© August 2024 | IJIRT | Volume 11 Issue 3 | ISSN: 2349-6002

A Deep Learning and Ai Based Personalized Yoga Trainer Using Pose Estimation Model

Hafsa Abdul Basith[1], Sridhar Gummalla[2], Mohammed Waheeduddin Hussain[3]
[1]PG Scholar, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086
hafsabasith211@gmail.com
[2]Professor, Department of Computer Science and Engineering, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086
Sridhar_gummalla@yahoo.com
[3]Professor, Department of Information Technology, Shadan College of Engineering and Technology, Hyderabad, Telangana, India - 500086
mwaheeduddinhussain@gmail.com

ABSTRACT
When embarking on a journey towards a healthy lifestyle, monitoring exercise is a fundamental and empowering practice. An unhealthy lifestyle is pretty much common in today's world, while most teens and adults' resort to gyms as a first-hand savior to improve physical fitness, we often ignore the potential, side effects and risks associated with machine exercises. In today's world, which is fueled by social media influence, we tend to incline towards maintaining a perfect physical fit. However, with our busy schedules, we often lack motivation and the push to exercise and go to the gym regularly. All exercises must be performed under the supervision of an experienced trainer since any incorrect or bad posture may lead to muscle soreness or injuries. Nowadays, people feel more at ease staying in their homes. This situation demands a technology-driven yoga practice. Developing a project that automatically detects a person's pose in an image frame is a difficult task as it depends on several aspects such as the scale and resolution of the image, background clutter, and clothing variations. This project concentrates on building an end-to-end custom yoga trainer. This yoga trainer detects the pose and suggests the correct yoga posture (if incorrect). The system monitors the set of exercises and evaluates the pose until the user achieves the desired and accurate pose. As consistency is the master key that unlocks the door to a healthier you, it also keeps track of the recursion round of the pose.

1. INTRODUCTION

Yoga, originating from India, is an ancient practice that combines physical, mental, and spiritual exercises to create harmony between the body and mind. As one of the oldest disciplines known, yoga offers numerous benefits for both physical and mental well-being. Regular practice can improve flexibility, energy levels, sleep quality, posture, muscle strength, cardiovascular health, and help alleviate chronic pain. However, incorrect postures during yoga can lead to injuries such as fractures, strains, sprains, joint dislocations, nerve damage, and even strokes. Artificial intelligence has made significant contributions across various sectors, including healthcare, mobility, and retail, and is now making its mark in the fitness industry. When people think of fitness, they often consider diet, cardio, gym workouts, and yoga, which remains a popular choice. For those who may not have access to personal yoga instructors, AI can serve as a helpful tool, offering guidance on the accuracy and timing of poses. This project aims to assist users by indicating whether their poses are correct, how long they can hold them, and their best performance in each pose.

Artificial intelligence (AI), a branch of machine learning (ML), is regarded as a crucial technology driving the Fourth Industrial Revolution. Deep learning (DL) technology, which stems from artificial neural networks (ANN), plays a significant role in various fields, including healthcare, image recognition, text analytics, cybersecurity, and beyond. Deep learning algorithms enable automatic feature extraction, allowing for the estimation of pose-related errors and providing users with the necessary feedback for improvement.

Incorporating deep learning with yoga can help individuals of all ages practice yoga poses correctly. This technology can be used to create a user-friendly, cost-free application that can be accessed from anywhere, making yoga more accessible to everyone.

2. RELATED WORK

Several companies have developed innovative technology products for the sports and fitness industry. For example, NADI X-Smart Yoga Pants, which sync with a mobile app to help guide users through proper exercise form. Another well-known product is a portable gym mirror that enables users to see both their workout and reflection simultaneously. SmartMat developed a smart yoga mat equipped with sensors that monitor posture pressure and provide real-time feedback via a mobile app. Additionally, YogaNotch created a wearable device that offers audio feedback on alignment during home yoga sessions.

Identifying yoga postures using a smartphone application presents challenges and requires careful monitoring to provide accurate feedback. The yoga mat with a pressure sensor proposed in was inadequate for correcting posture across the entire body. Similarly, previous products in fall short when it comes to exercises that involve reflection. Additionally, the portable gym mirror, which featured a reflection monitor, still has limitations as it only incorporates a camera without body posture detection or correction capabilities. The cost of using MIRROR, including the product price and monthly training class fees, may also be a barrier for some users.

Nagalakshmi et al. applied 3D landmark point learning from a single image using skinned multi-person linear (SMPL) models and an encoded architecture of classification models. These models included kNN, SVM, and various deep learning models such as AlexNet, VggNet, and ResNet. The proposed dataset, consisting of 13 different yoga postures, was used to train each model. The dataset included poses such as the half-camel pose, standing half-forward bend, bridge pose, seated forward bend, child pose, corpse pose, mountain pose, tree pose, triangle pose, and twisted pose. The classification accuracy was found to be 83% when tested.

Yoga classification plays a vital role in numerous studies that utilize various techniques, such as the OpenPose algorithm for posture estimation, along with machine learning and deep learning methods for pose recognition. In YogaNotch, yoga poses were classified for multiple individuals in real-time using 3D pose estimation from an RGB camera, focusing on 12 sun salutation asanas. These poses were recorded using a webcam, and the process involved constructing a skeletal model, extracting features, and classifying the yoga postures with machine learning models such as support vector machine (SVM), k-nearest neighbors (kNN), naive Bayes, and logistic regression. The method achieved an accuracy rate of 96%.

A computer-assisted self-training system was developed, which detected yoga poses such as the downward-facing dog, warrior 3, and tree pose while extracting the user's body contour. However, their studies achieved an overall accuracy of only 82.84%. Previous research introduced a self-training system that estimated yoga postures by extracting visual features from image data to analyze 12 different yoga poses, including tree pose, warrior 3, warrior 2, warrior 1, downward-facing dog, extended hand-to-big-toe pose, chair pose, full boat pose, cobra pose, side plank pose, plank pose, and lord of the dance pose. The system recorded a body map of the user to extract key feature points and provided visual guidance for adjusting to the correct posture. A yoga expert assessed the system's accuracy, scoring 93.45% of the frames as correctly recognizing the postures.

3. METHODOLOGIES

The system captures images using a camera, which can be a smartphone's built-in camera, a webcam, or a separate camera module—devices that are widely available. The camera serves as the input component of the system, capturing images and providing data to the model. An RGB (Red, Green, Blue) camera is well-suited for this project. On the screen, a reference square box is displayed, guiding the user to position themselves at a specific distance so that their entire body is within the square border. Throughout the routine, the built-in camera or an external camera module is used to continuously capture images of the user. These images are used by the system (phone or computer) for processing.

After capturing the image or visual input, we developed a model using Convolutional Neural Networks (CNN). The proposed system is designed to recognize a wide variety of poses. Consequently, we leverage datasets to their fullest extent. Pose detection is performed using MediaPipe once the user's input is received. This information allows us to create an accurate skeletal representation of the user. MediaPipe identifies key joints and positions on the human body, indexing them from 0 to 32, resulting in a total of 33 landmarks. The MediaPipe library provides a list of these key points in X, Y, and Z Cartesian coordinates. These points can be used in real-time to assess the general posture and orientation of the human body in any image or video stream. The MediaPipe library operates at a frame rate of 30 frames per second. This model will be implemented as either a web or Android application. To measure the angle between the user's joints, we use the NumPy library, and then compare it with image datasets.

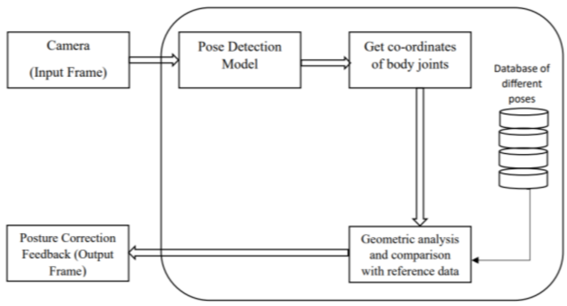

Fig. 2. Block Diagram of Proposed System

4. PROPOSED ALGORITHM

The proposed system is designed to accurately recognize yoga poses performed by a user in front of a webcam. As the user performs a pose, data points detected by MediaPipe are collected and stored in a CSV file. The system then utilizes CNN to compare the input data with the training data, which consists of correct yoga poses, to calculate the pose's accuracy and identify any errors. Feedback is provided to the user through voice assistance or text, helping them perform yoga in a more beneficial manner.
The approach can be broadly divided into four key steps, utilizing various Python modules and libraries. These steps are:

A. Pose Extraction
B. Key-Point Extraction
C. Application of ML Algorithm
D. Error Estimation and Feedback

A. Pose Extraction:
Python offers a range of libraries for image and video processing, with OpenCV being one of the most comprehensive. OpenCV provides numerous functions for handling image and video tasks. Using OpenCV, you can capture video or images from a camera or local storage. The cv2.VideoCapture() function accesses the webcam, while cv2.imread() reads specific frames from images or videos. The cv2.imshow() function displays frames from the webcam. Subsequent operations on the video or image can then be performed as needed.

B. Key-Point Extraction:
A webcam captures real-time images while performing the yoga pose. The pose data extracted with OpenCV is then used to identify 33 key points on the body, utilizing MediaPipe to classify and detect errors during the pose. MediaPipe's Pose Landmarker task identifies human body landmarks in an image, allowing for the detection of key body locations and the application of visual effects. This task employs machine learning (ML) models that can process either single images or a continuous stream of images. It outputs body pose landmarks in both image coordinates and 3-dimensional (x, y, z) world coordinates. MediaPipe also provides utilities to visualize the extracted pose. The identified key points are compared to reference values in the training data, based on their positions and angles. The 33 key points extracted are detailed below:

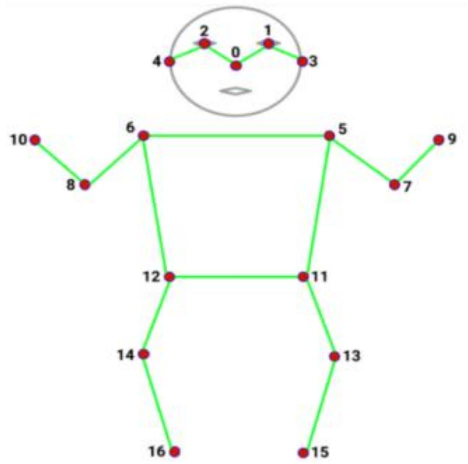

Fig. 1. Skeleton Key-Points

C. Application of ML Algorithm:
For pose classification, we utilize a Convolutional Neural Network (CNN). CNNs are well-suited for processing grid-like input data, such as images. A digital image, which is a binary representation of visual data, consists of a grid of pixels, each with a value that determines its brightness and color.
A typical CNN architecture includes three main layers: convolutional, pooling, and fully connected layers.

- **Convolution Layer:**

The convolution layer is the core building block of a CNN, handling the majority of the network's computational tasks. This layer performs a dot product between two matrices: one being the kernel, a set of learnable parameters, and the other representing the localized region of the receptive field.

- **Pooling Layer:**

The pooling layer replaces the network's output at specific locations by calculating an aggregate statistic from the surrounding outputs. This helps reduce the spatial dimensions of the representation, thereby decreasing the computational load and the number of weights required. The pooling operation is applied independently to each slice of the representation.

- **Fully Connected Layer:**

In this layer, similar to a standard Fully Connected Neural Network (FCNN), every neuron is fully connected to all neurons in the preceding and following layers. This allows it to be computed using standard matrix multiplication, followed by the addition of bias. The Fully Connected (FC) layer helps map the representation between the input and output. The Rectified Linear Unit (ReLU) activation function has gained significant popularity recently.

The function is defined as f(x) = max(0, x), meaning the activation simply acts as a threshold at zero.

D. Error Estimation and Feedback:

To evaluate the accuracy of the performed pose, we use the trained model to provide error estimation and feedback. MediaPipe helps calculate the angles between key joints. These computed angles are compared with the reference values from the training dataset. If discrepancies are found, users receive guidance through both text and voice messages, directing them to adjust their pose and correct any errors.

The methodology of the Yoga Trainer with AI project follows the steps outlined above. Users have two options to check their pose accuracy and receive feedback: either by using a webcam to capture an image or frame or by uploading an image/file via OpenCV. The MediaPipe library identifies the person within the frame, and the subject's x, y, and z coordinates are extracted based on their positioning. These coordinates are saved in a CSV file to aid in model training and to calculate the angles between key joints. The frame is then classified using the CNN and random forest classifier algorithms. The user's pose is identified, and the accuracy of the classification is returned. Errors are calculated based on the angles between critical joints, and feedback is provided to the user, either through voice or text, indicating how well the pose was performed or suggesting corrections.

5. CONCLUSION AND FUTURE ENHANCEMENT

The proposed system is designed to guide users in performing yoga effectively. It features a pipeline for pose detection and human body localization. Once errors are identified, the system utilizes bot technology along with a voice assistant or text format to provide the necessary feedback. The primary goal is to assist individuals in executing yoga postures more accurately without the need for an external instructor or trainer. Key points are extracted from the pose estimation module, MediaPipe, with joints being treated as key points. These key points are used to train the dataset for precise yoga pose correction.

The system is anticipated to deliver satisfactory results. However, based on our analysis, some improvements are needed, and we are actively working on them. Additionally, you can enhance the system's functionality by incorporating modules for additional yoga poses.

The primary Machine Learning module developed is highly adaptable and can be utilized across various platforms, including web applications, Android applications, IoT devices, and more. However, integrating this module on public platforms raises concerns regarding the privacy and security of user data. While this data is valuable for applications such as medical research, law enforcement, and national security, it is crucial to manage data storage and access efficiently while prioritizing security and privacy.

REFERENCES

[1]. Chhaihuoy Long, Eunhye Jo & Yunyoung Nam (2021) Development of a yoga posture coaching system using an interactive display based on transfer learning.

[2]. Chiddarwar GG, Ranjane A, Chindhe M, Deodhar R, Gangamwar P (2020) AI-based yoga pose estimation for android application. Int J Inn Scien Res Tech 5:1070–1073

[3]. Chen HT, He YZ, Hsu CC (2018) Computer-assisted yoga training system. Multi Tools App 77:23969–23991. https://doi.org/10.1007/s11042-018-5721-2 Vol-9 Issue-3 2023 IJARIIE-ISSN(O)-2395-4396 20201 www.ijariie.com 1068

[4]. Anusha M, Dubey S, Raju PS, Pasha IA (2019) Real-time yoga activity with assistance of embedded based smart yoga mat. In: 2019 2nd International Conference on Innovations in Electronics, Signal Processing and Communication, https://doi.org/10.1109/IESPC.2019.8902371

[5]. Pose Classification. https://google.github.io/mediapipe/solutions/pose_classification.html

[6]. Yoga Pose Perfection using Deep Learning: An Algorithm to Estimate the Error in Yogic Poses. Journal of Student Research, 10(3), 2021.

[7]. "Development of a yoga posture coaching system using an interactive display based on transfer learning."https://link.springer.com/article/10.1007/s11227-021-04076-w Long, Chhaihuoy, Eunhye Jo, and Yunyoung Nam The Journal of Supercomputing 78, no. 4 (2022): 5269-5284.

[8]. YOGANOTCH.https://yoganotch.com/ Accessed 16 July 2021

[9]. Karbasi M, Bhatti Z, Nooralishahi P, Shah A, Mazloomnezhad SM (2015) Real-time hands detection in depth image by using distance with Kinect camera. Int J Inter Things 4:1–6. https://doi.org/10.5923/c.ijit.201501.01

[10]. Charles J, Pfister T, Magee D, Hogg D, Zisserman A (2016) Personalizing human video pose estimation.

[11]. Verma M, Kumawat S, Nakashima Y, Raman S (2020) Yoga-82: a new dataset for fine-grained classification of human poses.

[12]. BlazePose: A 3D Pose Estimation Model- David Cochard https://medium.com/axinc-ai/blazepose-a-3d-pose-estimation-model-d8689d06b7c4

[13]. “Use of Data Mining Method for Secure Privacy in Social Networking Sites”, Kale Deepali Anil, “International Journal of Information Technology & Management”, Vol 13, Issue XVII, Nov-2016

[14]. Yoga Pose Detection using deep learning techniques Proceedings of the International Conference on Innovative Computing & Communication (ICICC) 2021.

[15]. Development of a yoga posture coaching system using an interactive display based on transfer learning. The Author(s), under exclusive license to Springer Science+Business Media, LLC, part of Springer Nature 2021.

[16]. Manap MS, Sahak R, Zabidi A, Yassin I, Tahir NM (2015) Object detection using depth information from Kinect sensor. In: 2015 IEEE 11th International Colloquium on Signal Processing & Its Applications. pp 160–163. https://doi.org/10.1109/CSPA.2015.7225638

[17]. AI-Based Yoga Pose Estimation for android application uploaded by International Journal of Innovative Science and Research Technology Pune- 411041, India Issue 9, September 2020.

[18]. Real-time Yoga recognition using deep learning Neural Computing and Applications volume 31, pages 9349–9361 (2019).

[19]. On-device, Real-time Body Pose Tracking with MediaPipe BlazePose - Valentin Bazarevsky and Ivan Grishchenko, Research Engineers, Google Research https://ai.googleblog.com/2020/08/on-device-real-time-body-pose-tracking.html

[20]. Review of deep learning: concepts, CNN architectures, challenges, applications, future directions, March 2021 Journal of Big Data 8(1) DOI: 10.1186/s40537-021-00444-8.

[21]. Pose Estimation Quality, Pose Detection Model (Blaze Pose Detector), Pose Landmark Model (Blaze Pose GHUM 3D), Cross-platform Configuration Options. https://google.github.io/mediapipe/solutions/pose.html

[22]. Yoga-Guru Real-Time yoga pose correction system using deep learning methods 2021 International Conference on Communication Information and Computing Technology (ICCICT), June 2021.

[23]. Wearable X. https://www.wearablex.com/. Accessed 16 July 2021

[24]. MIRROR.https://www.mirror.com/. Accessed 16 July 2021

[25]. SmartMat.https://www.smartmat.com/about/. Accessed 16 July 2021

[26]. Zhang Z (2012) Microsoft kinect sensor and its effect. IEEE Multi 19:4–10. https://doi.org/10.1109/MMUL.2012.24

[27]. Chun J, Park S, Ji M (2018) 3D human pose estimation from RGB-D images using deep learning method. In: Proceedings of the 2018 International Conference on Sensor, Signal and Image Processing, pp 51–55 https://doi.org/10.1145/3290589.3290591

[28]. Yoga Pose Estimation and Feedback Generation Using Deep Learning, Computational Intelligence and Neuroscience Volume 2022 Publisher-Hindawi Limited London, United Kingdom, 2022.

[29]. Pose Estimated Yoga Monitoring System https://papers.ssrn.com/sol3/papers.cfm?abstract_id=3882498 Anilkumar, Ardra, Athulya KT, Sarath Sajan, and Sreeja KA. Available at SSRN 3882498 (2021).