Real-Time Exercise Analysis and Corrective Feedback System

Prof. Samadhan Jadhav
PICT Assistant Professor
Pune, India
swjadhav@pict.edu

Neeraj Ghatage
PICT
Pune, India
neerajghatage@gmail.com

Atharva Joshi
PICT
Pune, India
atharvaavjoshi11@gmail.com

Atharva Jadhav
PICT
Pune, India
atharvaj898@gmail.com

Dhanashree Kamble
PICT
Pune, India
kambledhanashree04@gmail.com

ABSTRACT

Fitness activities offer numerous health benefits, but improper execution can lead to inefficiency and potential harm. When performing exercises incorrectly, individuals often fail to maintain the correct form or posture. In this research, a program has been developed to assess and provide feedback on the user's workout posture. Utilizing a camera, this system offers real-time interaction by detecting joints and evaluating workout accuracy through vector angles. The implementation leverages MediaPipe, a cross- platform ML framework compatible with Windows and Linux computers equipped with a webcam.

1. INTRODUCTION

Engaging in exercises is a commendable way to enhance one's fitness and overall health. However, the improper execution of exercises can pose significant risks, particularly when dealing with heavy weights that have the potential to cause severe muscle or ligament injuries. Many individuals enthusiastically participate in workout routines but often struggle to maintain the correct technique or posture. Such difficulties may stem from a lack of formal training, be it from fitness classes or a personal trainer, as well as the consequences of muscle soreness or attempting to lift weights beyond their capacity. Our goal is to offer a solution that helps individuals adopt proper exercise postures, thereby preventing injuries and enhancing the effectiveness of their workouts, using nothing more than a computer and a webcam.

In this research, we delve into the utilization of skeleton- based representations, which offer the advantages of being low-dimensional, interpretable, person-independent, and privacy- preserving. These skeletal representations allow us to concentrate on the essence of motion while making generalizations about an individual's appearance and background. We present a real-time body tracking pipeline that predicts the hand skeleton and the overall body concept, leveraging the capabilities of MediaPipe, a versatile framework for developing cross-platform machine learning solutions. Our approach involves posture estimation techniques, where the accuracy of exercise execution is assessed by fine-tuning the range of angles between relevant joints.

By combining these innovative techniques with the power of technology, we aim to revolutionize the way individuals approach exercise, making it safer, more efficient, and accessible to a broader audience. This system not only enhances the quality of workouts but also contributes to injury prevention and overall well-being.

2. LITERATURE SURVEY

[1] In this paper, the writer presents a body posture smart recom- mendation system, which detects user's posture and guides them according to the selected back exercise using a gyroscope sensory module embedded in the smart fitness suite. They proposed the system for two exercises, T-bar and bicep concentrated dumbbell curl. Along with this, a bicep curl muscle health detection feature is added to the proposed system, which detects muscle health in real-time. EMG sensor is used to stop the user from exercising in the extreme fatigue stage to prevent muscle injury. KNN model is used for the forward feature selection technique with 89% of accuracy. Subsequently, a user-guided recommendation feature is added which is based on the trained dataset over the android appli- cation using the text-to-speech feature in real-time. Future scope: In the future, the smart fitness suite could be trained for other body workouts as well. To increase the accuracy of the proposed system, they can consider the gyroscope drift issue to stabilize the signals for better classification of exercises. Finally, the proposed smart fitness suite can be made specifically for male or female users by collecting datasets separately.

[2] The detection of the 2D poses of many people in an image is done in this research using an efficient method. The method learns to associate body parts with people in the image using non- parametric representations known as Part Affinity Fields (PAFs). The technique uses the full image as the input for a two-branch CNN to jointly predict part affinity fields for parts association and confidence maps for body part detection. To associate body part candidates, the parsing stage executes a series of bipartite match- ings. Lastly, combine them into full-body positions for every per- son in the picture. No matter how many people are in the im- age, the architecture's global context encoding enables a greedy bottom-up parsing step that retains high accuracy. The architec- ture is designed to jointly learn part locations and their association via two branches of the same sequential prediction process. The method placed first in the inaugural COCO 2016 keypoints chal- lenge and significantly exceeds the previous state-of-the-art result on the MPII MultiPerson benchmark, both in performance and ef- ficiency. There is more improvement space in capturing spatial dependencies than in recognizing body parts appearances.

[3] For the pose estimation component, they utilize a pre- trained real-time system, called OpenPose, that can detect human body key-points in videos. They evaluate their posture identifier in different ways depending on the algorithm: for heuristic algo- rithms, they feed in all videos for evaluation, while for machine learning algorithms, they evaluate by splitting their video dataset into train and test sets, and report results on the test set. Pose Trainer application from a technical perspective as a pipeline sys- tem, consisting of multiple system stages as follows • Record and crop video • Pose estimation (OpenPose) • Normalize, clean and save key-points • Exercise perspective detection • Evaluate exer- cise using geometry or ML • Provide specific feedback on exercise The results of Pose Trainer on four different dumbbell (free mo- tion) exercises: bicep curl, front raise, shoulder shrug, and stand- ing shoulder press. For each exercise, they take both a geomet- ric/heuristic approach, as well as a machine learning approach us- ing dynamic time warping. They identified several extensions as strong opportunities for future work past this course project. One path would be to export Pose Trainer to smartphones, building an application that allows users to record a video and get pose feed- back 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. Also, they want to work on improved graphics, for instance, showing the user their labelled pose dia- gram, and comparing to the labelled pose diagram of a ground truth trainer.

[4] This Paper presents a groundbreaking application of Deep Neural Networks (DNNs) for human pose estimation, showcasing the advantages of DNN-based regression to joint coordinates and the use of a cascade of regressors. The approach achieves state-of-art or superior results on challenging academic datasets, indicating its potential for real-world applications. Additionally, the adaptation of generic convolutional neural networks for localization tasks demonstrates the versatility of neural networks across different domains. Future research will focus on the development of specialized

architectures for localization problems, particularly in the context of pose estimation.

[5] ERICA employs in-ear devices and inertial sensors to provide real-time feedback on exercise mistakes, achieving over 94% detection accuracy for common errors like improper range of motion. The system's unobtrusive design and multi-user support facilitate its use in shared gym environments, though it currently focuses on a limited range of exercises.

[6] The Recognition of Yoga Asana from Real-Time Videos using Blaze-Pose proposes a lightweight architecture that operates efficiently on low-end devices while achieving high accuracy (98.65% on test data). However, the limited training dataset and lack of feedback mechanisms for posture correction may affect its usability.

[7] The Infinity Yoga Tutor utilizes OpenPose and Mask R-CNN to provide real-time posture detection and correction through a mobile application. With a high pose detection accuracy of 99.91%, the system enhances user experience and safety, although its effectiveness may vary due to lighting conditions and it currently supports only six yoga poses.

[8] This paper concludes by summarizing the key contributions of the proposed approach, particularly the incorporation of discriminative part template predictors within a pictorial structure framework for robust human pose estimation. The demonstrated superiority over independent part templates and state-of-the-art methods using tree structures underscores the potential of this methodology in real-world applications.

3. METHODOLOGY

3.1. MediaPipe and Python

This model includes a whole-body and hand skeleton prediction pipeline for real-time on-device body tracking. MediaPipe, a framework for creating cross-platform ML solutions, is used to implement it.

With the help of BlazePose research, which also drives the ML Kit Pose Detection API, MediaPipe Pose is an ML solution for high-fidelity body pose tracking. It infers 33 3D landmarks and background segmentation masks on the entire body from RGB video frames. Modern methods mostly rely on robust desktop environments for inference. On the other hand, this approach achieves real-time performance on the majority of contemporary mobile devices, desktop/laptop computers, Python, and even the web. The following figure shows the 33 3D landmarks.

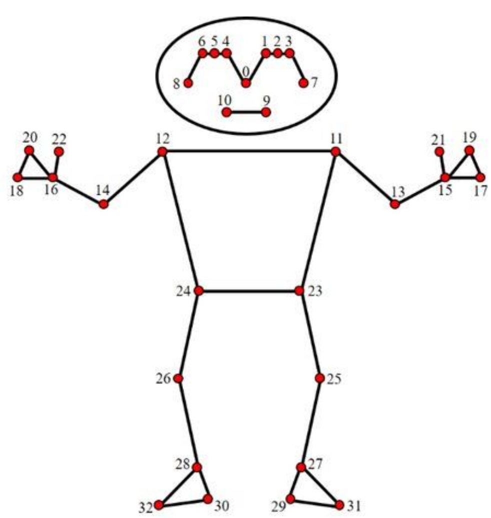

Figure 1: Pose Landmarks

It locates joints and assigns coordinates. It displays the connecting lines between the nearby coordinates. Calculating the necessary anglesbetween these vectors allows for exercise visualization and error verification.

3.2. Pose Estimation using Openpose

For the pose estimation component, a pre-trained real-time system, called OpenPose is utilized. which can detect the human body's key points in videos. The second part of the application involves detecting the quality of a user's predicted pose for a given exercise. heuristic-based and machine-learning models approach is used in this. A full application consists of the previously de- scribed two main components, combined into an end-to-end appli- cation that can take a video of an exercise and provide useful ex- ercise form feedback to the user. Pose training starts with the user recording a video of an exercise, and ends with the Pose Trainer application providing specific feedback on the exercise form to theuser.

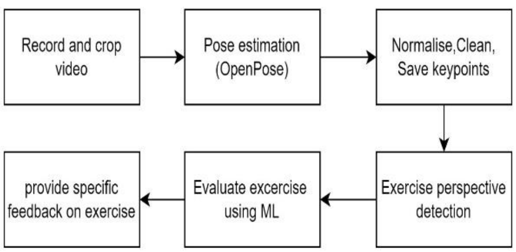

Figure 2: Block Diagram Pose Estimation and Correction

For pose estimation, deep convolutional neural networks (CNNs) to label RGB images is used. After experimentation with multiple state-of-the-art pose estimators, we choose to use the pre-trained model, OpenPose, for pose detection.

OpenPose output consists of lists containing the coordinate predictions of all key point locations, and their corresponding pre- diction confidence. It considers the predictions of 18 key points of the pose, which include the nose, neck, shoulders, elbows, wrists, hips, knees, and ankles.

Next, body vectors from key points of interest are computed, and use personal training guidelines and our own recorded videos to design geometric heuristics, evaluating the body vectors.

For instance, in a bicep curl, Two heuristics of interest are identified. First, the upper arm should be kept steady and not move significantly. this is done by calculating 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.

In the geometric algorithm, the range in angle between the upper arm and torso (measuring if the user rotates the shoulder when lifting), and the minimum angle between the upper arm and forearm (measuring how high the user lifts) are considered. If the range between the upper arm and torso angle is above 35 degrees, this signals as too much shoulder rotation. If the minimum angle between the upper arm and forearm is above 70 degrees, this is flagged as not curling the weight all the way up.

3.3. Posture recognition using 2D images and Convolutional Neural Network

The suggested approach is based on modifying a pre-trained CNN for our particular recognition challenge utilising 2D photos of human poses. Photos are classified into 1000 different object categories using the AlexNet network, which has been pre-trained using the ImageNet database. Our dataset is first pre-processed by scaling the images in accordance with the CNN input format and removing the background to extract the body silhouette before utilising posture images to retrain the network. Initial silhouette segmentation is carried out using depth image body parsing. We use the final images from the pre-processing step to retrain the CNN for posture classification.

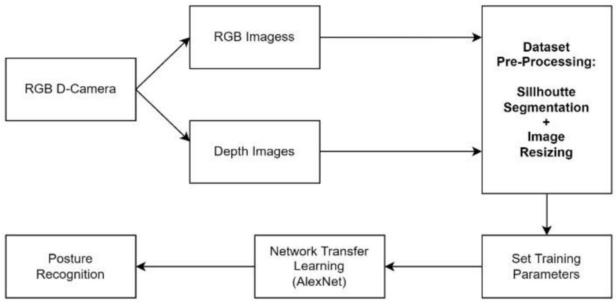

Figure 3: Posture recognition using 2D images and Convolutional Neural Network

3.4. Posture recognition using 3D joint-based features

Firstly, a 3D skeleton model is built using the skeleton detector of Shotton et al. for localizing different body parts and accurately modeling an articulated structure of connected segments. The ob- tained 3D skeleton formed by 25 joints represents the human pos- ture as illustrated in the figure.[9] Once joint positions are extracted, 2 types of features representing human body posture are computed

: the 3D pairwise distances between joints, and the geometrical angles of adjacent segments.

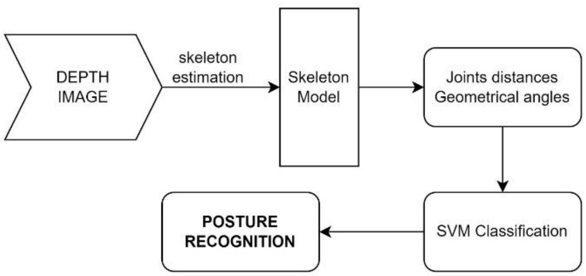

Figure 4: Posture recognition using 3D joint based features.

Geometrical angles defined by adjacent segments are directly estimated from joint positions in the 3D space. Once the body pose is modeled, posture recognition is performed using SVM classification.

4. WORKING

We followed method 3.1 for pose detection and correction i.e

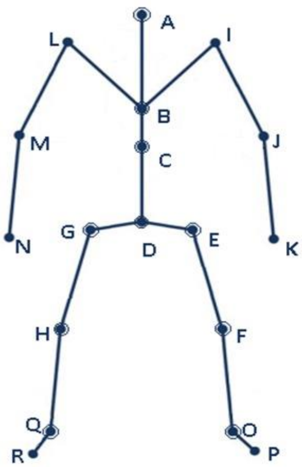

Figure 5: 2d skeleton

mediapipe and python. Mediapipe's pose model is used for this. Initially, we run our code on Jupyter notebook. When the code is executed camera starts running. It detects the coordinates on our body (33 Landmarks). The coordinates of the joints involved in a certain exercise are taken. After determining the joints and their coordinates respective angles are calculated. Setting the restriction on the angle involved in the exercise, the correctness of the exer- cise is determined. If the exercise is performed correctly then the counter is incremented else it remains the same.

For example Bicep Curl: In the Bicep Curl exercise shoulder, elbow, and wrist joints are involved. So the angle between these vectors is calculated i.e the angle at the elbow. While starting the exercise angle is between 160 degrees to 180 degrees and the stage is down. As the user starts performing the exercise angle starts de- creasing, If the angle reaches below or equal to 30 degrees and the stage is down counter is incremented it shows the exercise is per-formed correctly. And if the angle of the elbow is not reaching its limits (i.e 30 degrees and 160 degrees) counter is not incremented it signals you are not performing in the right way.

5.1 RESULTS

Real Time Performance Analysis

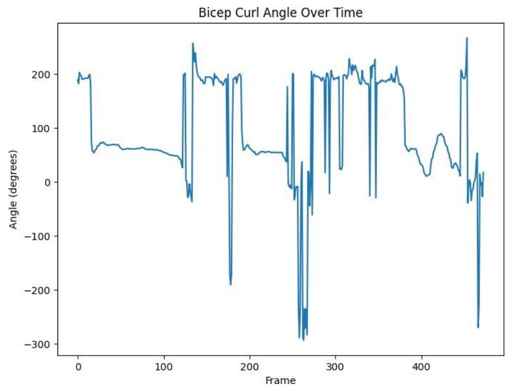

Figure 6: Bicep Curl Angle Over Time

Fig.6 chart is updated in real-time as new frames are processed, allowing us to observe how the angle changes as the personraises and lowers their arm during the bicep curl exercise.
It provides a visual representation of the dynamic changes in the angle of the arm during the exercise, which is used for tracking exercise performance and ensuring that the correct form is maintained. It allow to monitor their progress and makeadjustments as needed to perform the exercise accurately.

5.2 Angle Detection Accuracy

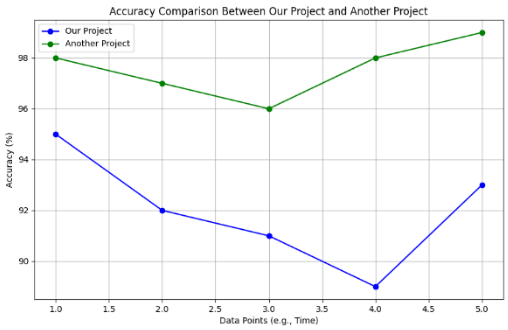

Figure 7: Evaluation of Bicep Curl Angle Detection Accuracy

Fig.7 enables a straightforward evaluation of the actual project's accuracy compared to the ideal benchmark, contributing valuable insights to the research.
It enables a straightforward evaluation of the actual project's accuracy compared to the ideal benchmark, contributing valuable insights.
By demonstrating the actual project's performance in comparison to the ideal benchmark, this contributes valuable findings to the broader field of human exercise correction and detection.

Real Time Feedback

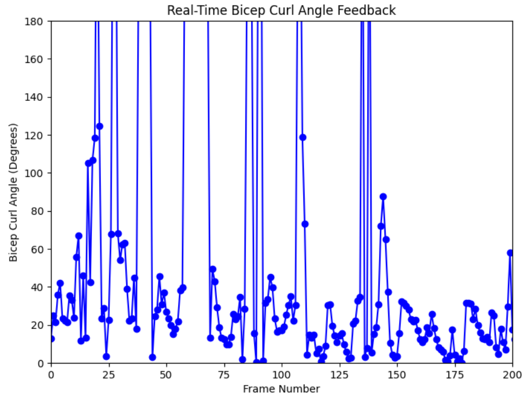

Figure 8: Real-Time Feedback Mechanism for Correction

Fig.8 feedback graph serves as a dynamic representation of the real-time feedback mechanism in the context of bicep curl angle detection.

It's an invaluable visual representation that demonstrates the real-time feedback mechanism's utility and its effectiveness in enhancing exercise quality and user experience. It is a vital component in the ongoing development of human exercise correction technologies.

5.3 Comparative Assessment of Project Performance

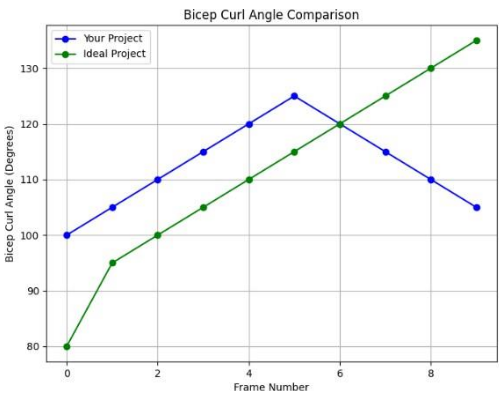

Figure 9: Project Effectiveness and
Performance Comparison.

In Fig.9, the comparison graph plays a pivotal role in the evaluation and benchmarking of your project, specifically in the context of bicep curl angle detection and correction. This visual representation allows for an in-depth examination of how this project performs compared to another project, often considered an ideal or reference model.

The primary purpose of this comparison graph is to objectively assess the performance of two distinct projects: y project (represented by the blue line) and another project (represented by the green line). This comparison emphasizes a data-driven evaluation, enabling clear and unbiased insights into their relative performance.

5. CONCLUSION

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 pose estimation is used to calculate human body key points during live interaction. Machine learning algorithm is used for deciding posture correctness and geometric algorithms for providing feed- back on exercise based on increment counter. One exercise is con-sidered which can be extended to many other exercises for future work. For increasing accuracy more than one angles can be con- sidered. This model can be extended to accommodate two or more people working out together.

6. REFERENCES

[1] A. Hannan, M. Z. Shafiq, F. Hussain, and I. M. Pires, *A Portable Smart Fitness Suite for Real-Time Exercise Monitor-ing and Posture Correction*. MDPI, 2021.

[2] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, *Realtime Multi- Person 2D Pose Estimation using Part Affinity Fields*, Cornell University ARXIV, 2017.

[3] S. Chen and R. R. Yang, *Pose Trainer: Correcting Exercise Posture using Pose Estimation*, Cornell University ARXIV, 2020.

[4] A. Toshev and C. Szegedy. Deeppose: Human pose estimation via deep neural networks. In CVPR, 2014.

[5] Meera Radhakrishnan, Darshana Rathnayake, Ong Koon Han, Inseok Hwang, Archan Misra1, *ERICA: Enabling Real-time Mistake Detection & Corrective Feedback for Free-Weights Exercises.* Digital Library ACM ,2020

[6] Sharfuddin Waseem Mohammed, Vignesh Garrapally, Suraj Manchala, Soora Narasimha Reddy and Santosh Kumar Naligenti, *Recognition of Yoga Asana from Real-Time Videos using Blaze-pose* , IJCDS, 2022

[7] M. Dantone, J. Gall, C. Leistner, and L. Van Gool. Human pose estimation using body parts dependent joint regressors. In CVPR, 2013.

[8] Bhat Dittakavi, Divyagna Bavikadi, Sai Vikas Desai, Soumi Chakraborty, *Pose Tutor: An Explainable System for Pose Correction in the Wild* . 2022 IEEE/*CVF.*

[9] *Federica Bogo, Angjoo Kanazawa, Christoph* Lassner, Peter Gehler, Javier Romero, Michael J. Black Keep it SMPL: *Automatic Estimation of 3D Human Pose and Shape from a Single Image.* Cornell University ARXIV 2016

[10] Deeptha Girish, Vineeta Singh, and Anca Ralescu. Understanding action recognition in still images. In 2020 IEEE/CVF Conference on Computer Vision and Pattern Recognition Workshops (CVPRW), pages 1523–1529, 2020.

[11] Simo-Serra, E., Quattoni, A., Torras, C., Moreno-Noguer, F.: A joint model for 2D and 3D pose estimation from a single image. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR. pp. 3634–3641 (2013)

[12] M. Young, The Technical Writer's Handbook. Mill Valley, CA: University Science, 1989.

[13] Yang, Y., Ramanan, D.: Articulated pose estimation using flexible mixtures of parts. In: IEEE Conference on Computer Vision and Pattern Recognition, CVPR. pp. 3546–3553 (2011)