Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

AI Posture Trainer Using Open CV and Media Pipe

Mr K Shankar Associate Professor, Dept. of Computer Science and Engineering,Nadimpalli Satyanarayana Raju institute of technology, Sontyam , Visakhapatnam

P Chandra Kiran 2 **, G Swaroop Ganesh** 3 **, Msvn Praveen** 4 **,CH Bhavya Shree** Student, Dept. of Computer Science and Engineering,Nadimpalli Satyanarayana Raju instituteof technology, Sontyam , Visakhapatnam

#219nu1a0588@nsrit.edu.in, #319nu1a0574@nsrit.edu.in, #4 20nu5a0506@nsrit.edu.in , #5 19nu1a05a9@nsrit.edu.in
#1 shankar.cse@nsrit.edu.in

---***---

Abstract - *Recently, a lot of people are facing many health issues due to the lack of physical activity. During the pandemic, with social distancing being the major factor, so many workout centers were closed. So people started training at home in the absence of a trainer. Training without a trainer often leads to serious internal and external injuries if a specific workout is not properly done. Our project offers multiple features that can benefit the users to attain their ideal body by providing them with a personalized trainer with a personalized workout and a customized diet plan. We also help our users to connect with people who have similar workout goals because the motivation behind working out increases if we have a workout partner. To achieve this, we are using OpenCV and Media Pipe to identify the user's posture and analyze the angles and the geometry of the pose from the real-time video, and we are using Flask to develop the front end of the web application.*

***Key Words*: BlazePose, Media Pipe, Open CV, COCO Data Set, Open Pose.**

1. INTRODUCTION

We are all aware of how important exercise is to our overall health. Additionally, it is crucial to exercise correctly. Exercising too frequently can result in major damage to the body, including muscle tears, and can even reduce muscular hypertrophy. Nowadays, home workouts are more popular. Additionally, it saves time and is quite convenient.

Also at that time of the pandemic, we understood how important is fitness and how such situations can let us work out in our homes. Sometimes people cannot afford a gym membership and are sometimes shy to work out in the gym and use weights. On the other hand, sometimes people can afford gym and trainers but because of tight schedules and inconsistency, they are not able to spend time on their body and fitness.

At that point, AI personal trainers started to appear. Since there are already so many digital fitness programmers and digital coaches, the term "AI personal trainer" is no longer a fresh one.

To begin, let's define an AI personal trainer for those who are unfamiliar with it. "AI personal trainers are artificial intelligence-powered virtual trainers who assist you in achieving your fitness goals. The computerized personal trainer may provide you with tailored training and diet regimens after gathering a few facts such as body measurements, current fitness level, fitness objectives, and more."

Nowadays, Every person needs a customized trainer, which takes time and money to provide. Artificial Intelligence technology, therefore be used to speed up the customizing process by determining the best exercise regimen for a certain student's demands or preferences.

Therefore, our goal is to create an AI-based trainer that will enable everyone to exercise more effectively in the comfort of their own homes. The aim of this project is to build an AI that will assist you in exercising by calculating the quantity and quality of repetitions using pose estimation. This project is intended to make exercise easier and more fun by correcting the posture of the human body and also by connecting with people having similar workout goals.

This study presented an objective discussion of the usage of AI technology to select a suitable virtual fitness trainer based on user-submitted attributes.

2. RELATED WORK

Initially, it started with body posture detection proposed by Jamie Shotton et al., [1] who used the Kinect Camera, which produces 640x480 images at a frame rate of 30 frames per second with a few centimeters of depth resolution. The depth image characteristics show that pixels are being identified. These characteristics work well together in a decision forest to distinguish all trained sections even though individually they merely provide a weak indication of which region of the body the pixel is located. However, the examples demonstrate a failure to recognize minute adjustments in the images, like crossed arms.

Later on, Sebastian Baumbach et al., [2] introduced a system that applies several machine learning techniques and deep learning models to evaluate activity recognition for sporting equipment in contemporary gyms. In terms of machine learning models, Decision trees, Linear SVM, and Naive Bayes with Gaussian kernel are used. For our exercise recognition (DER), they presented a Deep Neural network with three hidden layers with 150 LSTM cells in each layer. 92% of the time was accurate. The biggest issue, however, is the overlap of workouts for the same body component, such as the Fly and Rear delt as well as the Pull Down, Lat Pull and Overhead Press.

To improvise the system proposed previously, Steven Chen et al., [3] used deep convoluted neural networks (CNNs) to label RGB images. They made advantage of the trained model, Open-Pose, for pose detection. The model consists of multiple-stage CNN with two branches: one branch is used to learn the part affinity fields, while the other branch is used to learn the confidence mapping of a key point on an image. But this model has its own drawbacks too, i.e it works only for pre-recorded videos.

In order to provide real-time detection CE ZHENG et al., [4] suggested a model that is categorized into 3 different models, they are kinematic, planar, and volumetric. For 2d HPE, to ensure that there is only one person in each cropped area, the input image is first cropped. Regression approaches and heatmap-based methods are the two main categories for single-person pipelines that use deep learning techniques. Regression approaches use an end-end framework to learn a mapping from the input image to body joints or characteristics of human body models. Predicting approximate positions of body parts and joints that are supervised by heatmap representation is the aim of heatmap-based algorithms. This model relies on motion capturing systems, which are difficult to set up in a random setting and are required for high-quality 3D ground truth posture annotations. Additionally, person detectors in top-down 2D HPE approaches could be unable to recognise the edges of heavily overlapping images.

This problem of system failure to identify the boundaries of largely overlapped images, Anubhav Singh et al., [5] used a hereditary convolutional neural network can be created as a solution to the human posture estimation problem using a convolutional neural network and a regression approach. the CNN method for extrapolating 2D human postures from a single photograph. Many methodologies and procedures have recently been developed for the evaluation of human postures that use postures from physiologically driven graphic models. Using CNN, the arrangement of image patches is used to discover the positioning of the joints in the picture. This method achieves both the joint recognizable proof, which determines if an image contains freeze body joints, and also the joint constraint in an image plot indicates the proper location of joints. These joint limitation data are then pooled for posture surveying. In contrast to nearby locators, which are constrained to a specific part overlapped human bodies, CNN has the advantage of accepting the entire scene as an info motion for each body point.

To increase the prior models' accuracy, Amit Nagarkoti et al., [6] suggested a system that attacks the issue using methods from vision-based deep learning. To create a fixed size vector representation for a particular image, the network first employs the starting ten layers of the VGG19 network. Thereafter, two multi-step branches of CNN and OpenCV for optical flow tracking are used. However this system works only for motion along 2 dimensions.

Sheng Jin et al., [7] thought that since there are currently no datasets with whole-body annotations, earlier approaches had to combine deep models that had been independently trained on several datasets of the human face, hand, and body while contending with dataset biases and high model complexity. They introduced COCO-Whole Body, an extension of the COCO dataset that adds whole-body annotations, to close this gap. The hierarchy of the entire human body is taken into account using a single-network model called Zoom Net in order to address the scale variance of various body sections of the same person. On the suggested COCO-Whole Body dataset, Zoom Net performs noticeably better than the competition. The robust pre-training dataset COCO-Whole Body can be utilized for a variety of applications, including hand keypoint estimation and facial landmark identification, in addition to training deep models from scratch for whole-body pose estimation.

And after years, this methodology has been upgraded by classifying the human body into various keypoints. Valentin Bazarevsky et al., [8] used a detector tracker setup which, in real time, excels at a wide range of jobs., including dense face and hand landmark prediction. They have a pipeline with a network of pose tracking at the beginning and a light body pose recognition at the end. The tracker predicts a fine-grained zone of interest for the present frame, the person's presence on the present frame, and the coordinates of important points. All 33 of the person's important points are predicted by the system's pose estimate component. They used a heatmap, offset, and regression strategy all at once. Blaze Pose performs 25-75 times faster than the Open Pose and a t the same time the performance of Blaze Pose is slightly worse than the Open Pose.

Danish Sheikh et al., [9] abstracted the technique of exploitation create estimation abstract thought out-put as input for associate LSTM classifier into a toolkit referred to as Action-AI. For video process demo, Open-CV suffices. For create estimation they used Open-pose enforced with in style deep learning frameworks like Tensor-flow and Py-Torch. The user can start, stop, pause, and restart yoga by utilising the voice interface, which uses the Snips AIR voice assistant. This model produced good results with high precision, but in certain cases, when important points are rotated, angles do not change.

Gourangi Taware et al., [10] used java script, node js, and many libraries, including open cv, a library that utilises ML techniques in addition to various arithmetic and algorithms. This strategy employs an effective two step-tracker machine learning method. The location of the activity or posture in the live video is determined while utilising the tracker. It then forecasts the crucial moments in the targeted area using the input from most recent video frame. But it's important to keep in mind that the tracker is only activated at the start or when the model fails to recognise the body key points in the frame. However, in the real-time system, the programme is unable to catch numerous people in the frame.

Shwetank Kardam et al., [11] To make their detections possible, firstly they recolored the images because OpenCV renders the RGB image to BGR color format but for Media Pipe to work, they need to convert our BGR image back to RGB. Lastly, change the color format back to BGR format as OpenCV runs on BGR format, and then

they started the detections. There are 33 landmarks in total, starting from index 0. These represent the different joints within the pose. For instance, if they want to calculate the angle for our Right hand's bicep curl, they will require the joints of shoulder, elbow and wrist which are 12, 14 and 16 respectively.

Shiwali Mohan et al., [12] designed an intelligent coach. The coach has a sophisticated long-term memory that keeps records of their encounters with the trainee in the past as well as how they did on activities the coach had previously suggested. A relational database called POSTGRESQL and objective-C code are used to create this memory, which saves data gleaned from user interactions. The model is immutable. If the initial capability is grossly overstated, the coach may never be able to correct the mistake.

Neeraj Rajput et al., [13] suggested a set of Python bindings to address problems in computer vision. This library makes use of NumPy, and all the structures of arrays easily change between NumPy arrays and other formats. This implies that combining it with other Python libraries, such as SciPy and Matplotlib, won't be challenging. (these make use of NumPy). In order to obtain the correct points and the desired angles, they made use of the CPU's posture estimate. Based on these angles, a variety of actions are then found, such as the quantity of biceps curls. With just one line of code, they were able to calculate the angles between any three locations. But some pose estimation algorithms that performs the 2D detections don't have enough accuracy.

From the previous models, it was clearly understood that the use of OpenCV, key points for the human body and implementing the COCO dataset has delivered great accuracies, so Harshwardhan Pardeshi et al., [14] tracked the number of user performed exercises and spot flaws in the yoga position. They used the basic CNN network, the OpenPose Python package, and the COCO datasets. This model performed very well regarding the accuracy and performance but there is no variations in different Depending on their age category, there are categories for men, women, and kids.

Recently, Yejin Kwon et al., [15] a programme that estimates the posture of real-time photos to direct the content of the training (squats, push-ups), the number of workouts, and OpenCV and Media Pipe. Media Pipe simulates the real-time analysis and this model delivered excellent accuracy and performance, but the only problem with this model is there is no real-time support for the users if they have any queries and also there maybe a chance when user lacks the motivation of working out alone.So we are proposing a system that overcomes the important disadvantage of not being able to work out at home and at any time without guidance. The system provides us the opportunity to work out anywhere, anytime with guidance so we can do effective workouts. Mainly, This system overcomes the disadvantage of lack of real-time support to answer their queries and also this connects the users having similar workout goals so that they can workout together which improves their

3. PROPOSED METHODOLOGY

From the above information, we can say that AI can be used for creating a good personal trainer because by using AI we can not only suggest the exercises but also correct the posture of the person each and every time he poses a wrong posture.

Industrial Engineering Journal
ISSN: 0970-2555
Volume : 52, Issue 4, April : 2023

motivation. So, this System uses computer vision technology to execute the functions of our system. The system uses state-of-the-art pose detection tool known as “Blaze Pose” from “Media Pipe” to detect the form of the user’s body during the workout. OpenCV is used to mark an exoskeleton on the user body and display reps count on the screen. We are using Flask to develop the Front-end for the We also have chatbot that answers user's queries which we implement by using Chatter Bot.

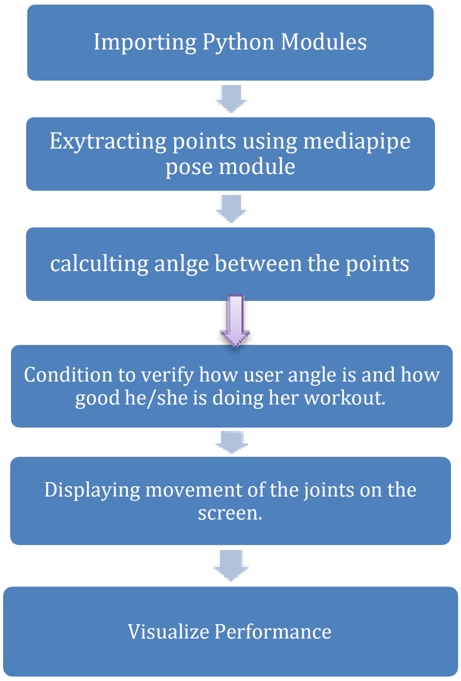
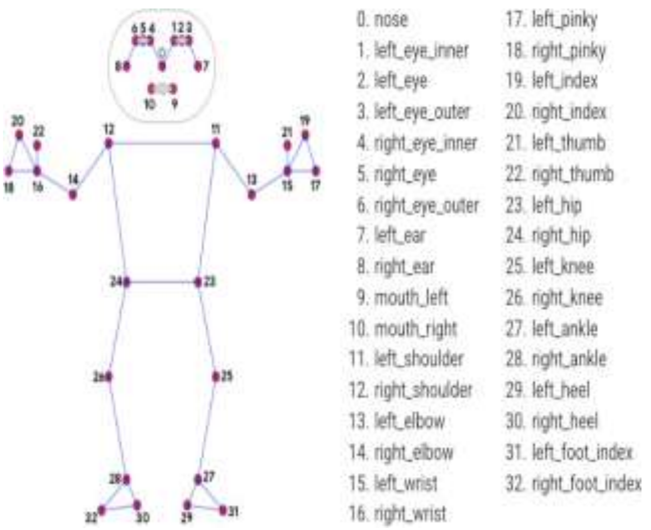


Fig.3.1 Proposed Methodology Framework



- The modules/packages which we are using for this project are Numpy, Media Pipe, Flask, ChatterBot, and Open CV.

- For extracting the data points There are 33 landmarks in total, starting from index 0. These represent the different joints within the pose. For instance, if we want to calculate the angle for our Right hand’s bicep curl, we would require the joints of the shoulder, elbow, and wrist which are 12, 14, and 16 respectively as referred to in the below figure.

To calculate the angle between joints first, We obtain the three joints' coordinates, which are necessary to calculate the angle. Then, using NumPy, we can determine the slopes of the joints. Angles are measured in radians and can be translated into degrees.

4. CONCLUSIONS

We conducted extensive research for this paper on AI Personal Trainer, looking at a wide range of research publications on Personal trainer at home. We learned that most of the models were not able to provide the real-time interactions with the user. We understood that there is no such platform where users can interact with others who have similar workout goals which can improve the efficiency and motivation of the user.

Additionally, we have learned about several proposed and existing systems through research publications, which has helped us develop a new model that would make translation much more efficient.

ACKNOWLEDGEMENT

Special thanks to our team guide, Mrs. Kriti Ohri, for all of her support and direction, which helped the literature survey portion of the project be successfully completed and yield positive results at the end.

REFERENCES

[1] Swarangi, Ishan, Sudha.(2022). Personalized AI Dietitian and Fitness Trainer.

https://www.ijraset.com/research-paper/personalized-ai-dietitian-and-fitness-trainer

[2] Ce Zheng, Wenhan Wu, Chen Chen, Taojiannan Yang, Sijie Zhu, Ju Shen, Nasser Kehtarnavaz, Mubarak Shah(2020). Deep learning based human pose estimation.https://arxiv.org/abs/2012.13392

[3] Jamie Shotton,Andrew Fitzgibbon,Mat Cook,Toby Sharp,Mark Finocchio,Richard Moore Alex Kipman,Andrew Blake(2011). Real-Time Human Pose Recognition in Parts from Single Depth Images.

https://www.microsoft.com/en-us/research/wp-content/uploads/2016/02/BodyPartRecognition.pdf

[4] ValentinBazarevsky,IvanGrishchenko,Karthik Raveendran,Tyler Zhu,Fan Zhang,Matthias Grundmann(2020).BlazePose: On-device Real-time Body Pose tracking.

https://arxiv.org/pdf/2006.10204.pdf

[5] Sheng Jin, Xu,Lumin Xu,Can Wang,Wentao luo,Chen Qian,Wanli ouyang(2020).Whole-Body Human Pose Estimation in thr wild. https://dl.acm.org/doi/abs/10.1007/978-3-030-58545-7_12

[6] Gourangi Taware , Rohit Agrawal , Pratik Dhende , Prathamesh Jondhalekar, Shailesh Hule(2021).AI Based Workout Assistant and Fitness Guide.

https://www.ijert.org/ai-based-workout-assistant-and-fitness-guide

[7] Aarti, Danish Sheikh , Yasmin Ansari , Chetan Sharma , Harishchandra Naik(2021).Pose Estimate Based Yoga Trainer.AI-based Workout Assistant and Fitness guide – IJERTz

[8] Harshwardhan Pardeshi, Aishwarya Ghaiwat, Ankeet Thongire,Kiran Gawande , Meghana Naik(2022).Fitness Freaks: A System For Detecting Definite Body Posture Using Open Pose Estimation.

https://link.springer.com/chapter/10.1007/978-981-19-5037-7_76

[9] Yeji Kwon,Dongho Kim(2022).Real-Time Workout Posture Correction using OpenCV and MediaPipe.

https://www.researchgate.net/publication/358421361_RealTime_Workout_Posture_Correction_using_OpenCV_and_MediaPipe

[10] Steven Chen, Richard R. Yang(2018). Pose Trainer: Correcting Exercise Posture using Pose Estimation.

https://arxiv.org/abs/2006.11718

[11] Amit Nagarkoti,Revant Teotia,Amith K. Mahale, Pankaj K. Das(2019).Real-time Indoor Workout Analysis Using Machine Learning & Computer Vision.

https://ieeexplore.ieee.org/document/8856547

[12] Ali Rohan,Mohammed Rabah,Tarek Hosny,sung-Ho Kim(2019) .Human Pose Estimation Using Convolutional Neural Networks

https://ieeexplore.ieee.org/document/9220146

[13] Shwetank Kardam,Dr.Sunil Maggu(2021). AI Personal Trainer using OpenCV and Python.

http://www.ijaresm.com/

[14] Sherif Sakr , Radwa Elshawi, Amjad Ahmed, Waqas T. Qureshi, Clinton Brawner, Steven Keteyian, Michael J. Blaha, Mouaz H. Al-Mallah(2018). Using machine learning on cardiorespiratory fitness data for predicting hypertension: The Henry Ford ExercIse Testing . https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0195344

[15] Alp Güler, R., Neverova, N., Kokkinos, I(2018).DensePose: dense human pose estimation in the wild. https://scholar.google.com/scholar?hl=en&q=Alp+G%C3%BCler%2C+R.%2C+Neverova%2C+N.%2C+Kokkinos%2C+I.%3A+DensePose%3A+dense+human+pose+estimation+in+the+wild.+In%3A+Proceedings+of+the+IEEE+Conference+on+Computer+Vision+and+Pattern+Recognition+%28CVPR%29+%282018%29

[16] Burgos-Artizzu, X.P., Perona, P., Dollár, P(2013). Robust face landmark estimation under occlusion. https://openaccess.thecvf.com/content_iccv_2013/html/Burgos-Artizzu_Robust_Face_Landmark_2013_ICCV_paper.html

[17] Cao, Z., Simon, T., Wei, S.E., Sheikh, Y(2017).Realtime multi-person 2D pose estimation using part affinity fields. https://openaccess.thecvf.com/content_cvpr_2017/html/Cao_Realtime_Multi-Person_2D_CVPR_2017_paper.html

[18] Gomez-Donoso, F., Orts-Escolano, S., Cazorla, M.Large-scale multiview 3D hand pose dataset. https://arxiv.org/abs/1707.03742

[19] Gross, R., Matthews, I., Cohn, J., Kanade, T., Baker, S(2010). Multi-pie. In: Image and Vision Computinghttps://www.sciencedirect.com/science/article/pii/S0262885609001711

[20] Insafutdinov EPishchulin LAndres BAndriluka MSchiele BLeibe BMatas JSebe NWelling M(2016).DeeperCut: a deeper, stronger, and faster multi-person pose estimation mode. https://link.springer.com/chapter/10.1007/978-3-319-46466-4_3