Deep Learning for Fitness

Mahendran N
Indian Institute of Technology Tirupati
Andhra Pradesh, India
csl9m008@iittp.ac.in

Abstract

We present Fitness tutor, an application for maintaining correct posture during workout exercises or doing yoga. Current work on fitness focuses on suggesting food supplements, accessing workouts, workout wearables does a great job in improving the fitness. Meanwhile, the current situation is making difficult to monitor workouts by trainee. Inspired by healthcare innovations like robotic surgery, we design a novel application Fitness tutor which can guide the workouts using pose estimation. Pose estimation can be deployed on the reference image for gathering data and guide the user with the data. This allow Fitness tutor to guide the workouts (both exercise and yoga) in remote conditions with a single reference posture as image. We use posenet model in tensorflow with p5js for developing skeleton. Fitness tutor is an application of pose estimation model in bringing a realtime teaching experience in fitness. Our experiments shows that it can leverage potential of pose estimation models by providing guidance in realtime.

1. Introduction

During these unprecedented times we are bound to be in home without much interaction to rest of the world. Staying physically active in stay-at-home period is a challenging task [7]. Practicing workout exercises or yoga can improve overall body health. The results will be exponential with proper guidance or with a mentor. However, this current situation makes it impossible for guiding physical work via video call as entire posture can’t be seen in 2D frame.

Recent inclusion of technology in fitness, can increase the workout results in less time. Technology can be used in fitness like workout applications, wearables, video conferencing with mentors. They analyse from the data collected from applications[5], fitness machines[3] and improvising the methods can produce better results. All of the existing applications play as support to the physical workouts, improvising physical workouts after workout is done. This is the major limitation of existing applications based on fitness. Doing exercise in home is good but without a coach, the posture, breathe in-out time may not be right. There are multiple implications like body pain, muscle soreness due to extra stress on muscles, wrong posture and so on. There are multiple methods that eliminates this problems even in mobile application like making live coaching, explaining the best posture in detail. Live coaching is possible, time flexibility can also be achieved. In live coaching individual care has to be taken while exercising or yoga but it can be costly and lot of time is needed for individual care. We propose an aspect of providing guidance to physical workout in real-time. The video conferencing guidance can be done for many physical workouts. This is highly effective approach. But it is not suitable for every exercise. Some exercises that exist in social media can be done without training but the correct posture should be maintained while doing the exercise. It is impossible to mentor every time when you train. Real-time suggestions are needed to improvise then and there and to get maximum output.

We propose Fitness tutor, an application to guide you when doing workouts. This applies pose estimation using deep learning to learn the body posture while exercising. Using video chats for posture correction with the mentor on other end, applications that provide animated real person exercising in video followed by voice commands may help but it wont predict the posture is right or wrong. We propose the method of correcting the posture with and without coach. We utilize deep learning PoseNet model in order to help correct the posture. The coach training image can be compared with the current posture that yoga enthusiast is doing. If the hand positions are wrong, the posenet model could identify with comparison with coach’s posture and tells the user where they needs to correct. With the help of AI posenet model, this is possible when exercising.

The proposed approach can be considered as application of PoseNet model in tensorflow in best way possible. PoseNet model accurately identifies the poses of human body parts. With the help of pose estimation model we extends the usage to fitness application. With our proposed approach, user can perform the exercise and this online tu-
tor helps in adjusting the posture.

The main contributions of this proposed paper are

• Posture correction in real-time without human intervention with reference image.

• Reference images can be of their own trainer, online workouts.

• We propose this as an web application using ml5js posenet model. Simple and no installation is needed.

2. Related Work

2.1. Fitness applications

Exercising can have direct impact on body and it is constant. Doing exercise not only enhances the physical aspect but also the mental aspect of human health. Doing workouts provide great relief to the mental stress. It increases productivity and feels fresh in every day to day activities. It increases creativity [16], mental health, life expectancy [4] and decreases health problems [13]. Steinberg et al. [16] research tells us that exercise enhances the positive mood and decreased the negative mood. Existing fitness applications can help in weight loss [11], nutrition and diet, relaxation and general work out information [1].

Researches suggest that participation in elite sport is generally favorable to lifespan longevity [6]. Doing workouts regularly can decrease almost all of the physical problems and gives an upperhand for dealing with the world mentally and physically.

2.2. Deep learning for Health

Researches in healthcare field using AI are booming. Research taking various forms like identification of disease in prior, classification into subcategories, analysing with pre-existing medical data, combining healthcare with robotics using deep learning paves the way to save human lives. Deep learning also implemented in fitness applications like fitbit, smart watches and mobile applications to monitor activity. This data can help predicting specialized suggestions for achieving their goals. We take the idea of implementing deep learning in fitness for impacting lives directly.

2.3. Action recognition

Understanding of the human actions and recognising the particular behaviour is a complex task. Existing researches in this field emerges with 98.69% accuracy models [2] for action recognition dataset UCF101 [15]. Action recognition plays an important role in human computer interaction, robotics. In our proposed method, since the pose estimation has to be in real time and quicker response, we use the ml5js posenet model for action recognition.

| S.No | PoseNet model | Fitness tutor |
|------|---------------|---------------|
| 1    | nose          |               |
| 2    | leftEye       |               |
| 3    | rightEye      |               |
| 4    | leftEar       |               |
| 5    | rightEar      | ✓             |
| 6    | leftShoulder  | ✓             |
| 7    | rightShoulder | ✓             |
| 8    | leftElbow     |               |
| 9    | rightElbow    |               |
| 10   | leftWrist     | ✓             |
| 11   | rightWrist    | ✓             |
| 12   | leftHip       | ✓             |
| 13   | rightHip      | ✓             |
| 14   | leftKnee      |               |
| 15   | rightKnee     |               |
| 16   | leftAnkle     | ✓             |
| 17   | rightAnkle    | ✓             |

Table 1. Table represents the body joints that are present in tensorflow and the joints that are used for comparison by fitness tutor.

3. Proposed Approach

3.1. Pose estimation

Pose estimation in humans using Deep learning research gives upperhand in designing and developing applications based on human pose. Pose estimation is useful in tracing the body parts of human. This helps in knowing the alignment of human body parts and knowing human characteristics better from images. This opens up an opportunity in developing applications based on human pose comparison, body part elevation and alignment.

The main idea of recognising the human actions with the body pose estimation and tracking can open up to a real world applications for scene understanding and vision applications like fitness. This approach completely depends on the pose estimation model [10, 9]. Their effort in proposing a pose estimation model and Dan Oved making it available for tensorflow.js makes this application for fitness possible. The tensor flow posenet model identifies 17 body parts as skeleton. Among them we use only certain body parts for comparison. Table 1 shows the body parts used for comparing the reference image with user data.

We use the posenet model developed in ml5js library. p5js helps in utilizing posenet model and doing calculations on the results. The scene understanding and pose estimation are fetched using their approach. Our proposed approach consists of pose estimation for referenced image and for user input. The exercising posture made by profes-
sionals is compared with the current posture done by the user. We gather the professional exercising posture and develop the skeletons in prior. When the user tries the pre-defined posture by professional, camera fetches the live user data and develop the skeleton for user. We use mathematical approach to compare the images. The resultant skeleton for reference image is stored as JSON file. With the referenced skeleton we match it with the user skeleton generated in real-time for suggesting the corrections. The architecture diagram of the Fitness tutor is shown in figure 1. Comparison between the body skeleton is done on the basis of slope between coordinates.

3.2. Dataset

This method of doing exercise with professional and comparing the posture with them does not have an existing dataset. We depend on online platform in order to fetch the posture images for comparison. Scraping the social media with keywords like exercise, workout results in images which are mostly not related, whole body is not covered in order to get the entire posture. So we utilize properly documented youtube videos and extract postures from randomly chosen youtube videos for comparison. Again in this social media we cant authenticate the exercise posture done by professionals or not. We assume some of the posture are authentic. Then we compare the other posture in order to meet the goal of finding the wrong one. This approach is similar to live camera feed. The live feed from p5js is converted to frames and only each frame is checked then and there. The results are provided for each frame.

3.3. Slope in body parts

Slope between the coordinates is calculated between some of the body parts. Some of the slopes are not needed. For example, the slope between left and right eye is unnecessary and there is not much effect even if it is measured. Thus we find slopes of some parts which have enough data about the user. Slope between two coordinates are calculated using this formula,

\[
Slope = \frac{(y_2-y_1)}{(x_2-x_1)}
\]

With the data, we find slope using above formula to match the posture upto ±0.5 value. Table 2 tells the slope between coordinates are calculated to match with the referenced skeleton.

From the tracking of 17 body parts by pose estimation model, we use specific body parts to match the pose. Slope between posture can help in training the user to match the exact posture. Since the pose estimation models provide high accurate position, the user cannot replicate the reference posture. So the slope values are accepted with a difference of ±0.5 which do not have drastic impact in posture.

| Bodypart A       | Bodypart B     |
|------------------|----------------|
| Left shoulder    | Left elbow     |
| Right shoulder   | Right elbow    |
| Left hip         | Left ankle     |
| Right hip        | Right ankle    |

Table 2. Table represents the body parts slope to be noted for pose comparison and tutoring

We provide suggestions based on slope. For example, if one arm is moved away from referenced posture, user is notified to move the arm up/down according to the slope value.

From the slope values obtained, we compare certain slopes made by body parts whether value is ±0.5 correct. If the slope made by hands, legs in certain posture is more or less correct, then the posture is right. If either of hands or legs is not in right posture, it displays a message of which part has to be corrected in order to get same posture as professional. This way there is a possibility that everyone can exercise in right way as it is to be. This results can suggest for arms to move up/down. Similarly, for both the legs the suggestions are provided to move the leg left or right according to the slope it makes. Figure 2 represents the results obtained by training different poses.

From the results obtained, it is clear that most of the tutoring can be done in this simple way. In the figure, although the posenet generated skeleton are misaligned, the posenet generated coordinates for joints are correct. Thus we use coordinates for correcting the posture.

4. Experimentation

We experimented the work on images taken from Youtube videos. We compared the same exercise/asana with the other images assuming as user. The realtime suggestions provided similar to results in the table. Although most of the experimented images provide correct results, there exists some difficulties in providing a correct posture always. Some of the wrong models provided are shown in this figure 3. Some results for referenced images or user inputs cannot be identified due to the wrong camera angle, no proper lighting and so on. So in order to reduce the misinterpretation of images, placing the camera in right position is recommended.

Also for some exercise postures, the skeleton cannot be identified. The figure tells us that skeletons developed from these posture cannot provide a suggestion to know the posture is right or wrong. This is the drawback of implementation of pose estimation in fitness.

5. Future Work

There is a setback of implementing this pose estimation algorithm for certain postures shown in Figure 3. This problem exists due to the 2D representation of body skeleton
Figure 1. This represents the workflow of the Fitness Tutor application. Get the referenced image and extract the skeleton. It also gets the camera feed frame and extract skeleton from input. Both the skeleton are compared with specific body joints using slope made between them. Results are made according to the slope for the body parts.

Figure 2. The results obtained from using the Fitness Tutor can be seen in this figure. The leftmost image in each row represents the referenced image. The rest of the images are the user input images, their skeleton and the suggestions obtained using pose estimation model. The slope and posture correction result is added for each user input images.
from referenced image and camera feed. This work can be extended to estimation of poses in 3D environment and mentoring in real-time.

Results

Fitness tutor have a direct impact for human lives. This is a mentoring application that learns from 2D reference image and guides the user to mimic the posture. Slope between two vectors are used to compare the body parts of referenced image and live camera feed. In the long term our approach could enable powerful tools for better inclusion of existing 3D pose estimation models for guiding the human workouts.

References

[1] J.P. Higgins. Smartphone applications for patients’ health and fitness. The American journal of medicine, 129(1):11–19, 2016.
[2] M.E. Kalfaoglu, S. Kalkan, and A.A. Alatan. Late temporal modeling in 3d cnn architectures with bert for action recognition. In European Conference on Computer Vision, pages 731–747, 2020.
[3] A. Kulkarni, A. Seetharam, and A. Ramesh. Deepfit: Deep learning based fitness center equipment use modeling and prediction. In Proceedings of the 16th EAI International Conference on Mobile and Ubiquitous Systems: Computing, Networking and Service, pages 394–403, 2019.
[4] I.M. Lee, E.J. Shiroma, F. Lobelo, P. Puska, S.N. Blair, P.T. Katzmarzyk, and Lancet Physical Activity Series Working Group. Effect of physical inactivity on major non-communicable diseases worldwide: an analysis of burden of disease and life expectancy. The lancet, 380(9838):219–229, 2012.
[5] P. Leijdekkers and V. Gay. Improving user engagement by aggregating and analysing health and fitness data on a mobile app. In International Conference on Smart Homes and Health Telematics, pages 325–330, 2015.
[6] S. Lemez and J. Baker. Do elite athletes live longer? a systematic review of mortality and longevity in elite athletes. Sports medicine-open, 1(1):1–14, 2015.
[7] S.M. Nyenhuis, J. Greiwe, J.S. Zeiger, A. Nanda, and A. Cooke. Exercise and fitness in the age of social distancing during the covid-19 pandemic. The journal of allergy and clinical immunology, 8(7):2152, 2020.
[8] S. Pagoto, K. Schneider, M. Jojic, M. DeBissasse, and D. Mann. Evidence-based strategies in weight-loss mobile apps. American journal of preventive medicine, 45(5):576–582, 2013.
[9] G. Papandreou, T. Zhu, L.C. Chen, S. Gidaris, J. Tompson, and K. Murphy. Personlab: Person pose estimation and instance segmentation with a bottom-up, part-based, geometric embedding model. In Proceedings of the European Conference on Computer Vision (ECCV), pages 269–286, 2018.
[10] G. Papandreou, T. Zhu, N. Kanazawa, A. Toshev, J. Tompson, C. Bregler, and K. Murphy. Towards accurate multi-person pose estimation in the wild. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition, pages 4903–4911, 2017.
[11] C.A. Pellegrini, A.F. Pfammatter, D.E. Conroy, and B. Spring. Smartphone applications to support weight loss: current perspectives. *Advanced health care technologies*, page 13, 2015.

[12] S.P. Ramocki. Creativity interacts with fitness and exercise. *The Physical Educator*, 59(1):8–18, 2002.

[13] C. Rominger, A. Fink, B. Weber, I. Papousek, and A.R. Schwerdtfeger. Everyday bodily movement is associated with creativity independently from active positive affect: a bayesian mediation analysis approach. *Scientific Reports*, 10(1):1–9, 2020.

[14] P.Á.L. Román, A.P. Vallejo, and B.B. Aguayo. Acute aerobic exercise enhances students’ creativity. *Creativity research journal*, 30(3):310–315, 2018.

[15] K. Soomro, A.R. Zamir, and M. Shah. Ucf101: A dataset of 101 human actions classes from videos in the wild. *arXiv preprint arXiv:1212.0402*, 2012.

[16] H. Steinberg, E.A. Sykes, T. Moss, S. Lowery, N. LeBoutillier, and A Dewey. Exercise enhances creativity independently of mood. *British Journal of Sports Medicine*, 31(3):240–245, 1997.