IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁜ Impact Factor 7.39 ⁜ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11792

Posture Correction using Human Pose Estimation

Shiva Shashank Dhavala[1], Vaibhav Porwal[2]

[1]Student, Department of ISE, RV College of Engineering, Bengaluru, India

[2]Student, Department of ISE, RV College of Engineering, Bengaluru, India

Abstract: Health and fitness have become a priority to a vast majority of the population ever since the COVID-19 pandemic has struck. People have resolved to eat healthy and exercise to ensure good health. The increase in the interest to exercise amidst public movement restriction led to the boom of the remote fitness industry. Various applications such Home Workouts, Nike Fit and Cult.fit prescribe a set of exercises to perform but do not assess how the exercise is being performed. Performing exercises using wrong postures might lead to injuries. Human pose estimation is a technique that uses Artificial Intelligence to detect the pose of a human and calculate angles between different parts of the body. Using human pose detection libraries like OpenPose, computer vision tools like OpenCV and geometry heuristics, appropriate posture correction suggestions can be given. This work aims to develop a solution that bridges the mentioned shortcoming of the remote fitness industry.

Keywords: Human Pose Estimation, Geometry Heuristics, Posture Correction, Remote Fitness Applications, Artificial intelligence, Computer Vision.

I. INTRODUCTION

Fitness exercises are very beneficial to personal health and fitness; however, they can also be ineffective and potentially dangerous if performed incorrectly by the user. Exercise mistakes are made when the user does not use the proper form or pose. This could be due to a lack of formal training through classes or a personal trainer, or could also be due to muscle fatigue or using too much weight. The incorrect form has become an increasingly problematic issue in the wake of COVID. Gyms are hotspots for viral transmission, and many beginners are electing to start their fitness journeys at home. However, without feedback from a personal trainer and minimal prior experience, it can be quite a difficult task to learn how to lift weights properly. There are many commercial solutions available that try to bridge this gap. Commercial applications have 2 shortcomings: 1) They're commercial and hence require payment. 2) They prescribe a set of exercises but never give feedback as to whether or not the exercise performed is right or wrong. This application is an attempt to overcome the disadvantages. Oftentimes, it is the case that people want to perform certain light exercises on their own without seeking the help of professionals. This is a booming industry - Remote Fitness. That does not mean that there is no threat presented. A guide to performing physical exercises in a safe way will reduce the number of injuries/ physical strains occurring. This application helps remote industry bridge all its gaps.

II. LITERATURE SURVEY

Authors in [1] gave a review paper on the most influential paradigms of HPE and how they are involved stage by stage. [1] describes different paradigms as Pictorial Structures for HPE, HPE using DNN, and Stacked Hourglass for HPE.
Authors in [2] gave a top-down approach by limiting the mistakes made while detecting highly entangled people using HrNet. [2] was carried out in three components: Clip Tracking Network, Video Tracking Pipeline, and Spatial-Temporal Merging procedure.This study resulted in successfully improving over the previous person detector approach of top-down approach by Spatial-Temporal Merging.This can recover over to 4-7% over missed predictions.
The idea of Part Affinity Fields (PAFs) which indicates the orientation of the line joining different joints was introduced by Authors in [3] which increased the accuracy of multi-person pose detection.[3] also has O(n^2) execution time.
Authors in [4] used a Bottom-up approach to identify image pose. Joints are identified using convolutional network architecture and distinguishing parts by hierarchy.In [4] improvement in HPE using high level spatial models took place.
Authors in [5] gave Pose estimation based on Deep Neural Networks (DNNs). The pose estimation is formulated as a DNN-based regression problem toward body joints. Localization of joints after normalization in sub-images.
Authors in [6] aimed at improving the user's exercise pose by using two different approaches of detecting by geometry and ML. [6] aims to work on four different exercises.It results in an end-to-end application that estimates using visual geometry and ML. This also provides Output with specific key points to improve.[6] uses Part Affinity Fields

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 7, July 2022
DOI: 10.17148/IJARCCE.2022.11792

Authors in [7] gave implementation of HPE for the niche of posture correction using OpenPose. [7] also implemented DTW. In [7] Posture implemented are Bicep curl,Front Raise, Shoulder Shrug, and Shoulder Press, a posture correcting application that specifies whether or not an exercise has been performed correctly. Not real time.Occlusions are not dealt with efficiently.

Authors in [8] Setup a system for Human Pose detection. Training of the dataset is done using CNN and deep learning. OpenCV is used for Image Processing and video processing.A posture detecting application that provides angle between different parts of shoulders and provides a graphical representation of different parts compared to the actual dataset. to detect multiple people during estimation. Previous models used a top-down approach or bottom-up approach to detect different people using bounding box techniques for the detection of people and ten joints.

III. DESIGN AND IMPLEMENTATION

The design implementation phase is a significant percentage of the overall design cycle of this project. It is critical that the implementation phase of the design be handled as efficiently as possible. The decisions before and during the design implementation phase can have a dramatic impact on the implemented design and project schedule.

A. Architecture

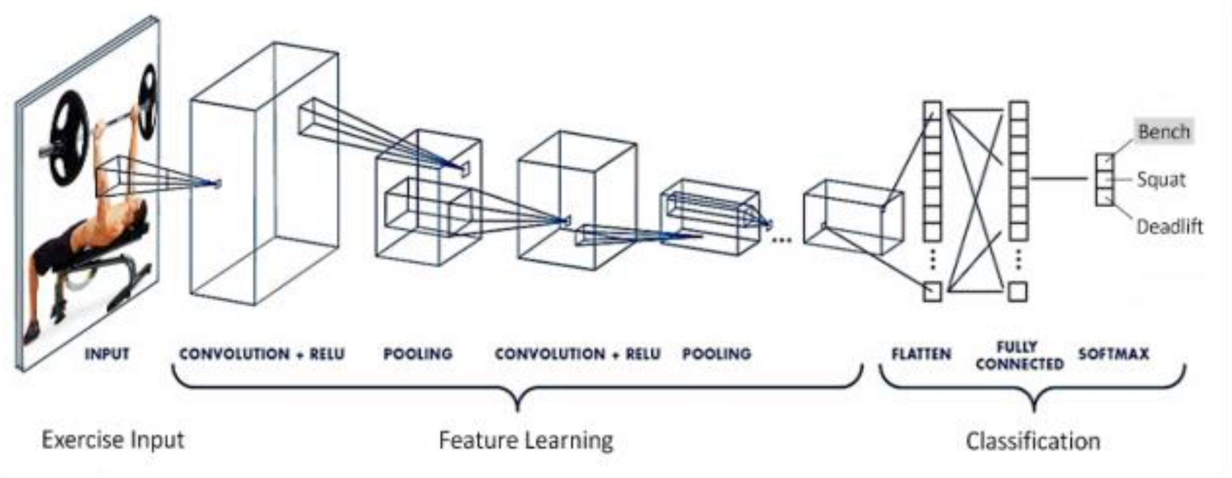

Fig. 3.1 Architecture of the basic functionality - Classification

This diagram shows the various stages at which convolution layers and pooling layers are used in order to arrive at the desired outcome. Both max pooling and average pooling are used here.

B. High - Level Diagram

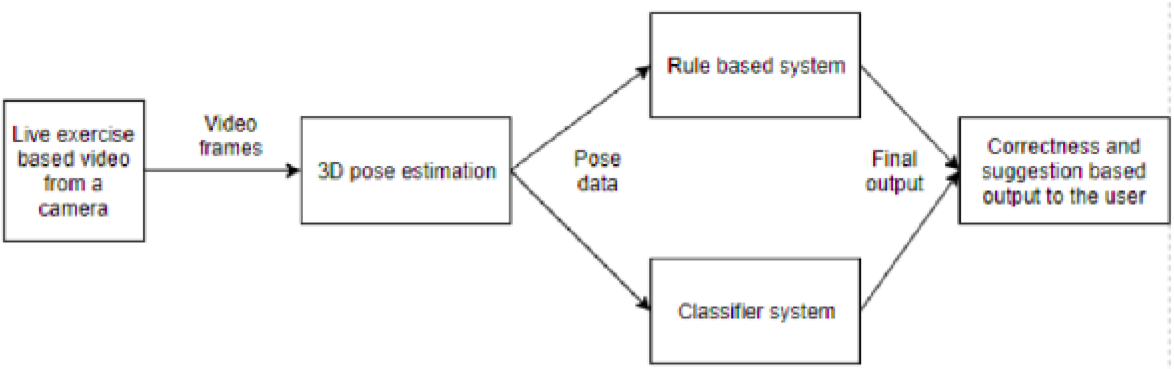

Fig. 3.2: Flowchart representation of steps involved in posture correction

This shows the flow of the decision-making process within the system.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⋇ Impact Factor 7.39 ⋇ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11792

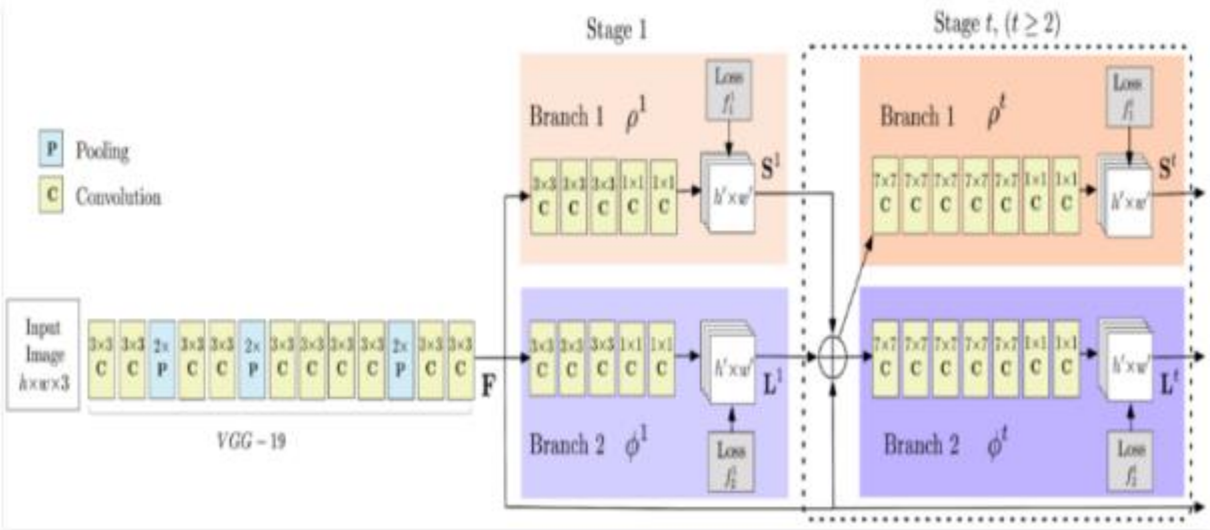

Fig. 3.3: CNN model (OpenPose) for pose detection

C. Low-Level Diagram

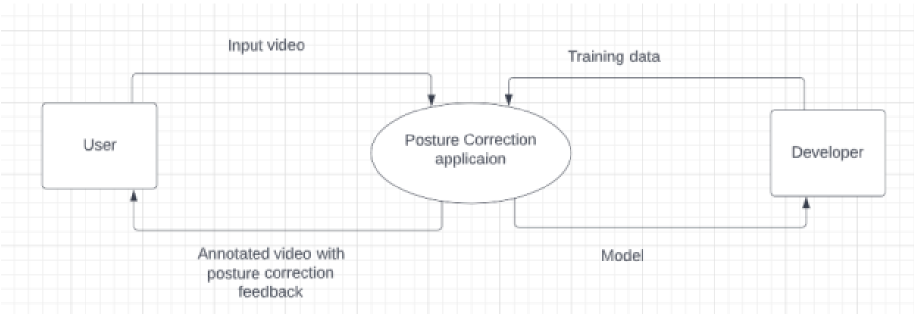

Fig. 3.4: Data flow diagram - Level 0 (for the project implementation)

D. Dataset

MPII: The MPII human pose dataset is a multi-person 2D Pose Estimation dataset comprising nearly 500 different human activities, collected from Youtube videos. MPII was the first dataset to contain such a diverse range of poses and the first dataset to launch a 2D Pose estimation challenge in 2014. Contains over 25k images with around 40k people.

E. Methodology

Our system will consist of a camera which may be of any type, for example, a mobile phone camera or a webcam from a laptop computer. The camera will be used to capture a normal RGB format video whose frames are fed into the 3D pose estimation model. The video will be such that the person performing an exercise routine is clearly visible. Pose estimation is performed to get the detailed pose data and this data is then fed to a rule-based system as well as a classification and comparison-based system to analyze the exercise routine. If the exercise is performed correctly, the repetition count of the exercise is automatically recorded and displayed in the user interface of our application. If the person does not perform the exercises according to the rules, then feedback is generated to notify the user and display the appropriate steps for correction.

General steps followed: Our project will consist of two major components - pose estimation and analysis of gym/exercise posture. The pose estimation part can be described as: Given a video of a person who is performing a given exercise, a 3d pose estimation model is built with the task of producing a 3D pose that matches the spatial position of the depicted person.

IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ≋ Impact Factor 7.39 ≋ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11792

Pose estimation model: OpenPose: OpenPose is one of the most popular bottom-up approaches for multi-person human pose estimation. As with many bottom-up approaches, OpenPose first detects parts (key- point) belonging to every person in the image, followed by assigning parts to distinct individuals. The OpenPose network first extracts features from an image using the first few layers (VGG- 19 in the above flowchart). The features are then fed into two parallel branches of convolutional layers. The first branch predicts a set of 18 confidence maps, with each map representing a particular part of the human pose skeleton. The second branch predicts a set of 38 Part Affinity Fields (PAFs) which represents the degree of association between parts. Successive stages are used to refine the predictions made by each branch. Using the part confidence maps, bipartite graphs are formed between pairs of parts (as shown in the above image). Using the PAF values, weaker links in bipartite graphs are pruned. Through the above steps, human pose skeletons can be estimated and assigned to every person in the image.

A rule-based system for exercise evaluation: For evaluating exercise posture, firstly, the output of pose estimation will have to be parsed into the proper format. The full skeletal estimation of the person is obtained and necessary normalization is performed. We can use the length of the torso as the standard for normalization. The perspective of the user is also determined. The rule-based system will be according to the standard guidelines of a particular exercise. Geometry-based heuristics and thresholds can be used to determine whether the given exercise has been performed correctly or not. Along with that, the actual mistake performed by the user can also be pointed out. For instance, in a bicep curl, we can identify two heuristics of interest. First, the upper arm should be kept steady and not move significantly. This can be quantified by the angle between the upper arm vector and the torso vector. If the upper arm is held steady, then it should be parallel to the torso with minor variations for the entire video. Second, a proper, complete curl requires the weight to be brought up until the bicep is fully contracted beyond the midway point (90between upper arm and forearm) that is commonly stopped at. This improper form is typically a result of the user using weights that are too heavy. We can quantify this problem by the minimum angle achieved between the upper arm and forearm. In the start position with the weight down, the angle should be nearly 180. As the weight is lifted, the angle should decrease until when the user stops, and increase again as the weight is brought down. Such rules can be used for many other exercises.

IV. CONCLUSION

OpenPose is a strong library used for human pose detection. A model with decent accuracy was developed that could detect the exercise being performed and give accurate suggestions based on the ML model with the help of geometry heuristics.

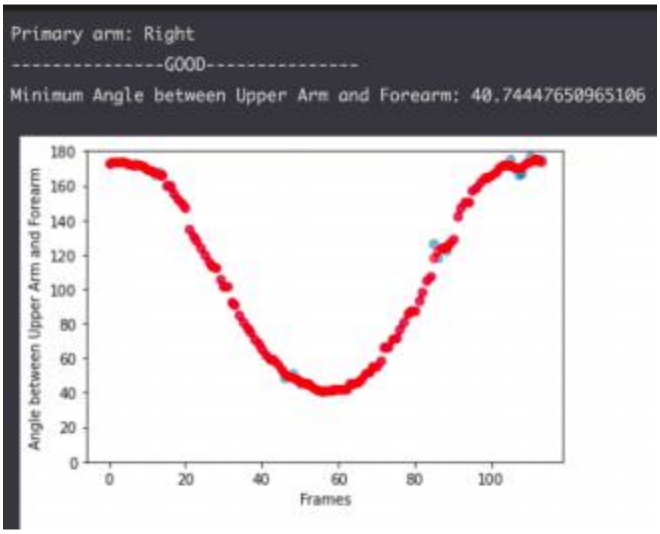

Fig. 4.1 Angle between upper arm and forearm for “Good” posture

Correct positioning involves training yourself to hold your body against gravity with the least strain and tension on supportive structures.In this project, an application is presented which provides feedback on human posture while performing exercises using pose detection, visual geometry, and machine learning. The output of this project is to locate the human body's key points from the video provided. ML algorithm is used for analyzing posture correctness and geometric algorithms for providing feedback on exercises performed.

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁂ Impact Factor 7.39 ⁂ Vol. 11, Issue 7, July 2022

DOI: 10.17148/IJARCCE.2022.11792

The problem of joint occlusion is still unsolved in very advanced models. Implementing this solution in real -time on a mobile device like a cell phone might prove to be very challenging. No framework has been developed that is lightweight to work on a mobile phone at the same time very efficiently.

We have identified several extensions as strong opportunities for future work. One path would be to export Pose Trainer to smartphones, building an application that allows users to record a video and get pose feedback at any place or time. Another direction would be to improve the pose feedback, providing specific suggestions on where the user's pose needs improvement (e.g., back, neck, shoulders), and suggesting targeted action. Finally, we could work on improved graphics, for instance, showing the user their labeled pose diagram, and comparing it to the labeled pose diagram of a ground truth trainer.

REFERENCES

[1] D. Berndt and J. Clifford. Using dynamic time warping to find patterns in time series. In AAAIWS, 1994.
[2] F. Bogo, A. Kanazawa, C. Lassner, P. Gehler, J. Romero, and M. Black. Keep it smpl: Automatic estimation of 3d human pose and shape from a single image. In ECCV, 2016.
[3] G. Papandreou, T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. Murphy. Towards accurate multi-person pose estimation in the wild. In CVPR, 2017.
[4] J. Shotton, A. Fitzgibbon, M. Cook, T. Sharp, M. Finocchio, R. Moore, A. Kipman, and A. Blake. Real-time human pose recognition in parts from single-depth images. In CVPR, 2011.
[5] Z. Cao, T. Simon, S. Wei, and Y. Sheikh. Realtime multi-person 2d pose estimation using part affinity fields. In CVPR, 2017.
[6] A. Newell, K. Yang, and J. Deng. Stacked hourglass networks for human pose estimation. In ECCV, 2016.
[7] A. Toshev and C. Szegedy. Deeppose: Human pose estimation via deep neural networks. In CVPR, 2014.
[8] P. Zell, B. Wandt, and B. Rosenhahn. Joint 3d human motion capture and physical analysis from monocular videos. In CVPR Workshops, 2017.
[9] Steven Chen & Richard Yang (2018). Pose Trainer: Correcting Exercise Posture using Pose Estimation.
[10] W. Yang, W. Ouyang, H. Li and X. Wang, "End-to-End Learning of Deformable Mixture of Parts and Deep Convolutional Neural Networks for Human Pose Estimation," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 3073- 3082, doi: 10.1109/CVPR.2016.335.
[11] X. Chu, W. Ouyang, H. Li and X. Wang, "Structured Feature Learning for Pose Estimation," 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2016, pp. 4715-4723, doi: 10.1109/CVPR.2016.510.
[12] T. L. Munea, Y. Z. Jembre, H. T. Weldegebriel, L. Chen, C. Huang and C. Yang, "The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation," in IEEE Access, vol. 8, pp. 133330-133348, 2020, doi:10.1109/ACCESS.2020.3010248.
[13] X. Chen, and A. Yuille. "Articulated Pose Estimation by a Graphical Model with Image Dependent Pairwise Relations." Eprint Arxiv(2014):1736-1744.
[14] W. Rawat and Z. Wang. (2017). Deep Convolutional Neural Networks for Image Classification: A Comprehensive Review. Neural Computation. 29. 1-98. 10.1162/NECOa00990
[15] S. Wei, V. Ramakrishna, T. Kanade and Y. Sheikh, "Convolutional Pose Machines," in 2016 IEEE Conference on Computer Vision and Pattern Recognition (CVPR), Las Vegas, NV, USA, 2016 pp. 4724-