Handwritten Text Recognition

|  |  |  |  |
| --- | --- | --- | --- |
| Avishkar Sudharma Dalvi  TE - *A.P.Shah Institute*  *of Technology* Thane, India  20202002@apsit.edu.in | Riddhi Jagdish Narkar  TE - *A.P.Shah Institute*  *of Technology* Thane, India  19102003@apsit.edu.in | Soumyojyoti Jyotimoy  Dutta  TE - *A.P.Shah Institute*  *of Technology* Thane, India  19102014@apsit.edu.in | Vedang Nilesh Gore  TE - *A.P.Shah Institute*  *of Technology* Thane, India  19102065@apsit.edu.in |

*Abstract*—The project focuses mainly on how effectively the user performs his/her exercise which helps them minimize the casualties that are associated with an improper form of their exercise. Human pose estimation is one of the important researches in the field of Computer Vision for the last few years. In this project, pose estimation and deep machine learning techniques are combined to analyze the performance and report a feedback on the repetitions of exercises performed in real time. Involving machine learning technology in the fitness industry could help the judges to count repetitions of any exercise. The proposed method divides respectively into three phases; pose tracker to identify and track user, exercise recognition to detect the name of the appeared exercises, and counter to count and indicate the correct and incorrect repetitions.

*Index Terms*—AQA: Action Quality Assessment

# I. INTRODUCTION

Create an Android application which will use your phone’s camera to count the number of repetitions you do for a single exercise and check how correctly you do them.

## A. Problem Definition

Due to the lockdown in the pandemic, everyone had to stay at home and weren’t able to access the gym. This caused people to seek the help of various apps available in the app store to help them out. Like everyone else, we too wandered through the App store to find an app that would fulfill our needs. After researching a lot we found out that there were plenty of apps that would suggest the user a new workout plan and these workout plans would instruct the user on how to perform the suggested workout using images or a video.However no attempts are made to recognise and correct the form of the user while performing these exercises. There are also no systems available that accurately count the no of reps that the user has performed. Without any proper assistance and guidance from a personal trainer, people are at a risk of injuring themselves even while performing basic exercises like a push up or a bicep curl. With this information in mind we scoped various research papers that could perform either of these task and we finally deduced two main ideas to tackle this problem.

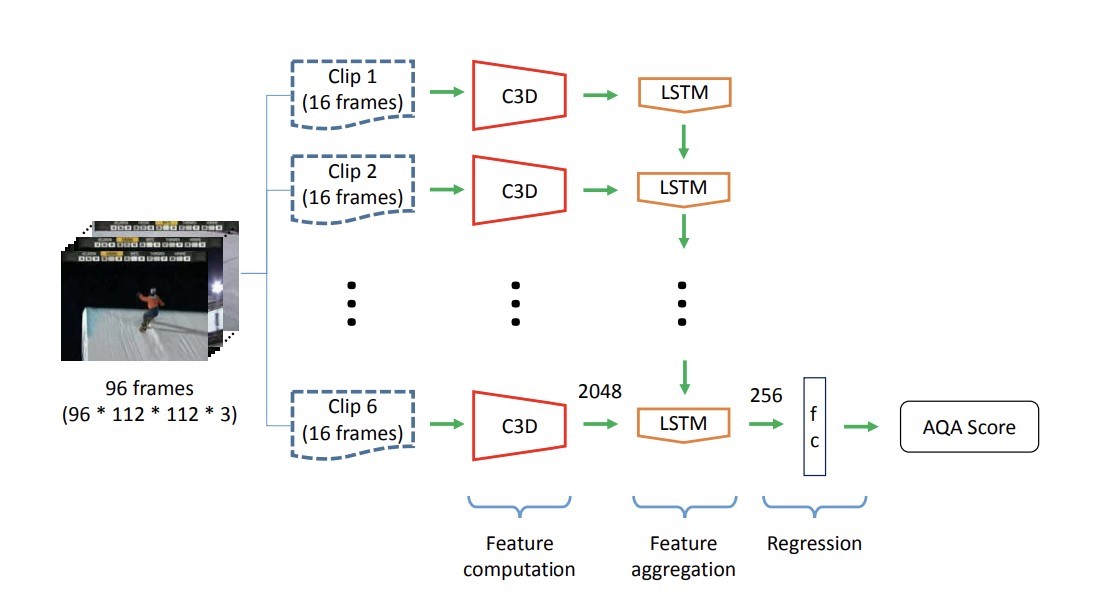


Fig. 1. C3D-LSTM network.

## B. Scope

The main idea of our project is pose estimation, pose correction and counting the number of times a particular action is repeated. Although we will be working with a limited number of exercises, our project, when modified correctly, can be used by Judges in Olympics, trained athletes, bodybuilders and many more. This also has wide application in accordance with the upcoming VR technologies. The basis of the project can be implemented in such a way that the action quality assessment can be used to create virtual simulation of surgeries or some simulation where the user needs to perform a specific task in quite accurate manner.

## C. Technology stack

Android (Flutter and Dart): It will be used to create the android application that will detect video. Python : Language which will be used to write the project. • PyOpenpose library (python) • Posenet library (from google teachable machine) Computer Vision: A technology that allows computers to deduce meaningful information from digital images, videos and other visual inputs. Video Database: The input data that we are going to use to train our model on.

# II. RELATED WORK

There has been only a handful of practical implementations that address on the subject of action quality assessment. Some of the credible works [1,2,3,4] highlight the different aspects of the domain.

Paritosh Parmar and Brendan Tran Morris from University of Nevada, Las Vegas worked on AQA-7 data set to recognize and measure the quality of differnt actions which include - Diving, Gym vault, Skiing, Snowboarding. The paper hypothesize that the actions have various commonalities which could be exploited despite individual differences such as judging criteria to share knowledge or learn a consistent model that can measure quality across multiple actions. To justify the hypothesis the paper proposes to use the C3DLSTM framework illustrated in (Fig.1).

The paper by Ofir Levy, Lior Wolf from The Blavatnik School of Computer Science focuses on the task of counting the number of repetitions for a unique sequence of actions. The proposed solution can handle live data and works by making use of a Convolutional Neural Network (CNN), coupled with a region of interest detection mechanism to minimize the noise in the input feed. Similar approach was followed in our solution as illustrated in (fig. 2.2)

# III. LITERATURE REVIEW

* Action Quality Assessment Across Multiple Actions : Inthis paper, they carried out experiments to see if knowledge transfer is possible in the action quality assessment (AQA) setting. Action quality assessment (AQA) is the process of quantifying how “well” an action was performed or computing a “score” representing the execution quality of an action.
* Live Repetition Counting : This paper talks about how whengiven an input video capturing a scene in which the same action is repeated multiple times in consecutive cycles, how we can count the number of repetitions. • Recognizing Exercises and Counting Repetitions in Real Time : This research paper provides a solution to count repetitions of a physical exercise in real time. The method used in this paper is pose estimation to track athletes, recognize their performed exercises, count the repetitions, and analyze the performance of the repetitions.

IV. PROJECT DESIGN *A. To create the Android Application:*

1) Flutter (framework) - Flutter is an open-sourced software development application developed by Google which is used to craft fast and attractive UI. Flutter is used for developing cross platform applications such as Android, iOS, Windows and many more from a single codebase. Flutter is quite useful since it works with existing code.

2)Dart (language) - Dart is a type safe general purpose language designed by Google to develop fast applications on multiple platforms. It is the primary language of the flutter framework which is used to develop mobile apps using for different platforms using the same codebase.

*B. To create Machine Learning Models:*

1) Action Quality Assessment 2) Repetition Counter

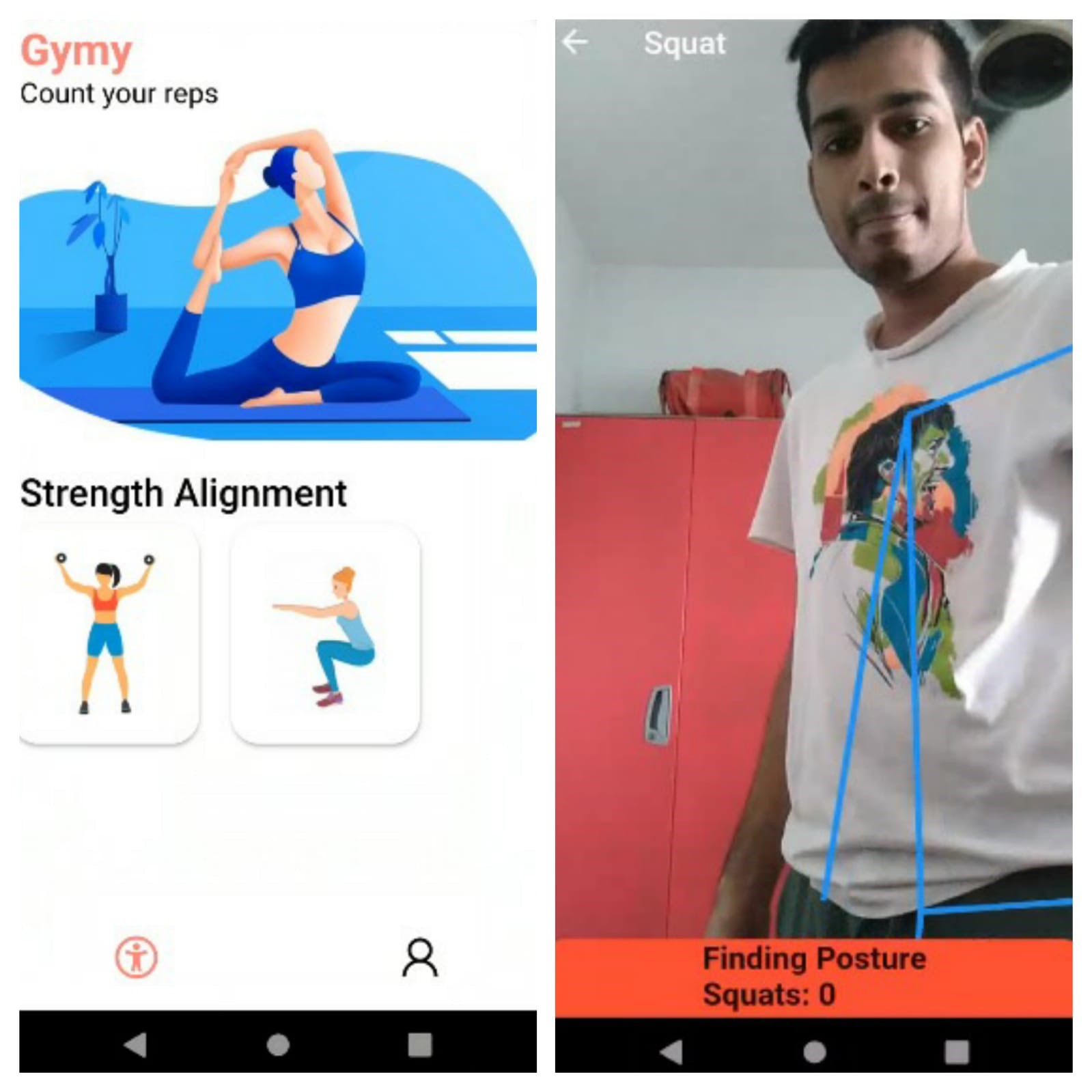
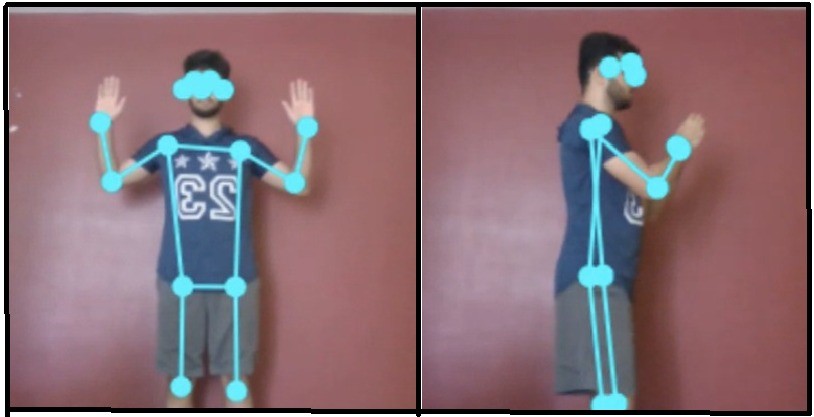
Fig. 2. Image showing the application of use of posenet library to detect different key point on the body

Fig. 3. 1.GUI of proposed Android Application 2.Image showcasing real-time pose detection

# V. PROJECT IMPLEMENTATION

## A. Android Application

The android application will act as an interface for the user to select the type of exercise that he/she wants to perform. On selection of exercise user would be redirected to exercise instance where repetition would be counted and exercise quality would assessed. Once the user successfully completes one repetition the exercise counter will get incremented automatically. If the user performs a wrong exercise a pop would notify the user to correct the posture.

Action Quality Assessment would be performed with help of a machine learning model. Machine learning model implemented using PyOpenpose library along with database consisting of videos of exercise and their corresponding label, to allow the model to learn the right and wrong postures of a given exercise. If the model detects that the user is performing wrong exercise then it will generate a pop to notify the user of the same.

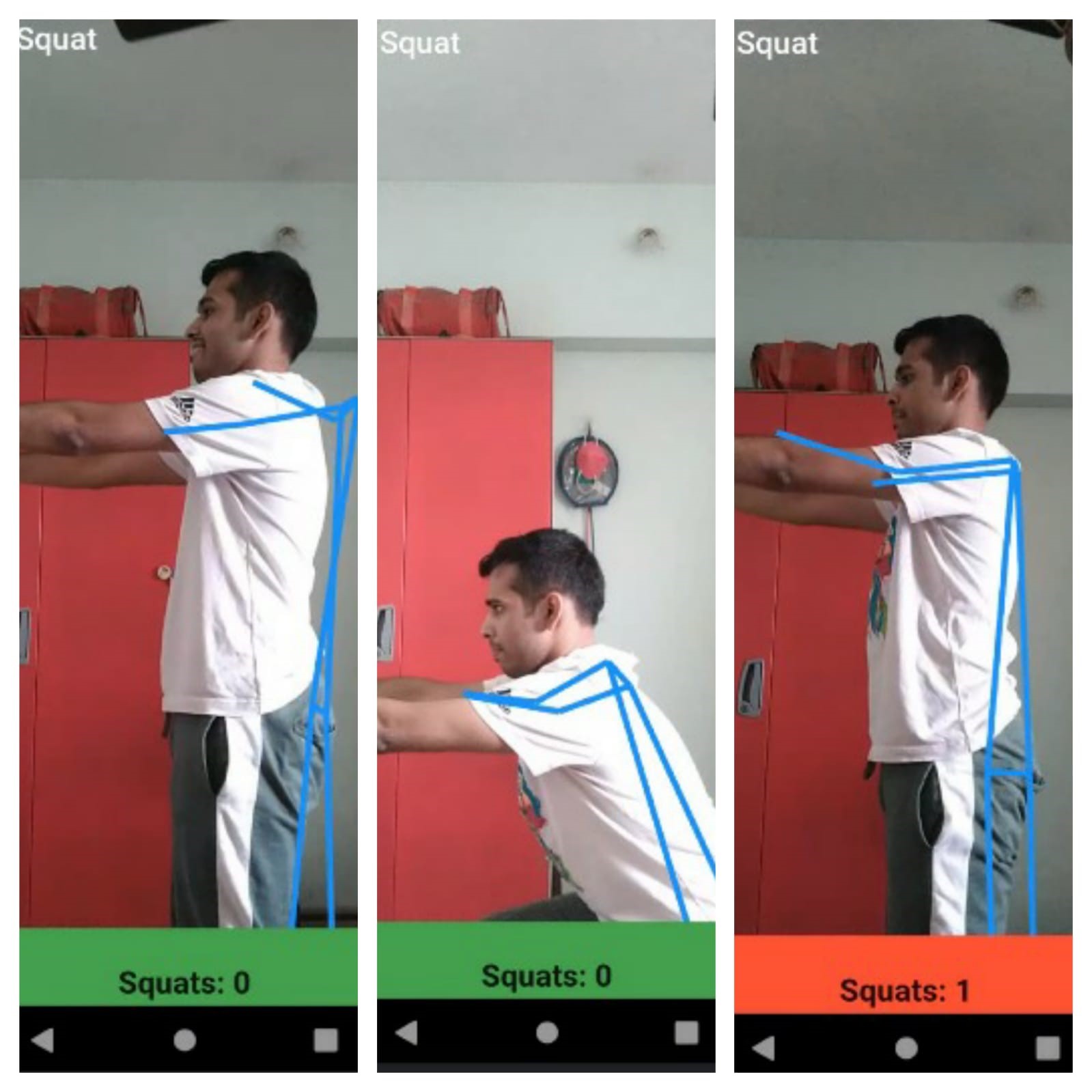
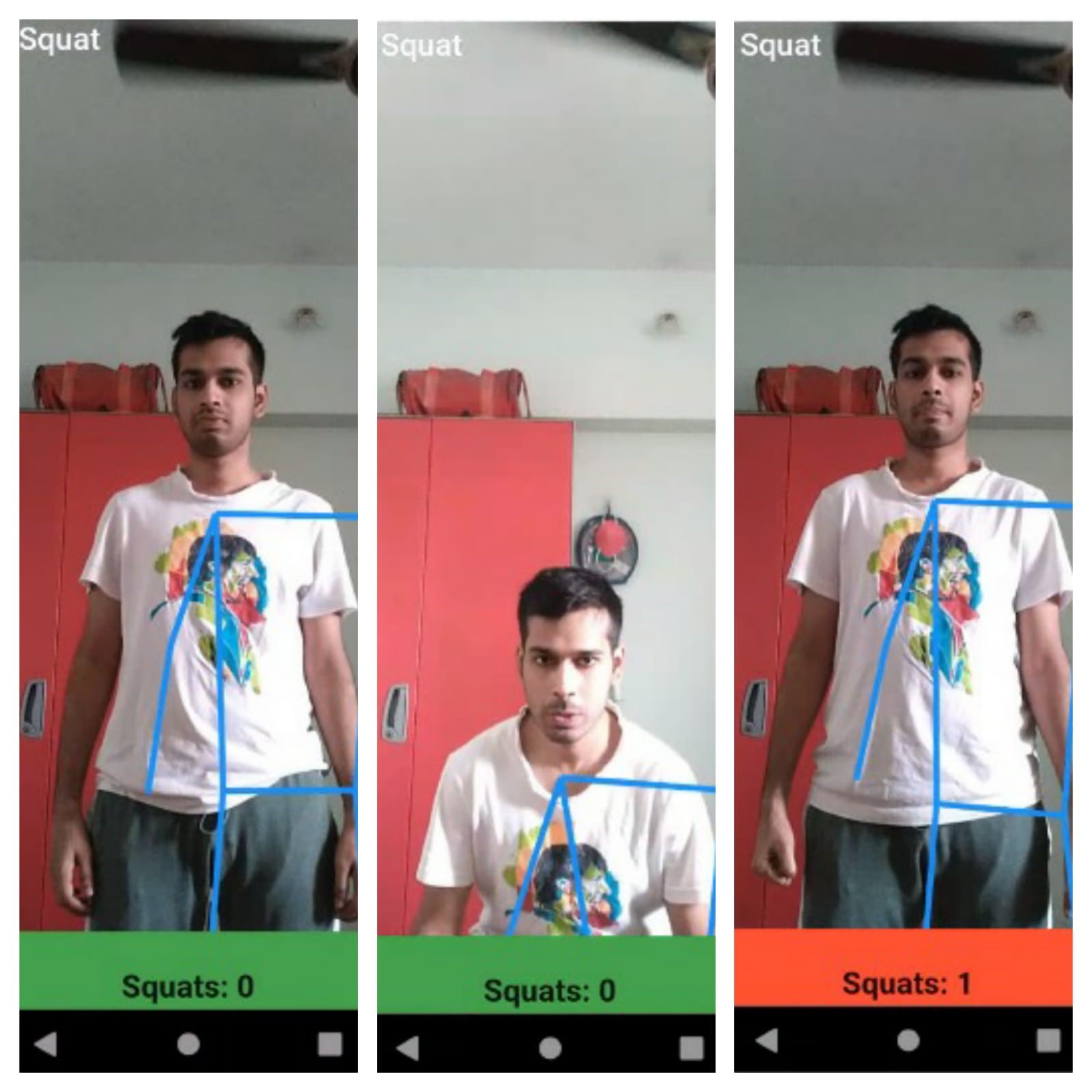
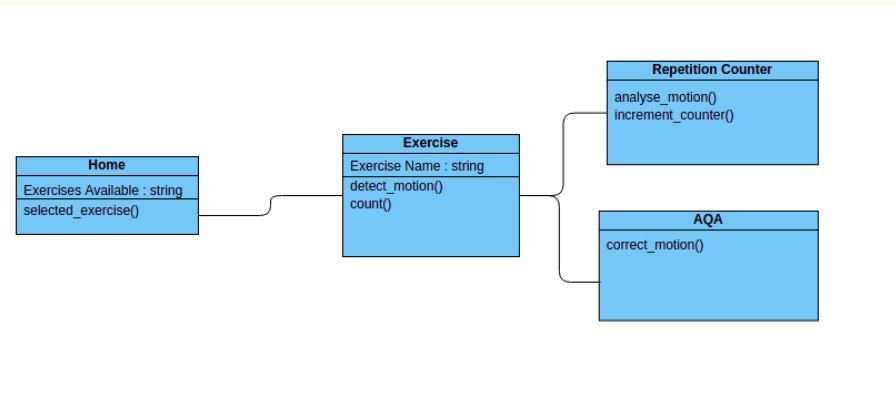
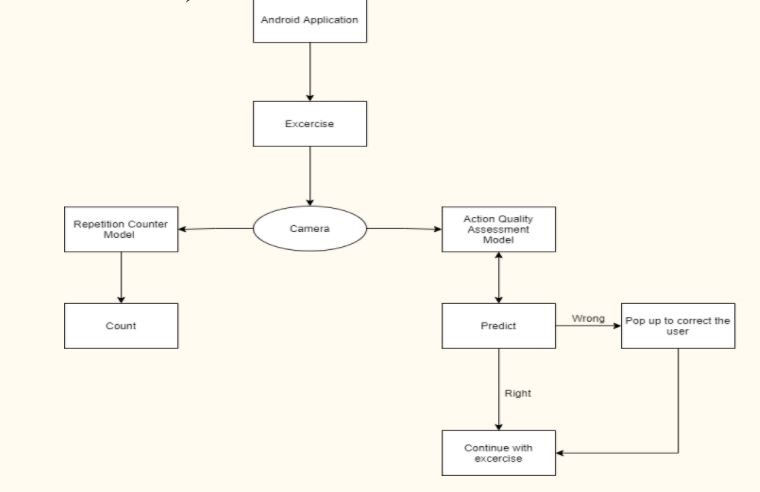


Fig. 4. Position 1 and Position 2

*B. Action Quality Assessment*



## C. Repetition Counter

Similarly, the exercise repetition counter would be a machine learning model. The model would be based on the Posenet library, which can be used to estimate the pose that the user is performing while also detecting the presence of the user. The model will detect when a user completes a specific exercise by detecting its pose and simultaneously increment the counter which would be displayed on application for the user to see. The image (Fig.4 and Fig.5 ) represents the different stages and steps followed for repetition counter.

## D. Pseudo Code

The below snippet of code focuses on tracking the different positions of body part. For exercise like squat positions (keypoints) of right and left shoulder and knee are tracked to count the repetition

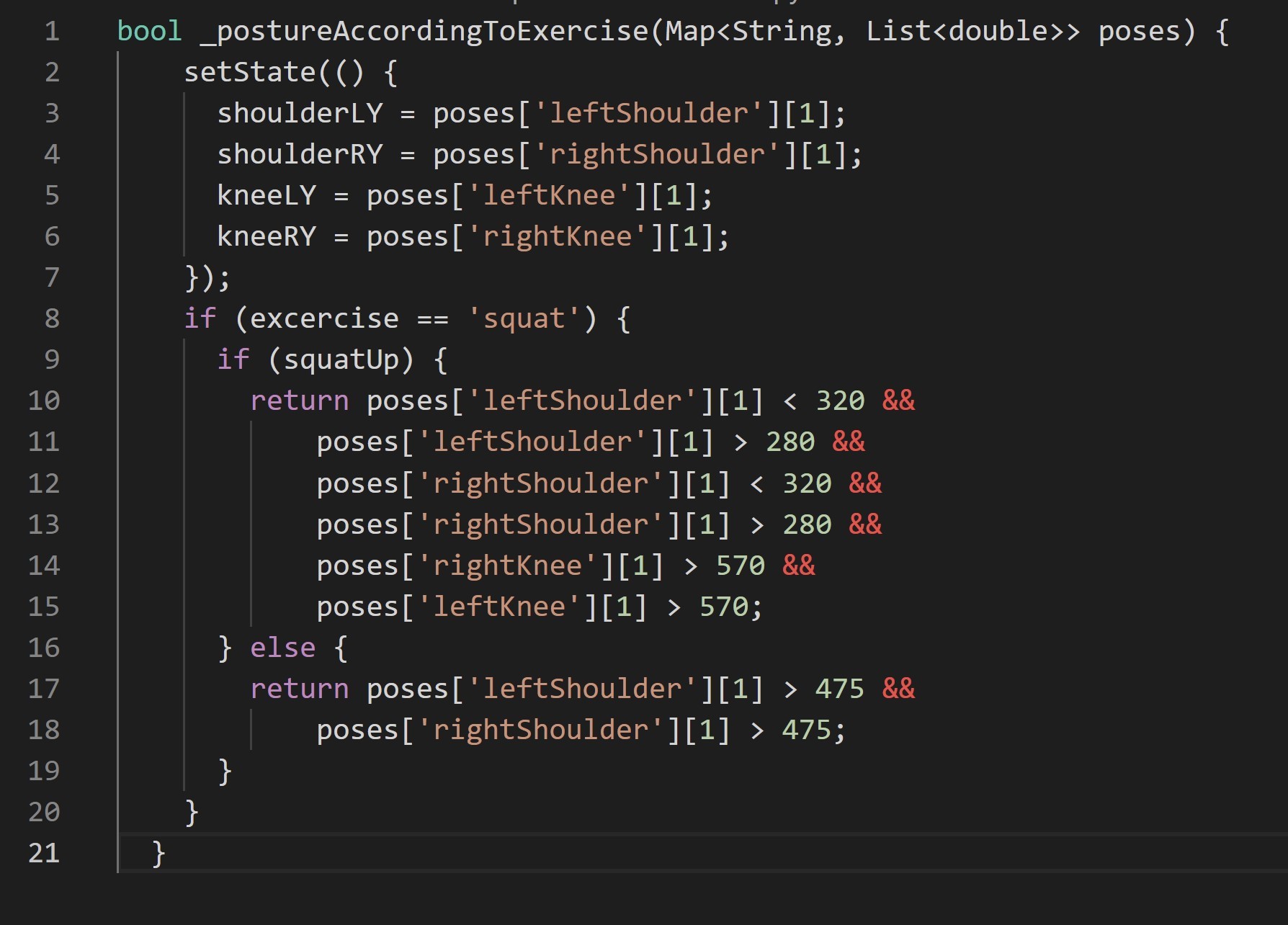


Fig. 5. Flow of Modules

VI.

# VII. CONCLUSIONS

* Created an Android Application which uses camera in-stance of users device to detect emotion of the user in real time.
* The motive behind the current implementation is to integratethe camera instance in an Android Application and also combine it with a Machine Learning Model.

VIII. REFERENCES

1. P. Parmar and B. Morris, ”Action Quality Assessment

Across Multiple Actions,” 2019 IEEE Winter Conference on Applications of Computer Vision (WACV), 2019, pp. 14681476, doi: 10.1109/WACV.2019.00161.

1. O. Levy and L. Wolf, ”Live Repetition Counting,” 2015 IEEE International Conference on Computer Vision (ICCV), 2015, pp. 3020-3028, doi: 10.1109/ICCV.2015.346.
2. Alatiah Talal, and Chen Chen. ‘Recognizing Exercises and Counting Repetitions in Real Time’. ArXiv:2005.03194 [Cs], May 2020. arXiv.org, http://arxiv.org/abs/2005.03194.
3. P. Parmar and B. T. Morris. Learning to score Olympic events. In Computer Vision and Pattern Recognition Workshops (CVPRW), 2017 IEEE Conference on, pages 76–84.

IEEE, 2017

1. P. Parmar and B. T. Morris. Measuring the quality of exercises. In Engineering in Medicine and Biology Society(EMBC), 2016 IEEE 38th Annual International Conference of the, pages 2241–2244. IEEE, 2016.
2. G. Bertasius, H. S. Park, X. Y. Stella, and J. Shi. Am i a baller? basketball performance assessment from first-person videos. In Computer Vision (ICCV), 2017 IEEE International Conference on, pages 2196–2204. IEEE, 2017.