International Journal for Research in Engineering Application & Management (IJREAM)
ISSN : 2454-9150 Vol-08, Issue-01, APR 2022

APT- Your Personal Trainer

[1]Aryan Ajay Singh, [2]Chinmay Manoj Marathe, [3]Dhruvin Bimal Kamani, [4]Yash Surendra Sampat
[1,2,3,4]UG Student, A.P.Shah Institute of Technology Thane, India.
[1]aryanam2000@gmail.com, [2]cmarathe1@gmail.com, [3]dhruvinkamani2201@gmail.com,
[4]yashsampat23154@gmail.com

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. One of the significant researches in the last few years belong in the Computer Vision field especially in the Human Pose Estimation. In this project, pose estimation and deep machine learning techniques are combined to analyze the performance and report feedback on the repetitions of exercises performed in real time. Machine learning when combined with the fitness industry could be useful for the users 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 tasks and we finally deduced two main ideas to tackle this problem.

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 have only been a handful of practical implementations that address 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 different actions which include - Diving, Gym vault, Skiing, Snowboarding. The paper hypothesizes that the actions have various commonalities which could be exploited despite individual differences such as judging

 DOI : 10.35291/2454-9150.2022.0106

criteria to share knowledge or learn a consistent

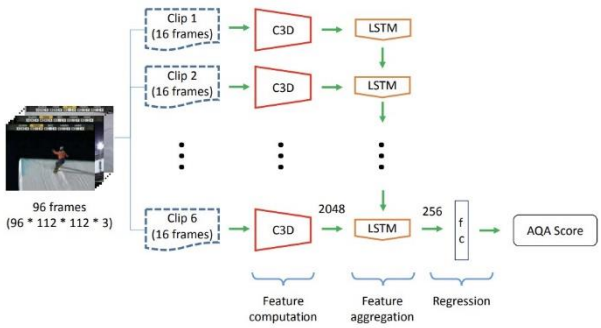

Fig.1 C3D-LSTM framework

model that can measure quality across multiple actions. To justify the hypothesis the paper proposes to use the C3D-LSTM 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

A. *Action Quality Assessment Across Multiple Actions :*

The paper describes the different approaches for developing an Action Quality Assessment model and also introduces a new dataset AQA-7 consisting of 1106 action samples from seven actions with quality as measured by expert human judges.

The paper explains different experiments carried out to see if knowledge transfer is possible in the action quality assessment (AQA) setting. Action quality assessment (AQA) can be defined as the process of quantifying how “well” a given action was performed by computing a “score” given for the execution quality of an action. [1]

B. *Live Repetition Counting :*

The paper talks about how when given 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. [2]

C. *Recognizing Exercises and Counting Repetitions in Real Time :*

The paper mentions the use of pose estimation and deep machine learning techniques to analyze the performance and report feedback on the repetitions of performed exercises in real time.

The method proposed in the paper achieves its target via three phases; pose tracker - to identify and track athletes to apply the algorithm, exercise recognition - to detect the name of the appeared exercises, and counter - to count the number of iterations and indicate the correct and incorrect repetitions.

The essential challenge faced by the authors of this paper was achieving higher frame per second (FPS) value using available computer hardware and equipment. The proposed model was run and tested on a MacBook Pro NVIDIA GeForce GT 750M 2 GB, with an integrated 720p FaceTime HD webcam. The hardware was limited and inefficient to implement deep learning models. Hence, the proposed method had been tested using only a prerecorded video. Another limitation was that the method was unable to distinguish between the competitors and judges and thereby resulting in unnecessary tracking and counting. It also sometimes just ignored any known movement for any person appearing in the frames. Another challenge faced was finding existing publicly available datasets of video recordings with its repetitions of various crossfit exercises. [3]

D. *Learning to Score Olympic Events :*

The paper targets the scoring system used for Olympic games by implementing techniques of AQA (Action Quality Assessment).

This work is based on three frameworks for evaluating actions, which utilizes spatiotemporal features learned using 3D convolutional neural networks. The proposed system is implemented by performing score regression with i) SVR ii) LSTM iii) LSTM followed by SVR. The Olympic events included are diving, vault, figure skating.

Limitations of this project are explained by the following example - a diver may perform perfect actions through the air but might fail to enter the water vertically and thus creating a large splash which is reflected as a poor overall dive score.

As a result, if the dive was scored solely on the basis of a brief clip taken while the diver was in the air, the perfect score that resulted would have no correlation with the actual action quality.

However, if the whole action clip were taken into account, the diver would have been penalized for erroneous entry into the water. [4]

E. *Measuring the Quality of Exercises :*

This paper traverses the problems posed by exercise quality measurement. Their main focus was on Large Amplitude Movement (LAM) exercise which is specifically designed for the children with Cerebral Palsy.

Authors use different machine learning techniques such as SVM, boosted trees, neural networks and dynamic time warping. Out of all the techniques, they got the best accuracy of 94.68% in AdaBoosted tree. The main obstacle faced by

 DOI : 10.35291/2454-9150.2022.0106

the authors were that they had a very uniform and limited size of training as well as testing dataset. [5]

F. Am I a baller? Basketball performance analysis from first-person videos :

Provides an algorithm to analyze a basketball player's performance through first-person videos.

The algorithm initially makes use of a convolutional LSTM model where each small fragment is given a decision about the atomic movement exhibited by the subject(player) based on its spatiotemporal semantics. Then a Gaussian Mixture model is used in order to group different features in order to make them compatible with a linear classifier. Lastly a CNN with 2 layers of 1024 neurons each was used to train the model with 250 labeled pairs of video samples.

However, there are 2 limitations to the solution at the bare minimum. First-person videos of basketball action are extremely volatile and deviate a lot. Hence producing optimal results becomes a challenge. Also, this is a problem that does not need to be really solved. The direction in which a player is looking is nowadays accurately sensed by cameras. [6]

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 applications 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 points on the body.

Fig. 3. Image shows demarcation of key body points as per the posenet model

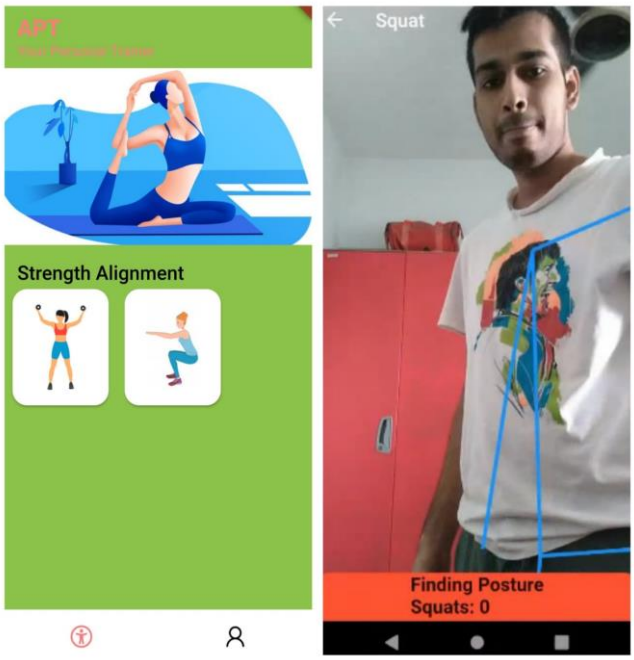

Fig. 4.1.GUI of proposed Android Application

Fig.4.2.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 users would be redirected to exercise instances where repetition would be counted and exercise quality would be assessed. Once the user successfully completes one repetition the exercise counter

 DOI : 10.35291/2454-9150.2022.0106

will get incremented automatically. If the user performs a wrong exercise a pop-up would notify the user to get into correct and requisite posture.

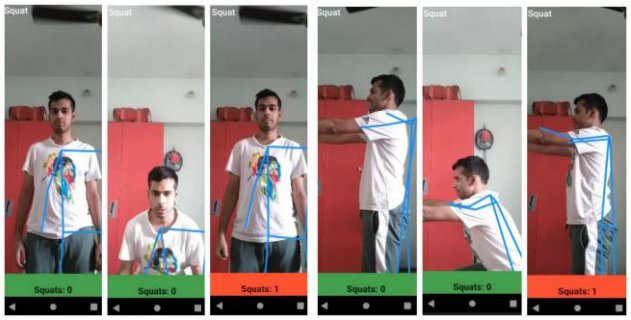

Fig. 5. Position 1 and Position 2

B. *Action Quality Assessment*

Action Quality Assessment would be performed with the help of a machine learning model. Machine learning model implemented using PyOpenpose library along with a 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. Every single frame of the training dataset is assigned a score decided by the judges of sporting events organized by esteemed institutions. The dataset is then split in a ratio of 3:1 and the model is trained. If the model detects that the user is performing the wrong exercise then it will generate a pop-up notification to remind the user of the same.

C. *Repetition Counter*

Similarly, the exercise repetition counter would prevail on a machine learning model based on the Posenet library. It is used to first detect the presence of the user and then to estimate the exercise that the user is performing. The model will detect when a user completes a specific exercise by detecting its pose and simultaneously incrementing the counter which would be displayed on application for the user to see. Incrementing a counter takes place in three major steps. First, the coordinates of respective joints are expected to waver between certain threshold values decided through trials and errors. As soon as the criterion is met, there is a variable called squatUp (in case of squat exercise) which gets set to false(initial value of squatUp is true). Then, the subject is expected to perform the next move of the exercise. Again there is a range of pixels that the coordinates of body points should fall in between in order to call the incrementCounter() function that is implemented using setState() function call of flutter to increment the displayed counter. As soon as the counter is incremented the squatUp variable is reset to true. These three steps keep on repeating cyclically as long as the user wants to perform the exercise post which the user can stop the camera instance in order to kill the executing process. The image (Fig.4.2 and Fig.5) represents the different stages and steps followed for the repetition counter.

D. *Pseudo Code*

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

```
bool _postureAccordingToExercise(Map<String, List<double>> poses) {
  setState(() {
    shoulderLY = poses['leftShoulder'][1];
    shoulderRY = poses['rightShoulder'][1];
    kneeLY = poses['leftKnee'][1];
    kneeRY = poses['rightKnee'][1];
  });
  if (excercise == 'squat') {
    if (squatUp) {
      return poses['leftShoulder'][1] < 320 &&
          poses['leftShoulder'][1] > 280 &&
          poses['rightShoulder'][1] < 320 &&
          poses['rightShoulder'][1] > 280 &&
          poses['rightKnee'][1] > 570 &&
          poses['leftKnee'][1] > 570;
    } else {
      return poses['leftShoulder'][1] > 475 &&
          poses['rightShoulder'][1] > 475;
    }
  }
}
```

Fig. 6. Pseudocode to detect whether correct posture has been achieved by the user.

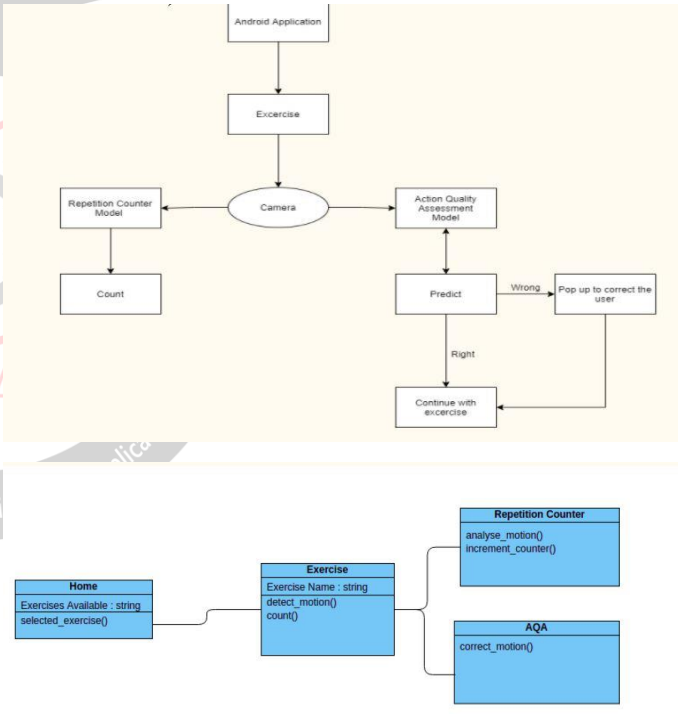

Fig. 7. Flow of Modules

VI. CONCLUSIONS

Created an Android Application which uses the camera instance of the user's device to detect the position of the user in real time. The motive behind the current implementation is to integrate the camera instance in an Android Application and also combine it with a Machine Learning Model. The problem our application solves is extremely unique and has never been tackled before by any existing solutions albeit being ubiquitous.The project is implemented using flutter

 DOI : 10.35291/2454-9150.2022.0106

framework a1nd dart language.The goal our project accomplishes is that users can now optimally perform their exercise without the fear of injuring themselves and having to worry about counting reps while performing the exercise and carrying a different set of mental load during the course of execution. The demand for a personal trainer will change and they will have to upgrade themselves to higher levels of refinement in their erudition. It makes use of the posenet model to detect the key points of the user's body which is stored using tflite extension. These key points are then used to detect the action of the user so as to detect the exercise being performed. These key points have X and Y coordinates which are stored in the form of a map between the abscissa and the ordinate and the name of the key point identified. The relative positioning of body parts is used for detecting the flow of exercise and is used for incrementing the counter. The project makes use of unsupervised learning which helps in real time implementation of repetition-counter using a mobile camera. The performance of the model depends on the quality of the camera as well as the computational power of the device. In the future we will develop a system to analyze complex exercises, which makes use of multiple limbs to be performed appropriately. Also fine tune the different parameters used for detecting the different exercises.

VII. 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. 1468-1476, doi: 10.1109/WACV.2019.00161.

[2] 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.

[3] 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.

[4] 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

[5] 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.

[6] 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.

[7] Chen, Steven & Yang, Richard., Pose Trainer: Correcting Exercise Posture using Pose Estimation, 2018.

[8] Kim I., Deep Pose Estimation implemented using Tensorflow with Custom Architectures for fast inference, 2018.

[9] Soro, A., Brunner, G., Tanner, S., & Wattenhofer, R. Recognition and Repetition Counting for Complex Physical Exercises with Deep Learning. 2019.

[10] Zhe C., Tomas S., Shih-En W., & Yaser S., OpenPose: Realtime MultiPerson 2D Pose Estimation using Part Affinity Fields, 2018.

[11] H. Pirsiavash, C. Vondrick, and A. Torralba. Assessing the quality of actions. In Computer Vision - ECCV 2014 - 13th European Conference, Zurich, Switzerland, September 6-12, 2014, Proceedings, Part VI, pages 556–571, 2014.

[12] V. Venkataraman, I. Vlachos, and P. K. Turaga. Dynamical regularity for action analysis. In Proceedings of the British Machine Vision Conference 2015, BMVC 2015, Swansea, UK, September 7-10, 2015, pages 67.1–67.12, 2015

 DOI : 10.35291/2454-9150.2022.0106