IJARCCE ISSN (O) 2278-1021, ISSN (P) 2319-5940

International Journal of Advanced Research in Computer and Communication Engineering

ISO 3297:2007 Certified ⁂ Impact Factor 8.102 ⁂ Vol. 12, Issue 4, April 2023

DOI: 10.17148/IJARCCE.2023.12491

AI Yoga Trainer: A Self-Training Yoga System

Kunal D. Patil [1], Siddhesh P. Patil [2], Nilesh D. Randive [3], Omkar P. Patne [4], Avinash R. Sonule [5]

Department of Computer Engineering, A. C. Patil College of Engineering, Kharghar, India[1,2,3,4,5]

Abstract: This paper presents system based on pose detection for yoga pose recognition. Yoga, which has its roots in India, is a form of physical activity that helps your body and mind work together properly for each of their individual functions. In many nations around the world, yoga of Indian heritage is primarily used to maintain health. As a result, the body's posture while practising yoga has a significant impact on one's health. Many medical professionals advised doing yoga to help patients recover from injuries more quickly and as the greatest way to combat mental health issues like depression, anxiety, and post-traumatic stress. Some yoga practitioners don't do their yoga postures correctly, which causes a variety of physical issues such joint discomfort, disc difficulties, shoulder pain, etc. We are launching a tool for categorising and improving your yoga. It will use MoveNet as a classification model to examine your yoga-related bodily motions.

Keywords: Non-Parametric Weight Feature Extraction (NFWE), Principal Component Analysis (PCA), TensorFlow, MoveNet, 17-Keypoints.

I. INTRODUCTION

With ageing and accidents, humans are more prone to musculoskeletal problems. Any kind of physical activity is required to help avoid this to some extent. Yoga, a form of physical and spiritual practise, has grown significantly in importance among scientists who study medicine. Without the use of medications, yoga has the power to totally cure illnesses and enhance both physical and mental health. Yoga is made up of several asanas, or physical static positions. Yoga involves complicated configurations of postures, making the use of pose assessment difficult. Furthermore, some cutting-edge techniques don't work effectively when the asana requires a horizontal body posture or when both legs are crossed. Therefore, there is a need to create a solid model that can aid in the spread of self-instructed yoga systems.

In computer vision, the study of methods and apparatuses that correct a body's posture is known as pose estimation. Body posture estimation is the process of locating human body components and joints in a given image for the purpose of the project. Data is used with a pre-trained pose estimate model that includes features like Face Identification, Face Mesh, Posture, Holistic, Hair Segmentation, Object Identification, and more. The major objective is to develop a yoga mentoring system that records and analyses the user's movements and postures in order to find problems with the yoga posture.

II. BACKGROUND ON EXISTING SYSTEMS

Use of kinetic sensors and depth cameras is one method for capturing an individual's posture. In 2018, Yu-Liang Hsu created an inertial sensor network that utilised an activity identification algorithm to precisely forecast daily human activity as well as sporting events. They suggested a technique involving the use of two inertial sensing units, each of which had a microcontroller, a triaxial accelerometer, and a gyroscope. To reduce the inertial signal's feature dimension, they combined non-parametric weight feature extraction (NWFE) and principal component analysis (PCA).[1]

In 2018, Hua-Tsung Chen proposed a human posture tracking system using kinetic console sensors in "Computer-Assisted Yoga Training System"[2]. In this method, they designed a yoga training system using a function-based approach. Using Kinect's inertial sensor to extract the contours of the user's body, they are then used to capture the body map and extract pose information such as dominant axis, skeleton and contour-based feature points.

In his article "Multi-Modal Posture Recognition System for Health care Applications," Siddharth Sreeni suggested an alternative approach. Here, they have combined a 3-D depth mapping technique with sensor-based, nine-degrees-of-freedom inertial measurements. These methods are used to record a person's posture parameters in real time, and to predict their pose, their parameters are compared to those of previously recorded poses.[3]
Other techniques include estimating poses based on contours. A SURF algorithm approach is used in Siddharth Patil's "Yoga Tutor: Visualization and Analysis using SURF Algorithm" to determine the pose of users in images. It uses an improved version of the SIFT algorithm is called SURF (Speed-ed Up Robust Feature). The method includes thresholding to remove the background and separate it from the user, as well as gray scaling, or converting the image into gray scale.

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12491

Erosion and dilation are used to smooth out the image's boundary noise. The Canny algorithm, which performs smoothing, gradient correction, and edge tracking, is also used in the prefinal stage. Finally, the practitioner video and reference video are compared using the SURF algorithm.[4]

YEAR	NAME	DRAWBACK
2018	Yu-Liang Hsu, Shih-Chin Yang, Hsing-Cheng Chang, Hung-Che Lai, "Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network"	Required inertial sensing devices which comprised of a micro-controller, a triaxial accelerometer and gyroscope.
2018	Hua-Tsung Chen, Yu-Zhen He, Chun-Chieh Hsu, "Computer-Assisted yoga training system"	Requires Microsoft Kinect which uses specialized hardware.
2018	Siddarth Sreeni, Hari S. R., Harikrishnan R., Sreejith V., "Multi-Modal Posture Recognition System for Healthcare Applications"	Required the use of IR and inertial sensors for depth mapping.
2011	Siddharth Patil, Amey Pawar, Aditya Peshave, Aamir N. Ansari, Arundhati Navada, "Yoga Tutor : Visualization and Analysis Using SURF Algorithm"	Uses computationally heavy algorithms unsuitable for real-time usage.

Figure 1*: Existing System's and their Drawbacks*

III. PROPOSED METHOD

Our system with the aid of TensorFlow models, will estimate poses. The system will be able to give the user feedback on the pose's accuracy based on the user's pose as seen in the video frames. Yoga asanas performed by the user will be automatically recognized in both live and recorded videos. The strategy uses 17 key body point monitoring on the user's live camera feed to automatically extract the user's body posture and compare it with a predefined collection of appropriate yoga postures. Our proposed system will do the following:

1. In the first step, the data is collected from a real time camera or recorded video.
2. Next, our Tensor Flow model recognizes the joint location.
3. And to Thirdly, output a 2-D skeleton of the body and also detect the accuracy of poses.

The procedure for yoga pose detection in our system consists of 3 stages, which are:

1. Pose Extraction

The user's Realtime video is first recorded by the device camera, which is then processed by the MoveNet model to separate the 17 main body points from the stream at the joints of the human body. Each of the 17 joint points for that point is labelled with the index value and the specific x and y coordinates for that joint. These coordinates show the joint's value in x and y.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12491

No.	Keypoint	No.	Keypoint
0	Nose	9	Right knee
1	Neck	10	Right foot
2	Right shoulder	11	Left hip
3	Right elbow	12	Left knee
4	Right wrist	13	Left foot
5	Left shoulder	14	Right eye
6	Left elbow	15	Left eye
7	Left wrist	16	Right ear
8	Right hip	17	Left ear

Figure 2*: Key points detected by MoveNet*

2. *Pose Identification*

It is the next step after extracting the important skeleton points and their coordinates for the human body. In this particular case, all 17 points are used for pose recognition. The x and y coordinates of each point are used to determine the human body's structure, which is then compared to the structure of each asana's ideal poses that were previously fed to the model. After recognizing the user's image, the model outputs the predicted asana with the accuracy rating.

3. *Error Estimation and Feedback*

After correctly identifying the user's pose, the model compares the derived key points of the user image to the predefined set of reference key points of that asana's ideal body structure. Each key point's position in relation to its neighbouring key point is checked, and if any errors or discrepancies are found, a text message instructing the user to adjust their current pose to correct the problem will appear.

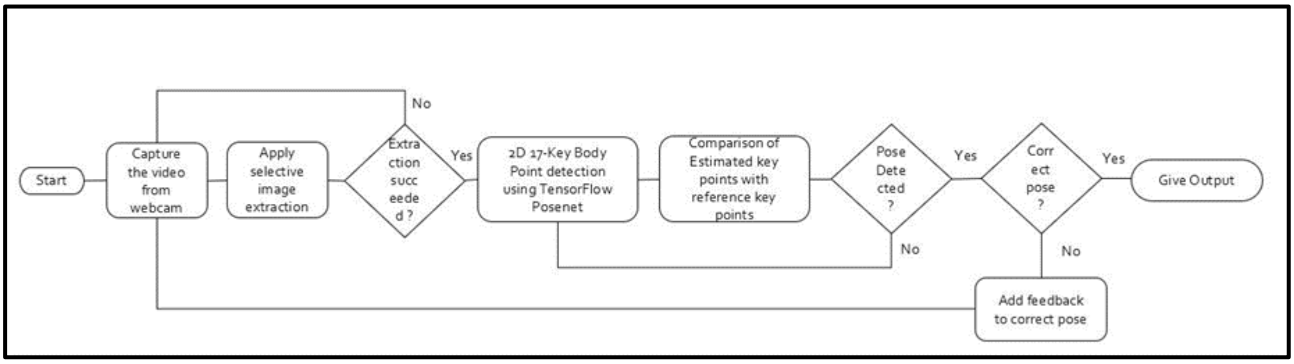

Figure 3*:Flowchart of Proposed System*

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁒ Impact Factor 8.102 ⁒ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12491

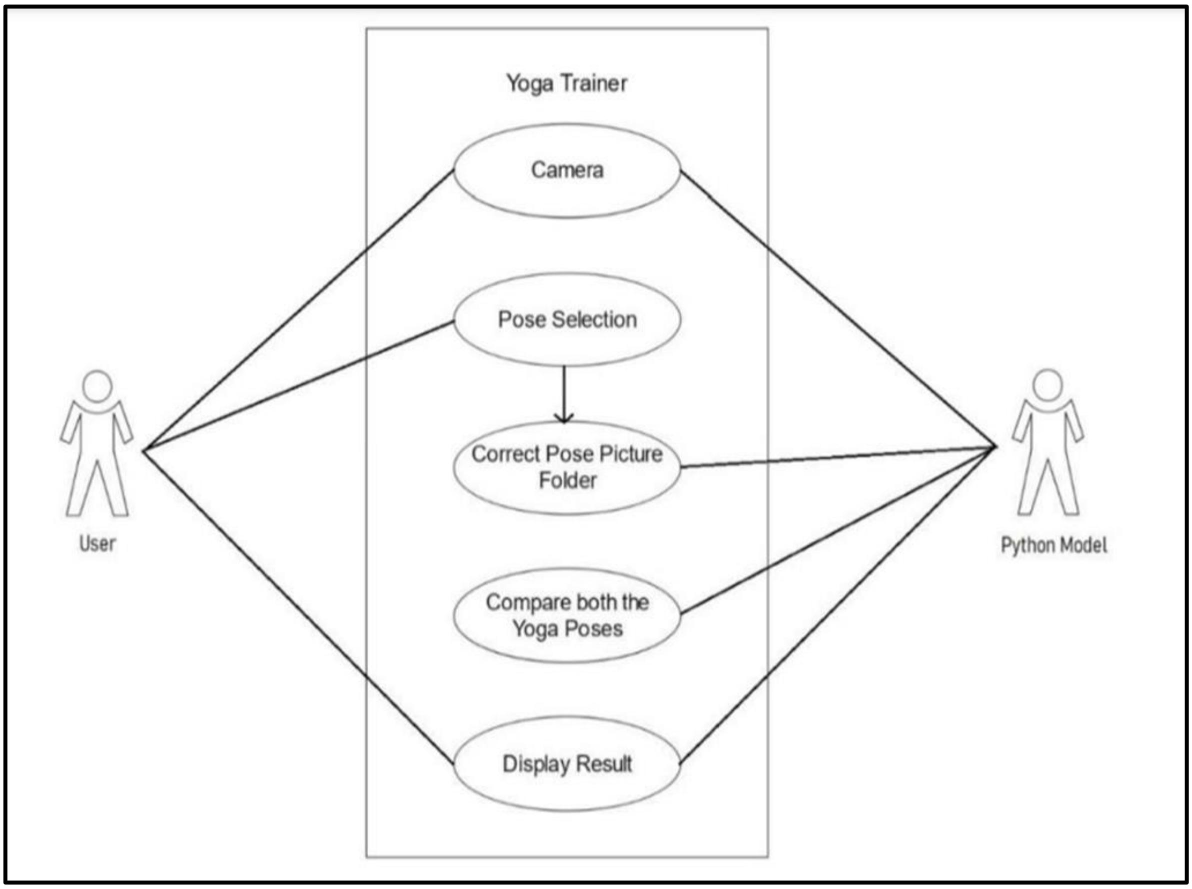

Figure 4*: Architecture of Proposed System*

[5] The feature extractor in MoveNet is MobileNetV2 with an attached feature pyramid network (FPN), which allows for a high resolution, semantically rich feature map output. There are four prediction heads attached to the feature extractor, responsible for densely predicting:

1. ***Person centre heat-map*****:** Predicts the geometric centre of person instances. The person centre heat-map is used to identify the centres of all individuals in the frame, defined as the arithmetic mean of all key-points belonging to a person. The location with the highest score (weighted by the inverse-distance from the frame centre) is selected.

2. ***Key point regression field*****:** Predicts full set of key-points for a person. An initial set of key-points for the person is produced by slicing the key-point regression output from the pixel corresponding to the object centre. Since this is a centre out prediction – which must operate over different scales – the quality of regressed key points will not be very accurate.

3. ***Person key-point heat-map*****:** Predicts the location of all key-points, independent of person instances. Each pixel in the key-point heat-map is multiplied by a weight which is inversely proportional to the distance from the corresponding regressed key-point. This ensures that we do not accept key-points from background people, since they typically will not be in the proximity of regressed key-points, and hence will have low resulting scores.

4. ***2D per-key-point offset field*****:** Predicts local offsets from each output feature map pixel to the precise sub-pixel location of each key-points. The final set of key-point predictions are selected by retrieving the coordinates of the maximum heat-map values in each key-point channel. The local 2D offset predictions are then added to these coordinates to give refined estimates.

ISSN (O) 2278-1021, ISSN (P) 2319-5940
International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁜ Impact Factor 8.102 ⁜ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12491

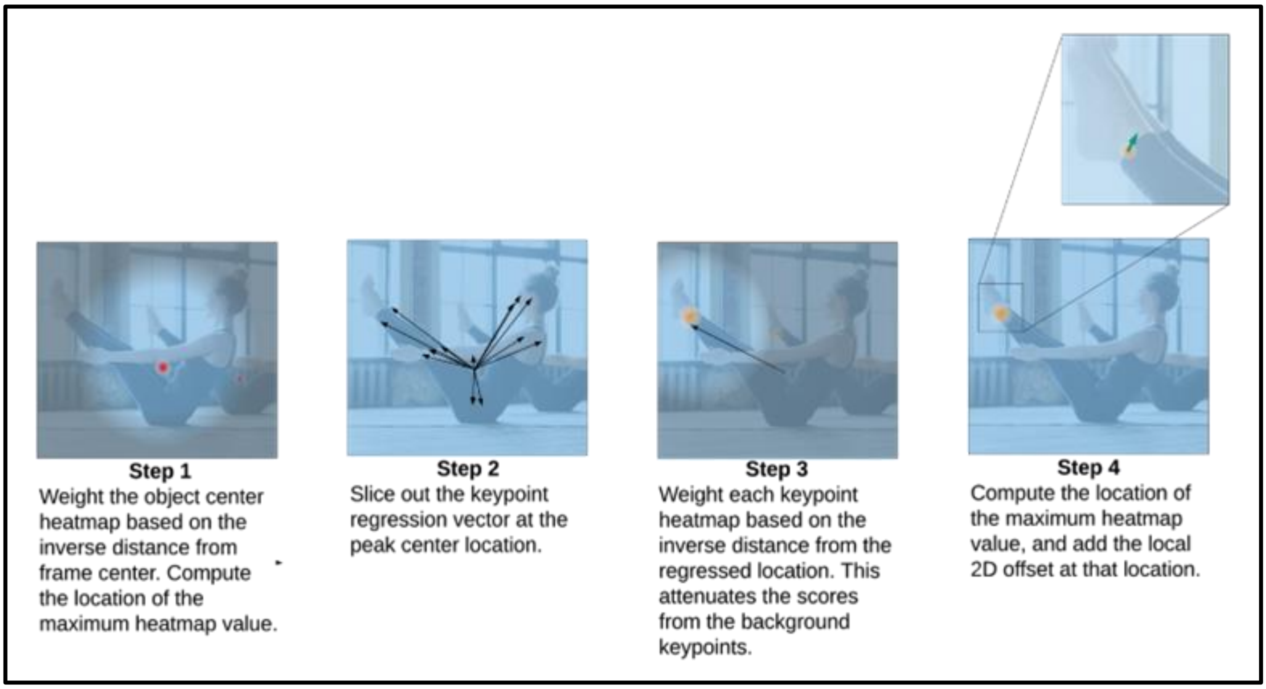

***Figure 5**: Steps for detecting Key-points*

IV. RESULTS

The outcomes of our project are presented below, displaying the body's 2D skeleton while also identifying Yoga poses.

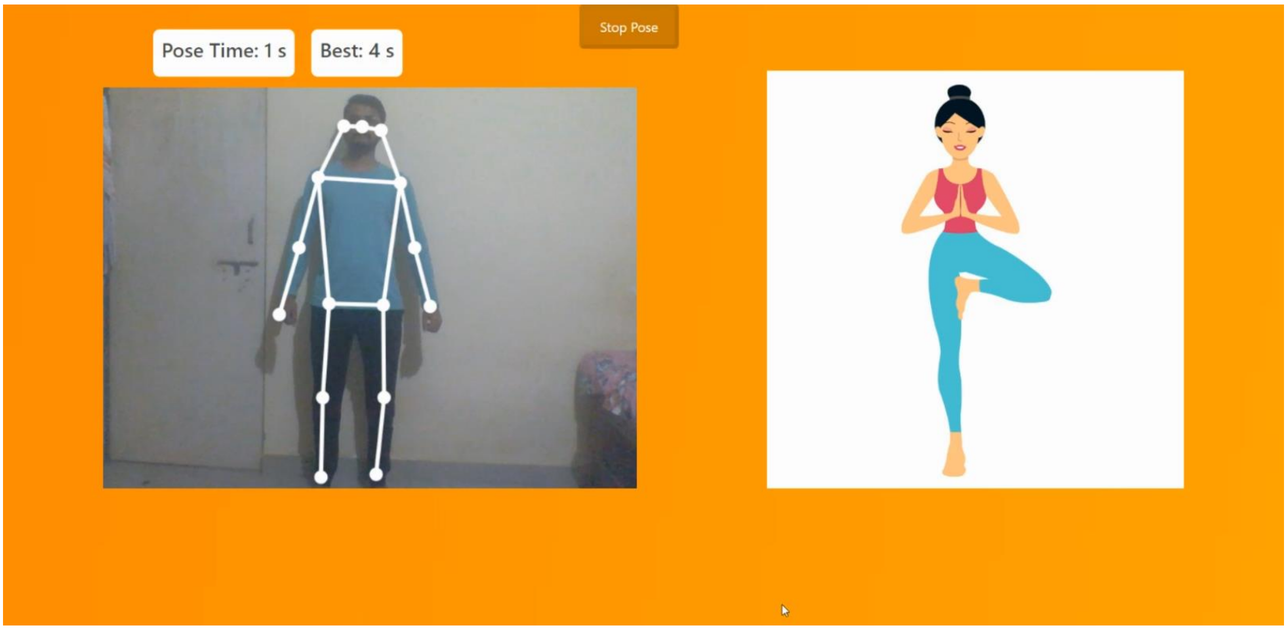

***Figure 6**: Tree Yoga Pose*

International Journal of Advanced Research in Computer and Communication Engineering
ISO 3297:2007 Certified ⁙ Impact Factor 8.102 ⁙ Vol. 12, Issue 4, April 2023
DOI: 10.17148/IJARCCE.2023.12491

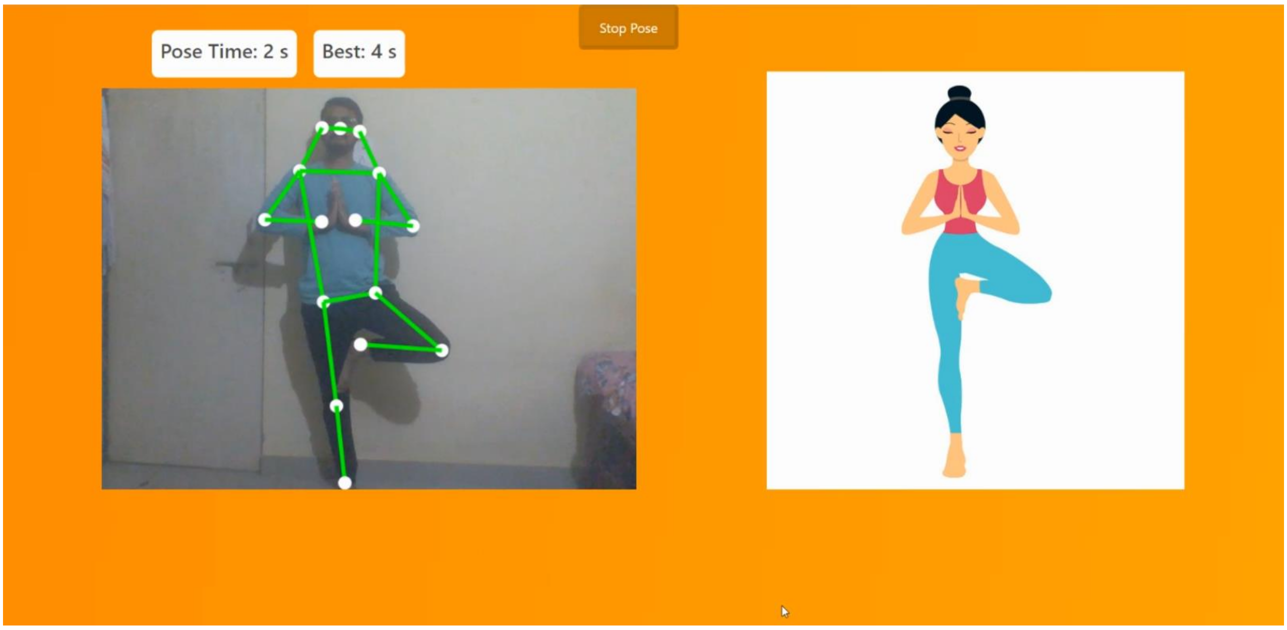

***Figure 7**: Results*

V. CONCLUSION

We tested cutting-edge image classification techniques like CNN while developing this system, and they performed well due to the vast amount of data in the data set. There are still a ton of possibilities to investigate in the domain.

It should be noted that our method can be used with input from a regular RGB camera, eliminating the need for Kinect or any other specialized hardware for Yoga posture identification. Future research can include more asanas and a bigger dataset made up of both images and videos. Additionally, the system can be put into use on a portable device for self-training and real-time predictions. Activity recognition is demonstrated in our work.

REFERENCES

[1]. Chen, H., He, Y. Hsu, C. "Computer-assisted yoga training system" Multimedia Tools Appl 77, 23969–23991 (2018). https://doi.org/10.1007/s11042-018-5721-2.
[2]. S. Sreeni, H. S. R, H. R and S. V, "Multi-Modal Posture Recognition System for Healthcare Applications," TENCON 2018 - 2018 IEEE Region 10 Conference, Jeju, Korea (South), 2018, pp. 0373-0376,
doi: 10.1109/TENCON.2018.8650294.
[3]. S. Patil, A. Pawar, A. Peshave, A. N. Ansari and A. Navada, "Yoga tutor visualization and analysis using SURF algorithm," 2011 IEEE Control and System Graduate Research Colloquium, Shah Alam, Malaysia, 2011, pp. 43-46, doi: 10.1109/ICSGRC.2011.5991827.
[4]. Y. -L. Hsu, S. -C. Yang, H. -C. Chang and H. -C. Lai, "Human Daily and Sport Activity Recognition Using a Wearable Inertial Sensor Network," in IEEE Access, vol. 6, pp. 31715-31728, 2018, doi: 10.1109/ACCESS.2018.2839766.
[5]. https://blog.tensorflow.org/2021/08/pose-estimation-and-classification-on-edge-devices-with-MoveNet-and-TensorFlow-Lite.html.