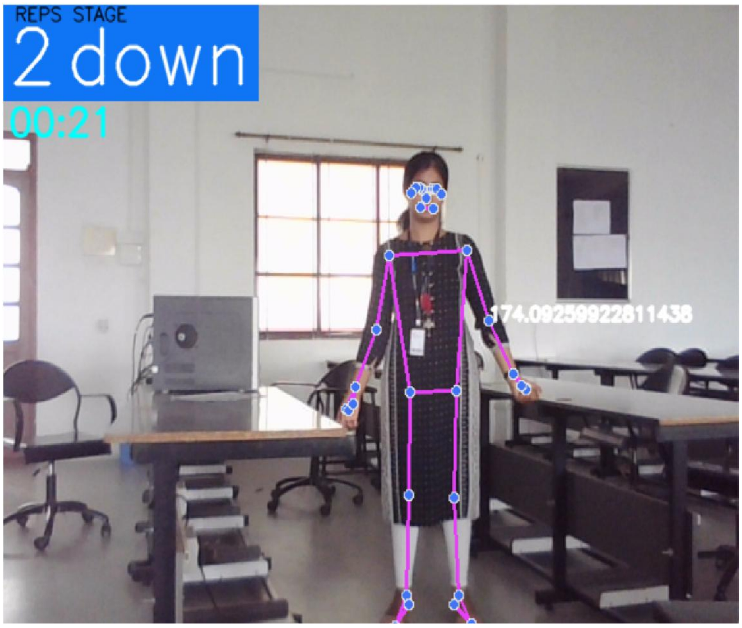


Figure 5: Curl counter -Down

IV. CONCLUSION

People today struggle to maintain a healthy lifestyle due to their hectic schedules. In addition to being healthy, being fit helps prevent diseases. Using artificial intelligence to motivate people to start exercising and maintain a healthy weight is a novel approach. This project made the most of MediaPipe by using it. Although this effort is still in its early stages, it already has a very user-friendly and practical system. There is always room for development and expansion. Coding conventions are adhered to during the creation of this for simple expansion and maintenance. Future improvements to the work could include other exercises like squats and jump rope. This application can be improved to meet additional system needs even though it was designed for the bare minimum.

REFERENCES

[1]. Deep Learning-based Human Pose Estimation using OpenCV By V Gupta.

[2]. Jatin Goel, Harshita Jain, Prabhjot Kaur, " Bicep Curl Count: Computer Vision Based Counting", Ijraset Journal For Research in Applied Science and Engineering Technology, 2022.

[3]. Robust articulated-icp for real-time hand tracking was done by A.Tagliasacchi, M.Schroder, A.Tkach, S.Bouaziz, M.Botsch, and M.Pauly in Computer Graphics Forum.

[4]. DeepPose: Human Pose Estimation via Deep Neural Networks (August 2014) A.Toshev, C.Szegedy.

[5]. Pose Trainer: Correcting Exercise Form using Pose Estimation performed by S.Chen, R.R. Yang Department of CS., Stanford University.

[6]. Composite fields for human pose estimation by S Kreiss, L Bertoni, and AAlah in 2019. [7] Common objects in context by T Y Lin, M Maire, S Belongie, J Hays, P Perona, D Ramanan, P Dollar, and C Lawrence Â´ Zitnick in 2014.

[7]. Robust 3d hand pose estimation in single depth images: from single-view CNN to multi-view CNNs by L.Ge, H.Liang, J.Yuan, and D.Thalmann in 2016.