Body Posture Detection and Motion Tracking using AI for Medical Exercises and Recommendation System

Anuj Patil[1 a)], Darshan Rao[1 b)], Kaustubh Utturwar[1 c)], Tejas Shelke[1d)] and Ekta Sarda[1e)]

[1] Department of Computer Engineering, Ramrao Adik Institute of Technology, D Y Patil Deemed to be University, Navi Mumbai

Abstract. Exercises are highly essential in our everyday lives, especially when patients are in the middle of a healing process and need to speed up their body's recuperation. Exercise has become more important in our lives as a result of this. They provide the cornerstone for improving human capacities and extending their lives. Artificial Intelligence and Image Processing can be utilized to improve and supplement the workout process without the need for professional supervision. A software-based motion tracker can keep track of all the exercises you've done and provide you feedback on your posture while you're working out. Through computing data and analysis, the exercise's beneficial efficiency will be increased. The MediaPipe framework could be utilized for this application; in this machine learning model, points are plotted at several joints of the human body posture, and movement is tracked, stored, and analyzed. This detailed analysis of the body tracking could be used in the implementation of an application that could keep a track of the medical exercise of a registered individual. The software could be further improvised in such a manner that the registered user could be mapped to an authentic verified doctor having the access to the diagnosis reports and exercise history of the mapped patient using databases.

Keywords: MediaPipe, BlazePose, BlazeFace, Bicep curls.

1 Introduction

Weighted exercises are one of the most useful and widely used types of physical activity in the world of healthy exercise. Some disorders can be cured by performing workouts that target specific muscle groups. Free weight workouts, on the other hand, can be frightening to beginning users and can have serious bodily consequences if done incorrectly or without professional assistance. Despite these advantages and high-risk activities, little to no effort has been made to use technology to assist new learners in learning new exercises. Our objective is to develop an easy-to-use interface technology that assists users who need to perform workouts to speed up their recovery or healing without the need for active professional supervision [11]. Counting repetitions, tracking your form, offering information on users' sessions, and recommending new and better exercises under passive supervision of health professionals are just a few of the features. When exercising with equipment or weights, this can greatly reduce the risk of injury [8]. Simple biceps curls are an example of this, as your back posture must be perfect to avoid straining your shoulder muscles and spine. [10] For such problems, we intended to create unobtrusive technology that could track users without requiring them to wear anything or place anything on their bodies, and that could subsequently be analyzed by health specialists for future treatment [9].

Push-ups are one of the most frequent bodyweight workouts. This is because it has several health advantages and requires the usage of muscles across the upper body. In terms of how this exercise is carried out, there is a lot of variation.[14] As a result, it was selected as a proof of concept for recording bodyweight exercises and repetitions performed.

The goals of this system are to provide easy access to the software, which can work on any system with x86 or arm-based CPU-s in a for all intents and purposes big way and

[a] Electronic Mail: anujpatil200@gmail.com
[b] Electronic Mail: darshan.rao120501@gmail.com
[c] Electronic Mail: k.utturwar@gmail.com
[d] Electronic Mail: tejasshelke040402@gmail.com
[e] Electronic Mail: ekta.sarda@rait.ac.in

to easily provide a hands-free interface to increase its practicality and usage, which really is quite significant.[13] Preparation of the entire dataset or data hub covering most of the exercises, depending on the diseases or other types of critical health conditions. Apart from this, keeping track of various points on the body to derive an analysis, track and calculate the human body posture percentage could be considered one of the significant objectives.
Lastly, providing tips to improve the body posture for optimal exercise benefits along with recording and preserving the exercise history to provide chronological analytics on your daily/weekly/monthly exercise habits [12].

2 Literature Survey

There is currently no instance of improving health condition-based exercises. Personal trainers and Physiotherapists are the ones that monitor, manually. This has disadvantages for the two types of people. Firstly, the people who live in remote places do not have access to these personal trainers and hospitals. The second type of people is the senior citizens, who need continuous attention for their daily exercise and diet plans.[5] Even if they have their treatment in the hospitals, there is no one to check if their exercise form is correct. [7] There is a need for a person to be allotted for these activities, which increases the additional fees. Sensors such as Xbox Kinect had come up which were able to track body gestures [2]. These sensors were used for games and entertainment and did not enter the field of health. The main reason for this is the expensive build due to which it was not affordable for ordinary citizens. There is a dearth in the exercise to be assigned for the specific types of diseases. Even though some software for detecting gestures, the proper integration with the disease-specific exercise is still in the development stage.[4]

Projects like a real-time on-device hand tracking solution for AR/VR applications that predicts a human hand skeleton from a single RGB camera.[1] These were mainly used for mobile devices and used for hand tracking.[8] Research on metabolic health and nutrition are made which have given a brief idea of biomarkers that can be used the making nutrition plans.[3]

3 Proposed Methodology

The hand tracking method makes use of an ML pipeline comprised of two models that collaborate.:

1. A palm detector that locates palms using an aligned hand bounding box and operates on a whole input picture.
2. A hand landmark model that works with the palm detector's chopped hand bounding box to yield high-fidelity 2.5D landmarks.

Firebase was used as the database framework to store the information about the exercise and easily fetch them as needed. Django framework would handle the front end of the application.

The general architecture of the proposed system is shown in **Fig.1** which shows the generalized flow that includes capturing the live video feed from the patient's camera and displaying it on the dashboard. Further, the exercise tracker tracks the workout performed and pushes the results in the database which are further passed to the doctor's dashboard which is also used to input the instructions for the next exercise set and diet chart and send it back to the patient’s dashboard via the database.

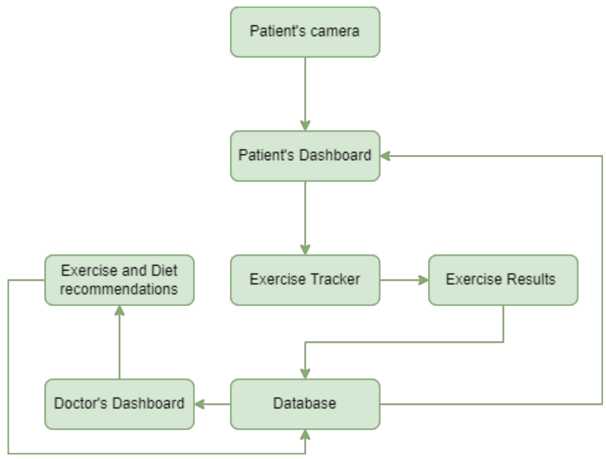

Fig. 1. The generalized flow of the system

There are two kinds of detection models in the MediaPipe framework - BlazePose detector and BlazeFace detector. In BlazePose, there are two machine learning models: a Detector and an Estimator. The Detector removes the human region from the input image, while the Estimator returns key points from a 256×256 resolution image of the detected individual. Similar to BlazeFace, which is also available in MediaPipe, this single shot detector type is suited for mobile real-time applications. Hand gesture detection is a challenging task. The model must detect both occluded and self-occluded hands while functioning with a wide range of hand sizes and a large scale span (20x). The contrast pattern on the face is high around the eyes and mouth, whereas hands do not, making it more difficult to distinguish them just based on their visual characteristics. The system employs three separate validation datasets that span different sectors to compare the quality of the models to other well-performing publicly accessible solutions: Yoga, Dance, and HIIT. There is only one person in each shot, who is 2-4 meters away from the camera. The model simply evaluates 33

key points from the COCO topology to be compatible with previous solutions.
Fig.2 is an example of how the BlazeFace detector detects the human body posture and can perform segmentation masking over it with very high precision and accuracy.

Fig. 2. Example of MediaPipe Pose real-world segmentation mask. [1]

In **Fig.3**, the algorithmic flow of the MediaPipe framework for hand movement detection is as follows: image snapshots from the given camera are pipelined into the BlazeFace, which then recognizes body points using large datasets. The image is then mapped, cropped, and projected on the screen.

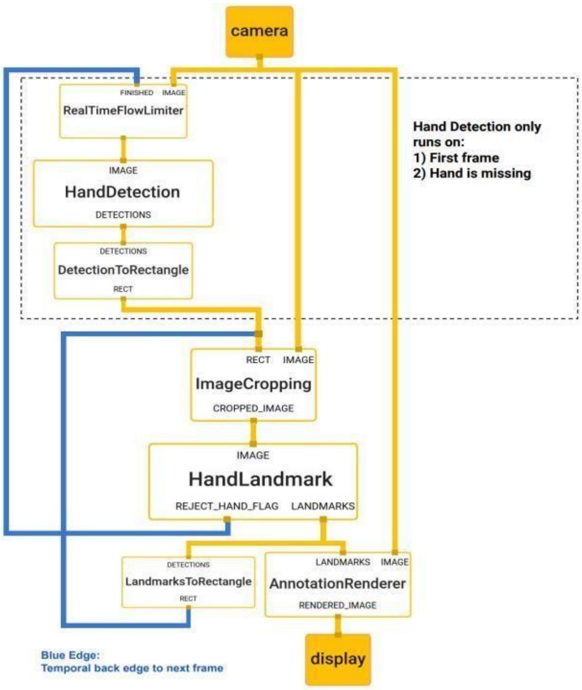

Fig. 3. MediaPipe pipeline flow. [1]

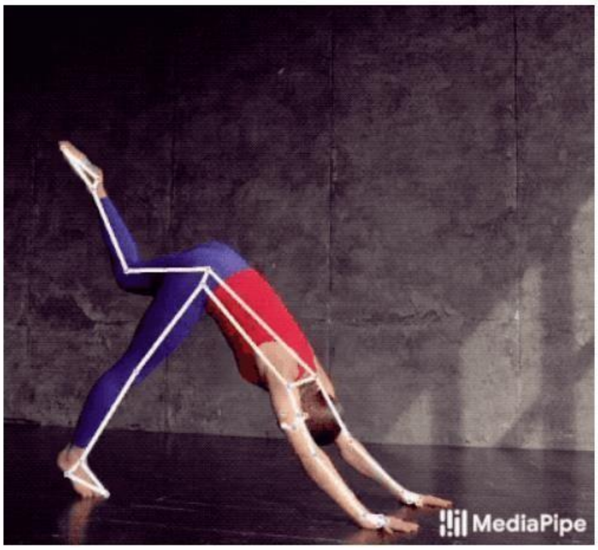

Fig. 4. Example of MediaPipe Poses for pose tracking.[7]

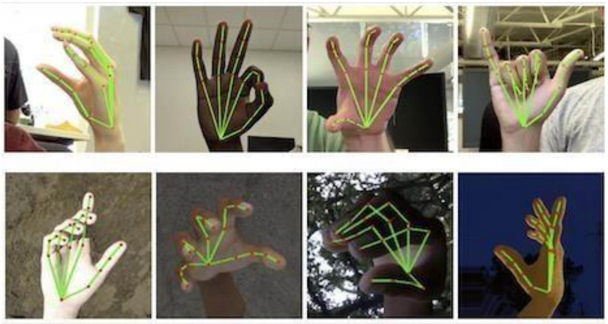

Fig. 5. Hand gesture tracking outputs [7]

Fig.4 and **Fig.5** shows how the MediaPipe model plots the joints and lines on the human body posture and palm for movement detection using the angle between two joints.[6]

4 Implementation and Results.

The Health Tracker is integrated to a web application using Django.
Django is a high-level Python web framework that promotes rapid development and simple, practical design. It's built by professional developers to take care of a lot of the problems of web development, so the user can concentrate on developing your application instead of reinventing the wheel.
The mapping between these two portals would be done using Google's firebase. The patient's portal would consist of a navigation bar along with the profile details. Further, regarding the main components of the portal, the dashboard would consist of a window that is integrated with the system's camera and shows the live feed of the

user's environment. Beside the window, allotted daily tasks would be displayed, which would consist of the several exercises allotted by the doctor who is mapped with that particular patient in a list format. The patient needs to follow the daily workout schedule referring to the list and his/her exercise count would be tracked using the machine learning model which is being implemented, and the count would be displayed using a counter container on the corner of the screen for the patient to refer to. Once the allotted instructions are completed by the user, he/she would push the results in the database using the front-end component provided. Apart from the daily exercise recommendations, a section with a detailed description of the daily diet provided by the respective doctor would be also displayed dynamically on the webpage.

The doctor's portal would again consist of the navigation tabs, similar to the patient's portal with the respective profile information which would consist of professional-medical biodata of the doctor. Further, the doctor's dashboard would be mapped with all the patients which are registered to him/her using the database framework, that is, on the webpage, there would be a section where the doctor could view and surf through all the patients mapped with him/her, on the selection of any of the patients, a detailed dashboard consisting of the patient's entire medical journey with that doctor would be displayed in a proper format. The detailed dashboard would consist of several tabular, diagrammatic, and graphical forms of data representations based on the patient's chronological diagnosis to date. On this web page, doctors can dynamically push the daily exercise and diet recommendations in the system's database, which is further pushed on the patient's dashboard.

The Exercise Tracker is the main component of the system. The video captured is further processed in MediaPipe. The MediaPipe uses the BlazePose Detector model for detecting the person. After detecting the person, the landmark model in MediaPipe Pose uses prediction methods to detect the location of 33 pose landmarks as shown in **Fig.6**. The exercise used in this application requires only hand and leg landmarks.

All landmarks are joined, and a function is used to calculate the angle between the joints. The difference in the joint angles caused due to movements is used for calculating the reps of the exercises.

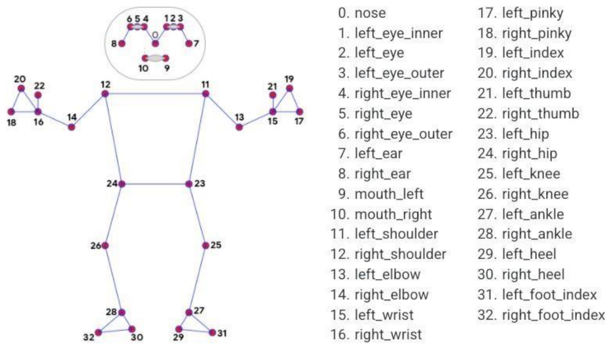

Fig. 6. Landmark identification on the body[7]

For test cases, exercises such as biceps curls and push-ups are used. The biceps curls use angles between the landmarks {16, 14, 12}(right hand) and {11,13,15}(left hand) and angles a difference between timestamps is used for counting reps. Two instances of bicep curls detection and counting are shown in **Fig.7** and **Fig.8**.

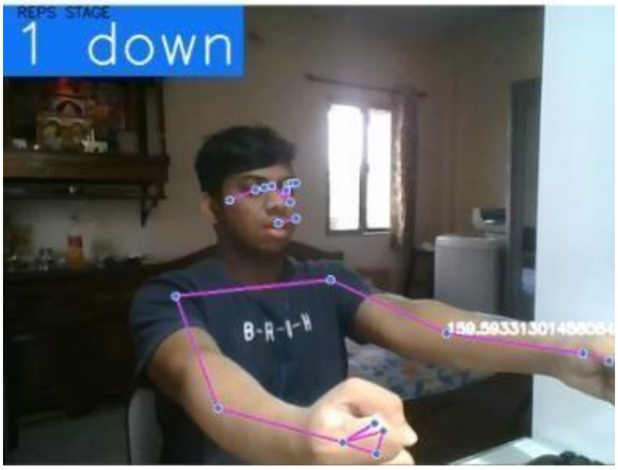

Fig. 7. Image showing bicep curls (Instance 1)

Fig. 8. Image showing bicep curls (Instance 2)

Fig. 9. Prototype User Interface showing repetition and body posture.

Both of the portals, that is the patient's portal and the doctor’s portal which are supported by the backend source code needs to be properly mapped together and backed with the collected and stored data in the data warehouse for the application to run and function according to the expected requirements. These requirements could be efficiently and optimally fulfilled by Firebase’s Firestore. Google Firebase is an application developed by Google for a development platform that allows creators to create applications for iOS, Android, and the web. Firebase delivers analytics tracking, reporting, and app issue fixes, as well as marketing and product experimentation capabilities.

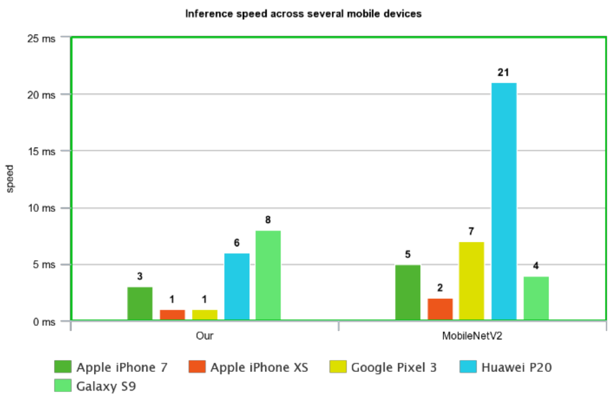

Fig. 10. Graph showing the inference speed across various mobile devices with different operating systems and hardware.

To identify the face or the body, a frontal camera or webcam is used. The model allows the user to keep track of his posture and count the overall number of reps he has completed. Only faces that occupied more than 20% of the image area were considered for the frontal camera model, according to the use case. The 5% criterion was set for the rear-facing camera model. The model was generated after it was trained on a dataset of 66k photos. The developed system has used a 2k image collection from a private, geographically diversified dataset. The mapping of the movements done is of a very high precision because of the usage of the dataset and the robust model, which in comparison to other models on the market is quite accurate **[Fig.9]**. The lighting in the space in which the activities are performed is really important. The tracker was unable to detect the joins while testing in lower illumination settings, resulting in some incorrect findings. The camera's resolution made a difference in the tracking of the workouts; the higher the camera's resolution combined with enough lighting, the closer the precision obtained under such ideal conditions is to perfection.

Model	**Average Precision**	**Inference Time, ms (iPhone XS)**
MobileNetV2-SSD	97.95%	2.1
Ours	**98.61%**	**0.6**

Table 1. Frontal camera face detection performance

The outcomes of certain test cases are shown in **Table.1** above. The tracking which was implemented at first was tested with the existing model MobileNetV2-SSD, but the results were a millimeter off. The model which was developed later on produced better results than the previous model when used. The iPhone XS was the device utilized to test our first case, and the test results were better than the previous ones.

Continuing the testing on the iPhone XS, the testing was conducted on a variety of devices, ranging from phones with lower processing power, such as the iPhone 7, to mobile devices with better computing speed and graphical processing, such as the iPhone XS(iOS) and Pixel 3. (Android).

Accuracy	Exercises
92%	Biceps Curl
84%	Squats
82%	Push-Ups

Table 2. Accuracy of Exercises

The testing of the app for each exercise was done for 50 reps, hence the percentage accuracy of each exercise is the percentage of reps counted per 50 reps. 10 test cases were made for push-up exercise **[Fig.9]** which gave an average accuracy of 92% **[Table.2]**.

5 Conclusion

The system introduced a free weight exercise tracking software that allows users to learn and correct their form in this research. The squat exercise was chosen for this study because it has extremely specific needs that could be tracked and modified easily. Form specifications were used to transmit these requirements, and an indicator of a red X and a green check was used to demonstrate whether they were proper or incorrect. As the user performed the push-up, these parameters, and indicators allowed them to see and adjust their form. In our pilot study, participants were able to enhance their push-ups and correct their form to the point where they felt comfortable performing the exercise.

6 Future Work

The system has a lot of areas for improvement. Further research will be required on the types of exercise used by patients with specific types of diseases. The research requires consulting a physician or Physiotherapists who can verify the exercise is proper or not. Some exercises such as push-ups had lesser accuracy, mainly due to the camera quality used for the application. The less accuracy can be improved by making an android app that can use its higher quality camera for detection. As the system created is only limited to diet plans and health tracking, there is a lot of space for adding more features such as blood tests, vaccine reports. The system can be more versatile by adding vernacular languages, as for countries like India, which has different languages in every state. The application can also be made women-specific that can keep track of women's health, such as checking periods and suggesting women-specific exercise. The UI needs to make it more user-friendly as the targeted patients are thesenior citizens and the people living in rural areas who areobtuse to technology.

References

[1] Zhang, F., Bazarevsky, V., Vakunov, A., Tkachenka, A., Sung, G., Chang, C. L., & Grundmann, M. (2020). Mediapipe hands: On-device real-time hand tracking. arXiv preprint arXiv:2006.10214.

[2] Conner, C., & Poor, G. M. (2016, May). Correcting exercise form using body tracking. In *Proceedings of the 2016 CHI Conference Extended Abstracts on Human Factors in Computing Systems* (pp. 3028-3034).

[3] Lee, E. C., Fragala, M. S., Kavouras, S. A., Queen, R. M., Pryor, J. L., & Casa, D. J. (2017). Biomarkers in sports and exercise: tracking health, performance, and recovery in athletes. *Journal of strength and conditioningresearch*, *31*(10), 2920.

[4] Van Der Zee, M. D., Van Der Mee, D., Bartels, M., & De Geus, E. J. (2019). Tracking of voluntary exercise behaviour over the lifespan. *International Journal of Behavioral Nutrition and Physical Activity*, *16*(1), 1-11.

[5] Shpigelman, L., Lalazar, H., & Vaadia, E. (2008). Kernel-ARMA for hand tracking and brain-machine interfacing during 3D motor control. *Advances in neural information processing systems*, *21*, 1489-1496.

[6] Ahmad, A., Migniot, C., & Dipanda, A. (2019). Hand pose estimation and tracking in real and virtual interaction: A review. *Image and Vision Computing*, *89*, 35-49.

[7] Rahman, S. U., Afroze, Z., & Tareq, M. (2014). Hand Gesture Recognition Techniques For Human Computer Interaction Using OpenCv. *International Journal of Scientific and research publications*, *4*(12).

[8] Pedersen, B. K., & Saltin, B. (2015). Exercise as medicine–evidence for prescribing exercise as therapy in 26 different chronic diseases. *Scandinavian journal of medicine & science in sports*, *25*, 1-72.

[9] McGloin, R., Embacher, K., & Atkin, D. (2017). Health and exercise-related predictors of distance- tracking app usage. *Health Behavior and Policy Review, 4*(4), 306-317.

[10] Voth, E. C., Oelke, N. D., & Jung, M. E. (2016). A theory-based exercise app to enhance exercise adherence: a pilot study. *JMIR mHealth and uHealth, 4*(2), e4997.

[11] Mohamed Abdur Rahman, Ahmad M. Qamar, Mo- hamed A. Ahmed, M. Ataur Rahman, and Saleh Basalamah. 2013.Multimedia Interactive Therapy Environment for Children Having Physical Disabilities. *In Proceedings of the 3rd ACM Conference on Inter- national Conference on Multimedia Retrieval (ICMR'13). ACM, New York, NY, USA,* 313–314.

[12] Derrick Cheng, Pei-Yu Chi, Taeil Kwak, Björn Hart- mann, and Paul Wright. 2013. *Body-tracking Camera Control for Demonstration Videos. In CHI '13 Extended Abstracts on Human Factors in Computing Systems (CHI EA '13). ACM, New York, NY, USA,* 1185– 1190. *DOI:http://dx.doi.org/*10.1145/2468356.2468568

[13] Han Ding, Longfei Shangguan, Zheng Yang, Jinsong Han, Zimu Zhou, Panlong Yang, Wei Xi, and Jizhong Zhao. 2015. *FEMO: A Platform for Free-weight Ex- ercise Monitoring with RFIDs. In Proceedings of the13th ACM Conference on Embedded Networked Sensor Systems (SenSys '15). ACM, New York, NY, USA,* 141–154.

[14] Kyle Rector, Cynthia L. Bennett, and Julie A. Kientz. 2013. Eyes-free Yoga: An Exergame *Using Depth Cameras for Blind & Low Vision Exercise. In Pro- ceedings of the 15th International ACM SIGACCESS Conference on Computers and Accessibility(ASSETS'13). ACM, New York, NY, USA, Article* 12, 8 *pages.*