IJCRT.ORG ISSN : 2320-2882

INTERNATIONAL JOURNAL OF CREATIVE RESEARCH THOUGHTS (IJCRT)

An International Open Access, Peer-reviewed, Refereed Journal

Virtual Assistance For Physical Fitness Using HumanPose Estimation

[1]Radha R, [2]Dhanush Kumar R, [3]Shalini M B, [4]Sharmila A, [5]Sumalatha B R

[1]Assistant Professor, [2]Student, [3]Student, [4]Student, [5]Student[1]Information Science and Engineering,
[1]Sai Vidya Institute of Technology, Bangalore, Karnataka ,India

Abstract: Approximately 39 percent of adult people on the planet are overweight. The fact presented above makes it clear how important and necessary exercise is. Exercise helps us maintain a healthy weight, a fit body, and a calm mind in addition to helping us lose weight. Regular exercise also keeps us active and improves blood circulation to keep our bodies in the same condition as when we used to visit the gym. It is expensive or not within everyone's reach to train under a trainer, visit the gym, take yoga courses, or both. Self-training is an additional option that provides pre-recorded yoga practice steps without any feedback. Without proper feedback about our postures, injuries can happen and it will do more harm than good and that's exactly where our project comes into play. These days, human position estimation is a widely pursued project in computer vision. The study of strategies and systems that retrieve an articulated body's stance is known as articulated pose estimation in computer vision. The course of determining the human body's location parts and joints in a given image is known as "articulated body pose estimation" in the context of this study. We study the many applications that we may put into practice with the data, which we received using a pre-trained posture estimation model called MediaPipe. Motion capture, gait analysis, anomaly detection, sign language recognition, and other uses are among them.

***Index Terms* -** Human pose estimation, pose detection, pose estimation survey.

I. INTRODUCTION

By feeding the system with sample data or trained models, artificial intelligence or machine learning may be used to construct human posture estimation. This will allow the system to localize joints in the human body during the course of one video or picture. Now that we know where the joints in the human body are located, we can utilize this information for a variety of purposes, such figuring out a person's walking gait cycle or following a professional athlete around to learn about the tactics and methods they employ to attain their success. We might ascertain a person's location at significant points by employing pose estimation algorithms. We would therefore be able to assess orestimate the human body's posture using this and provide comments accordingly. The input video from the camera and the database's captured information are compared in our project. The image that was entered and its pose will be judged to be accurate if both are the same; if not, appropriate instruction is going to be provided to correct it.

II. LITERATURE SURVEY

The following journals and research papers were surveyed for the project. They provided information on the current state of pose estimation in general additionally other key areas of image processing and semantic segmentation regarding our project additionally the various approaches taken to achieve the goal of real-time articulated pose estimation.

Santosh Kumar Yadav et al. [1] built a system where they recognized the various Yoga asanas by using deep-learning algorithms. Here they used Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) architecture to be able to recognize yoga poses by using a stream of realtime images. A model that incorporates temporal data makes use of the data from earlier frames to get a reliable and accurate outcome.

S Khan et al. [2] proposed a system that can detect people with the help of multiple uncalibrated cameras. That is when a person is detected in one camera, the system is able to identify in which all cameras this same person is present. It presents a method to automatically recover these lines by observing motion in the environment and analyzing those images, furthermore once these lines are initialized the homographs between the views can also be recovered, it presently provides results on indoor and outdoor or sequence containing persons and vehicles.

Chen Qian et al. [3] proposed a system for robust real-time hand tracking using depth sensors. Hand tracking is a highly complex task because of small size of the fingers and the speed at which the fingers move under a large view-point. The system tracks fully articulated hands in real-time under a large view-point with a performance of about 25 frames per second without using a Graphical Processing Unit.

C. Sminchisescu et al. [4] proposed a system which uses a simple kinetic reasoning for the forwards or backward flips so it could speed up within its linked group of minima named kinematic jump process for monocular 3D human tracking. This can be utilised deterministically for jump diffusion style research. This model uses various coordinates for the purpose of representing the joints. For the easiness of the system, they use algorithms to produce the result.

Shruthi Kothari et al. [5] proposed yoga pose detection using deep learning where in, with the help of deep learning and machine learning yoga poses are classified with the assistance of pre recorded video additionally in reality. A self-instruction exercise system that enables people to learn and perform exercises correctly on their own can be developed using human pose recognition.

Zhe Cao et al. [6] proposes the OpenPose real time multi-person 2D pose estimation using Parity Affinity Fields (PAF). The 2D pose of multiple people can be detected in an image using a real-time approach. This system achieves accuracy in high level and real time performance. It uses a non-parametric representation method to be able to identify body parts of each individual in the given image.

Alex Kendall et al. [7] proposes Posenet which is a conventional network. It uses a six-degree monocular relocalization system which is also real time. This contains neural networks that are convolutional to be able to obtain an image that is single. The algorithm can function indoors as well as outdoors.

III. METHODOLOGY

3.1 Basic Block Diagram

The fundamental block diagram in Figure 3.1.1 provides information about the relationships and connections among the different system components .As seen in the figure, an image is captured by a camera and then delivered to the system for processing.
A basic block diagram for human pose correction includes several key components. It starts with an input image containing a person in a specific position. This image is processed by a pose detection model, which identifies key points on the person's body, such as joints.

The next step involves calculating the angles between these key points to support the pose. Depending on predefined criteria, a pose assessment module determines if the pose needs correction. If correction is necessary, a correction algorithm determines howto adjust the pose.

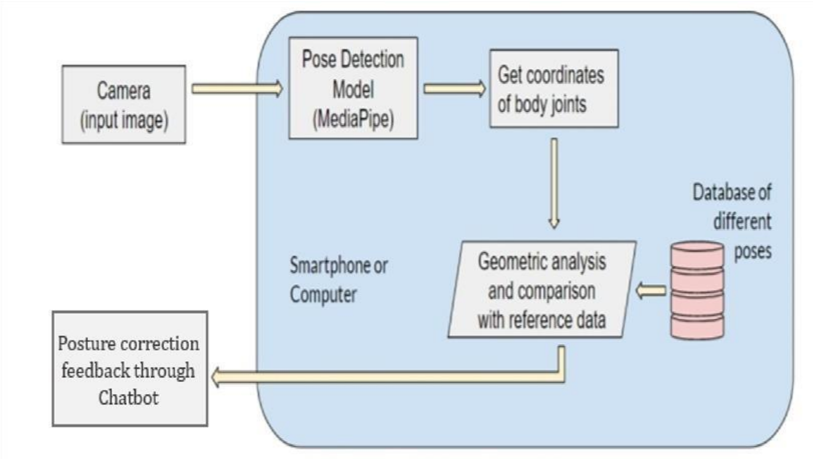


Figure 3.1.1: block diagram

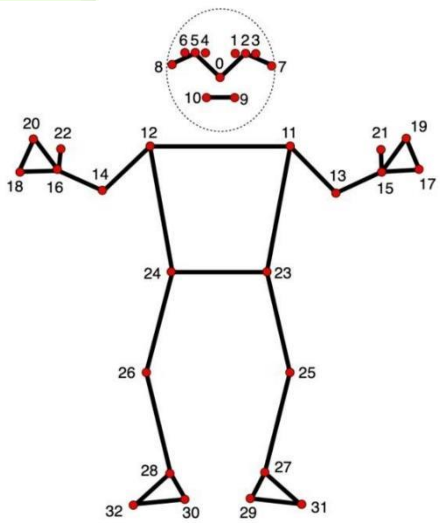
The pose adjustment module then physically adjusts the key points to fix the pose. Each component in the block diagram plays a vital role in the course of human pose correction, ensuring that the final pose is correct and meets the desired criteria.

3.2 Image acquisition

The image is taken as input through a camera which can be either a separate camera module for the raspberry pi board, namely the py-camera or a smartphone camera which is now ubiquitously available or a webcam which is also a good way for capturing images since most people have any one of this type of camera input solution. Camera acts as the input component of the system. The camera source can be from a webcam, mobile camera or a separate camera module. Camera is employed for image acquisition and as data input to the model. Our project is compatible with an RGB (Red, Green, Blue) camera. A reference square box is shown on screen and the user is asked to stand at a particular distance so that his/her whole body is within this square boundary. The built-in camera or a separate camera module is employed in order to obtain image of the user continuously during the routine which is then sent to the system (smartphone or computer) for processing.

3.3 Pose estimation

To identify the user's body's main points, the input picture is sent to the MediaPipe library. A list of coordinates in the X, Y, and Z axes for 33 essential vital places on the human body is the result. Each main body part's arrangement in the input picture can be determined by the resulting set of coordinates. By using



these data, we are able to precisely construct the user's skeleton competition orientation**.**

Figure 3.3.1 Pose Landmark Detection

the markers point to the main limbs and body regions of humans. As indicated in Table 1, they are ranged from 0 to 32, significance that there are 33 landmarks in total that are output by the MediaPipe library. The visage land marking procedure uses the first 11 landmarks, labeled 0 through 10.

table 1:the key points denoted in mediapipe documentation.

0. nose	17. left pinky
1. left eye inner	18. right pinky
2. left eye	19. left index
3. left eye outer	20. right index
4. right eye inner	21. left thumb
5. right eye	22. right thumb
(1) 6. right eye outer	23. left hip
7. left ear	24. right hip
8. right ear	25. left knee
9. mouth left	26. right knee
10. mouth right	27. left ankle
11. left shoulder	28. right ankle
12. right shoulder wrist	29. left heel
13. left elbow	30. right heel
14. right elbow	31. left foot index
15. left wrist	32. right foot index
16. right wrist	

These landmarks, additionally referred to as key points, enable us to recognise the face and its trajectory in an image. The upper body recognition technique is accustomed to figure out the next 11 landmarks, designated 11 through 22. The upper body comprises the elbows, wrists, hands, and an estimated three fingers from each hand: the thumb, index finger, and pinky finger. The lower body, comprised of the hips, knees, legs, and foot, is outlined by the final 11 significant points or landmarks, which run from 23 to 32. When combined, they provide a comprehension of the human body's direction in three dimensions additionally the overall framework in the visualization.

3.4 Database of Yoga Pose

Instead of keeping graphics of these stretches in the database, each stance is represented by a sequence of angles made at various joints. For every yoga stance, these angles have been saved as text data in the database. To train, the user chooses a yoga stance from the screen. The frame is taken once the user strikes a posture facing the camera, and patterned evaluation is performed using the MediaPipe Library's data. The mathematical evaluation yields the angles formed between each joint. The reference angles for an individual yoga establish that are already in the database are subsequently contrasted to these angles.

3.5 Pose Comparison

All that is included in the result of the MediaPipe library are the coordinates for the user's primary picture key locations. A function is coded within the program to obtain this coordinate values and subsequently compute the angles at every joint, such as the elbows, shoulders, hips, and so forth. With the application of analytical geometry, we can quickly determine the angle formed between the two lines given three essential points. Prior to data compilation, each yoga pose's angles are computed and entered into the database. To obtain every aspect for a certain yoga position, this analysis is performed. The user commences by deciding the yoga stance they might desire to execute. The angles produced by the position's creation are taken into consideration by the user and in contrast with the angles data of that one particular yoga contention that has previously been saved in the database for reference. In real time, the user acquires announcement on the display when their perspective angle stray from the reference data.

3.6 Feedback to the user

It is crucial to provide suggestions to the user so that they comprehend what that they are doing wrongly. This aids in demonstrating to the user whether to develop appropriate posture and accomplish the yoga orientation. Via the display, the user obtains real-time feedback concerning their performance. The user receive notification when their variance exceeds the threshold value. For the purpose of to correctly perform the yoga exercise, users can witness the correction and adapt the posture as desired.

IV. FLOW CHART

An input image of a person in a specific stance is the first step in the flowchart for human pose correction. To identify important body parts, such joints, this image is processed using a posture detection model. To evaluate the pose, the angles between these crucial locations are computed after they have been detected.

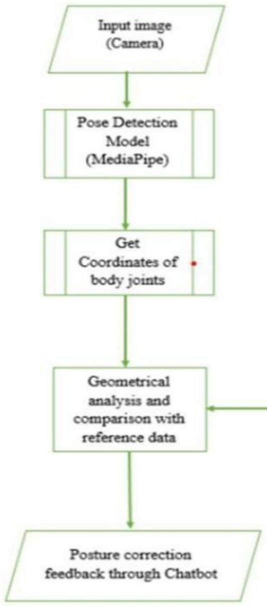


Figure 4.1: Flow chart for human pose detection and correction

Subsequently, the position is evaluated to ascertain whether correction is required, based on predetermined criteria such as angles falling outside of a specific range. If the pose has to be corrected, a plan is made to change it, which can include moving some joints.

After that, the correction is physically applied to the identified key spots to fix the stance. Lastly, the output image shows the adjusted pose. This flowchart shows the fundamental steps involved in identifying and adjusting human position, which may be applied to a variety of applications like animation, sports analysis, and fitness tracking.

V. USE CASE DIAGRAM

Fig 5.1 depicts the project's use case diagram-Virtual Assistance for Physical Fitness using Human Pose Estimation. this use case diagram illustrates the various interactions between the user and the system components involved in the procedure of human pose correction.

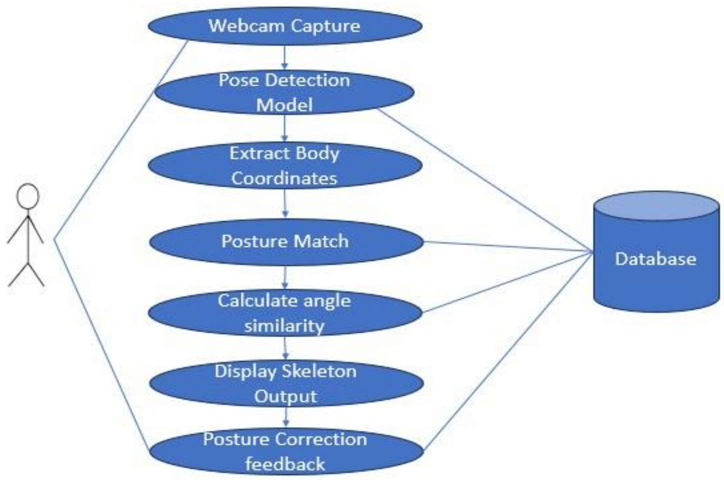


Figure 5.1: Use case diagram for human pose detection and correction

The use case for human pose correction involves the user interacting with a system designed to correct poses in images. The user provides an input image containing a person in a specific position. The system utilizes a pose detection model to identify key points on the person's body, such as joints. Using these key points, the system calculates the angles between body parts to assess the pose.

Based on predefined criteria, such as angles being outside of a certain range, the system determines if the pose needs correction. If correction is necessary, a correction algorithm is employed to calculate the adjustments needed to correct the pose. These adjustments are then applied to the detected key points, physically adjusting them to correct the pose.

Finally, the corrected pose is displayed to the user in an output image. The user can then view the corrected pose and assess its accuracy. If the pose is not corrected satisfactorily, the user may choose to adjust the input image or settings and repeat the process.

VI. CONCLUSION

Our system will make it easier to do exercises without the need for a special trainer. Reduce injuries due to improper technique. By implementing the system into a website, and selecting the required exercise in the webpage thereby improving the accuracy of detecting the pose with much more precision which in turn helps better angle calculation for comparison. Feedback for correction will be in the form of audio through chatbot.

REFERENCES

[1] Santosh Kumar Yadav, Amitojdeep Singh, Abhishek Gupta and Jagdish Lal Raheja, ``On Real-time Yoga recognition usingdeep learning," 20 May 2019 .
[2] S. Kan, O Javed, Z. Raheed, M. Shah, ``Human tracking in multiple cameras," 7 July 2001 .
[3] Chen Qian, Xiao Sun, Yichen Wei, Xiaoou Tang, Jian Sun, ``Realtime and Robust Hand Tracking from Depth,"June 2014 .
[4] C. Sminchisesu; B.Triggs, ``Kinematic jump processes for monocular 3D human tracking," 18 June 2003.
[5] Shruthi Kothari, ``Yoga Pose Classification Using Deep Learning" 20 May 2020 .
[6] Zhe Cao, Gines Hidalgo, Tomas Simon, Shih-En Wei, Yaser Sheikh, ``OpenPose: Realtime Multi-Person 2D Pose Estimationusing Part Affinity Fields," May 2019 .
[7] BlazePose, ``On-device, Real-time Body Pose Tracking with MediaPipe," 20 May 2019 .
[8] Zhichen Zhao, Huimin Ma, Shaodi You, ``Single Image Action Recognition Using Semantic Body Part Actions," December2016 .

[9] Imad Rida; Xudong Jiang; Gian Luca Marcialis, ``Human Body Part Selection by Group Lasso of Motion for Model-FreeGait Recognition," January 2016 .