Real-Time Yoga Pose Detection Using a Fully Connected DNN and Mediapipe Pose Landmarks

Azizur Rahman Ankon
Bangladesh University of Business and Technology (BUBT)

Zarin Akter
Bangladesh University of Business and Technology (BUBT)

***Abstract*—Sustaining mental and physical health has become essential in the modern digital age, as many individuals spend a lot of time on mobile devices and participate in fewer physical activities. Yoga, an age-old discipline that encourages flexibility, mental clarity, and physical well-being, is a useful remedy. In this research, a unique method for real-time yoga posture detection is presented, combining a deep learning classifier with Mediapipe's Pose Estimation model. The five fundamental yoga poses—Desert Dog, Goddess, Plank, Tree, and Warrior—are identified by the system. The suggested model has an amazing 95% accuracy rate and instantaneous feedback, which makes it perfect for posture correction and real-time yoga instruction. This technology aims to promote healthier lives by promoting alignment and appropriate technique during yoga sessions.**

Keywords: Yoga Posture Detection, Deep Learning, Pose Estimation, Mediapipe, Real-Time Feedback,Real-Time Classification.

INTRODUCTION Yoga is much more than just a physical workout; it is an age-old discipline that unites the mind, body, soul, and consciousness in harmony. It improves general health, strength, and flexibility, but only when the postures are done correctly can it be truly beneficial. Every yoga pose has its own set of advantages, therefore it's important to align and perform them correctly. Yoga saw a sharp increase in popularity during the COVID-19 pandemic as more individuals looked for at-home health solutions. Yoga is traditionally taught under the supervision of a qualified teacher, but not everyone can afford or readily obtain expert assistance. Because of this, a lot of people learn and practise yoga from mobile apps or internet videos, frequently without realising if their postures are correct. The likelihood of incorrect form is increased by this absence of instruction. which, if continued over time, can result in major difficulties like joint discomfort and other health complications.

Artificial intelligence (AI) systems that use human position estimate techniques can help detect and correct yoga poses in real time, reducing these dangers.

Digital Object Identifier 10.1109/MCE.YYYY.Doi Number
Date of publication DD MM YYYY; date of current version DD MM YYYY

Skeletal representations are used in the deep learning discipline of human pose estimation to distinguish different regions of the human body. However, because there are so many different yoga postures, degrees of flexibility for movement, and variances in body appearance, real-time yoga pose identification offers a unique problem.

This study suggests using Deep Neural Network (DNN) which is a type of (FNN) Feedforward Neural Network to develop a yoga stance detection model that is both affordable and effective. During testing, the study used 8,034 photos that represented five different yoga poses: warrior, plank, goddess, downward dog, and tree. It obtained a 95% accuracy rate. The outcomes give a thorough assessment of the method and provide comparisons with accepted practices today. Deep learning is very good at analysing photos, videos, and unstructured data since it requires less constant human intervention. Deep learning's ability to carry out feature engineering on its own is one of its main advantages. The system increases recognition and classification accuracy of yoga poses while speeding up the learning process by scanning data and finding pertinent information.

RELATED WORK Yoga Asana Identification: A Deep Learning Approach Technological and scientific developments have made it possible to investigate a wide range of opportunities in interdisciplinary fields. Modern technologies like computer vision, machine learning, and artificial intelligence are currently being used to create a wide range of real-time applications. Yoga, a popular form of exercise for preserving one's physical and mental health, has profited from these developments as well. To provide practitioners with a virtual trainer in this situation, researchers have suggested an automatic yoga posture identification system that uses pictures or videos. Conventional image categorisation techniques like Convolutional Neural Networks (CNNs) were tested during development. Unfortunately, these techniques could not produce adequate findings because of the small amount of data that was available. Transfer learning was used with the VGG16 architecture, which was pre-trained on ImageNet, in conjunction with a Deep Neural Network (DNN) classifier to increase performance. This strategy had an 82% prediction accuracy, which is encouraging. There is still a great deal of room for investigation in this field. Future research could examine the movement and accuracy of yoga poses using video analysis in addition to static picture analysis. Long Short-Term Memory (LSTM) networks, 3D CNNs, Deep Pose Estimators, and Gated Recurrent Units (GRUs) are a few examples of architectures that work well for sequence analysis and could be used to improve real-time video-based posture identification and correction. iYogacare: Real-Time Yoga Recognition and Self-Correction for Smart Healthcare The YOGI dataset, created specifically for this study, consists of five mudras (about 500 photos each mudra) and ten yoga postures (400–900 images per stance). It was possible to extract features from skeletons using hand mudras and body postures. For the purpose of constructing skeletons, two different algorithms were utilised: one for mudras and another for yoga positions. In order to extract joint angles as characteristics for deep learning and machine learning models. With RandomSearch CV, XGBoost produced the best accuracy of these models, at 99.2%. Their post goes into detail about the entire system. Yoga Pose Classification: A CNN and MediaPipe-Inspired Deep Learning Approach In this study, a novel architecture for low-latency classification of yoga positions into five categories is presented. Before supplying images to the network, the suggested model skeletonises them by identifying body keypoints using the MediaPipe library. The paper shows how skeletonised input improves accuracy by comparing performance of skeletonised vs non-skeletonized deep learning models. VGG16 obtained the highest validation accuracy (95.6%) on non-skeletonized images, followed by InceptionV3, NASNetMobile, YogaConvo2d (89.9%), and InceptionResNetV2. Using skeletonised images, the suggested YogaConvo2d model outperformed VGG16, InceptionResNetV2, NASNetMobile, and InceptionV3, achieving the greatest accuracy of 99.62%.

Recognition of Yoga Postures Using Deep Learning.Applying the Y PN-MSSD Model to Yoga Instructors

Because tracking important body locations based on predetermined human body components is a unique difficulty, human posture assessment has drawn a lot of study attention. The Y_PN-MSSD model, which can identify seven yoga asanas with an astounding accuracy of 99.88%, is presented in this research. The model combines MobileNet SSD for person detection in every frame with Pose-Net for key point detection. Live tracking is made possible by this, giving users the ability to adjust their posture instantly. Y_PN-MSSD provides better accuracy than the typical Pose-

Net CNN model. The model's potential uses extend beyond yoga, including activity recognition in sports, healthcare, and surveillance. Input augmentation can further improve this model for self-training and real-time feedback.

METHODOLOGY

Pose Estimation Using Mediapipe

The Mediapipe library contains a strong pre-trained Pose model that can be used to estimate 33 important landmarks on the human body with a mobile camera. The locations of the head, shoulders, elbows, hips, knees, and ankles are captured by these markers, which are crucial for precise posture detection.

These landmarks, which we can extract from each frame of an image or video feed using Mediapipe, serve as the cornerstone of our system for detecting and correcting yoga posture. The retrieved landmarks are normalised and converted into a feature vector for every body point, which represents the joints' 3D coordinates (x, y, and z). This makes it possible to estimate poses precisely and gives users immediate feedback to help them get better at their form.

Data Collection

The "Yoga Poses Dataset" and "Yoga Pose Classification" datasets from Kaggle served as the source of the dataset for yoga posture recognition. Additional data was gathered from other web sources. We will concentrate on the following five poses: Warrior, Plank, Goddess, Downward Dog, and Tree. We processed up to 950 photos each position.

The dataset was expanded using common methods like rotation, zoom, and scaling to increase the model's generalisation. This improved the model's capacity to recognise and categorise yoga positions across a range of variants.

Feature Extraction and Classification

Every yoga stance frame was processed to obtain 99 characteristics (33 landmarks × 3 dimensions) for classification and feature extraction. After that, these features were compressed into a single vector, which was used as the deep neural network (DNN) classifier's input.

The input layer, hidden layers, and output layer are the three main parts of the deep learning model architecture. The 99-dimensional feature vector corresponding to the pose landmarks is fed into the input layer. The activation functions of the hidden layers, which are made up of many dense layers, alternate between Tanh and ReLU in order to create non-linearity and improve feature extraction.

Lastly, a softmax layer with five nodes—Downward Dog, Goddess, Plank, Tree, and Warrior—represents each yoga posture in the output layer. The activation function of softmax makes sure that Pose categorisation is made possible with accuracy by the model's output, which is a probability distribution over the five classes.

Model Training

A crucial stage of the data science development lifecycle is model training, where the objective is to minimise the loss function by optimising the machine learning algorithm through weight and bias adjustments. In our situation, to guarantee effective gradient descent during backpropagation and minimise total loss, we employed the Adam optimiser in combination with a categorical cross-entropy loss function.

We used early halting and learning rate reduction strategies to make sure there was no overfitting and faster convergence. The model belongs to a subclass of Feedforward Neural Networks (FNNs) called Deep Neural Networks (DNNs), which are distinguished by the presence of many hidden layers. This TensorFlow and Keras-built model meets the requirements for a fully connected DNN because it has several hidden layers.

The sequential architecture of the DNN is made up of thick layers, which make it has a strong capacity to pick up intricate representations and comprehend the subtleties involved in classifying yoga poses.

Evaluation Metrics

In order to fully analyse the model's classification effectiveness, important measures like accuracy, precision, recall, and F1-score were used to assess its performance. In order to show how well the model distinguished between the various yoga postures, a confusion matrix was also created to visualise the classification findings for each pose.

Confusion matrix

$$\begin{bmatrix} 1121 & 7 \\ 13 & 284 \end{bmatrix}, \begin{bmatrix} 1123 & 10 \\ 25 & 267 \end{bmatrix}, \begin{bmatrix} 1118 & 21 \\ 12 & 274 \end{bmatrix},$$

$$\begin{bmatrix} 1117 & 42 \\ 47 & 219 \end{bmatrix}, \begin{bmatrix} 1088 & 53 \\ 36 & 248 \end{bmatrix}$$

FIGURE 1. Goddess

FIGURE 2. Tree

FIGURE 3. Downdog

FIGURE 4. Plank

FIGURE 5. Warrior

RESULTS

Model Performance

On both the training and test sets, the model achieved an overall classification accuracy of 95.18% and 91%, respectively. For each of the yoga poses—Downward Dog, Goddess, Plank, Tree, and Warrior—we assessed accuracy, recall, and F1-score. These metrics are summarised in the adjacent table, which shows how reliable and consistent the model is in its performance across all postures.

Real-Time Feedback

On a typical CPU, our system showed real-time performance with a delay of about 20 ms per frame. Additionally, it is designed to function well on smartphones, allowing for fluid video processing and immediate feedback when practicing yoga in real time. The technology guarantees a smooth and responsive experience, which makes it perfect for interactive, live yoga practice, especially considering the significance of high-quality cameras and minimal latency in real-time applications.

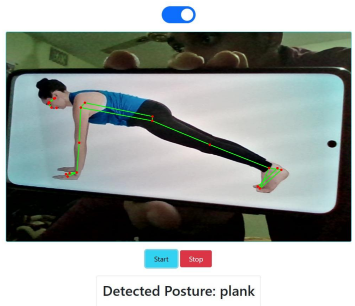

FIGURE 6. Plank

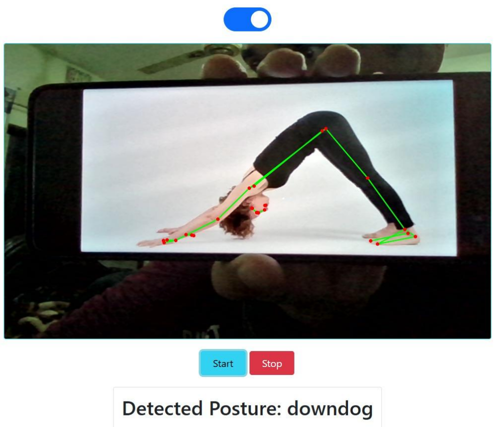

FIGURE 7. Downdog

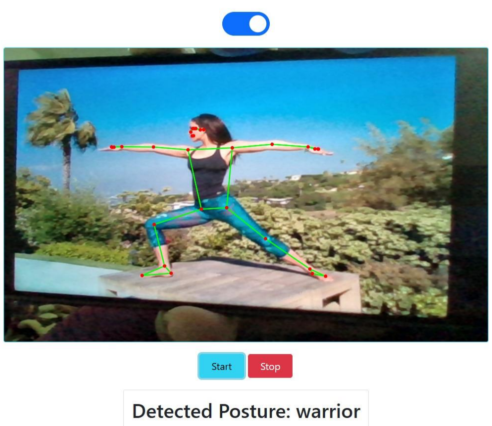

FIGURE 8. Warrior

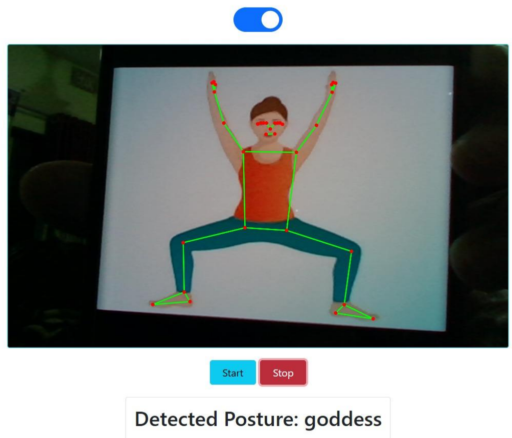

FIGURE 9. Goddess

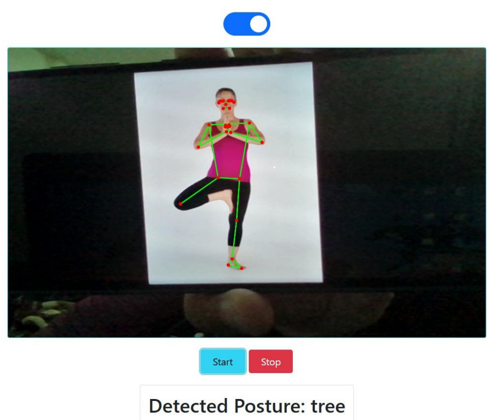

FIGURE 10. Tree

CONCLUSION In order to detect yoga postures in real time, this research offers a novel deep learning-based approach that achieves high classification accuracy across five important postures. The system is very efficient for real-time applications since it integrates deep neural networks for classification with Mediapipe for pose prediction. Future improvements will involve extending the range of postures that are recognised

TABLE 1. Performance Metrics for Yoga Poses

Yoga Pose	Precision	Recall	F1-Score
Downdog	0.98	0.96	0.97
Goddess	0.96	0.91	0.94
Plank	0.93	0.96	0.94
Tree	0.84	0.82	0.83
Warrior	0.82	0.87	0.85

and including audio feedback to improve the user experience.

REFERENCES

1) Josvin Jose and S Shailesh 2021 IOP Conf. Ser.: Mater. Sci. Eng. 1110 012002
2) A. Sharma, Y. Shah, Y. Agrawal and P. Jain, "iYogacare: Real-Time Yoga Recognition and Self-Correction for Smart Healthcare," in IEEE Consumer Electronics Magazine, vol. 12, no. 4, pp. 47-52, 1 July 2023, doi: 10.1109/MCE.2022.3171054
3) Garg, S., Saxena, A. & Gupta, R. Yoga pose classification: a CNN and MediaPipe inspired deep learning approach for real-world application. J Ambient Intell Human Comput 14, 16551–16562 (2023)
4) Upadhyay, Aman, Niha Kamal Basha, and Balasundaram Ananthakrishnan. "Deep learning-based yoga posture recognition using the Y_PN-MSSD model for yoga practitioners." Healthcare. Vol. 11. No. 4. MDPI, 2023.
5) Saurav, S., Gidde, P. & Singh, S. Exploration of deep learning architectures for real-time yoga pose recognition. Multimed Tools Appl 83, 81621–81663 (2024).
6) A Sharma, Y Shah, Y Agrawal, P Jain - arXiv preprint arXiv:2201.07594, 2022 - arxiv.org
7) K. Elissa, "An overview of decision theory," unpublished. (Unpublished manuscript)
8) Jiang, S.; Lv, B.; Guo, W.; Zhang, C.; Wang, H.; Sheng, X.; Shull, P.B. Feasibility of Wrist-Worn, Real-Time Hand, and Surface Gesture Recognition via sEMG and IMU Sensing. IEEE Trans. Ind. Inform. 2018, 14, 3376–3385
9) Sakshi, Saini, S. Yoga with Deep Learning: Linking Mind and Machine. SN COMPUT. SCI. 5, 427 (2024).
10) Huang, X.; Wang, F.; Zhang, J.; Hu, Z.; Jin, J. A Posture Recognition Method Based on Indoor Positioning Technology. Sensors 2019, 19, 1464 . Kumar Rajendran and S. Chakkaravarthy Sethuraman, "YogiCombineDeep: Enhanced Yogic Posture Classification Using Combined Deep Fusion of VGG16 and VGG19 Features," in IEEE Access, vol. 12, pp. 139165-139180, 2024, doi: 10.1109/ACCESS.2024.3414654
11) Hachaj, T.; Ogiela, M.R. Rule-based approach to recognizing human body poses and gestures in real time. Multimed. Syst. 2014, 20, 81–99.
12) A. K. Rajendran and S. C. Sethuraman, "A Survey on Yogic Posture Recognition," in IEEE Access, vol. 11, pp. 11183-11223, 2023, doi: 10.1109/ACCESS.2023.3240769
13) S. Madan, J. Sembhi, N. Khurana, K. Makkar and P. Byati, "Yoga for preventive health: A holistic approach", Amer. J. Lifestyle Med., pp. 1-6, Jan. 2022.
14) J. Palanimeera (1Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamil Nadu, India) and K. Ponmozhi (1Department of Computer Applications, Kalasalingam Academy of Research and Education, Krishnankoil, Tamil Nadu, India)
15) Jain, S., Rustagi, A., Saurav, S. et al. Three-dimensional CNN-inspired deep learning architecture for Yoga pose recognition in the real-world environment. Neural Comput & Applic 33, 6427–6441 (2021)
16) Gayakwad, Milind, et al. "Web Service Implementation" AI Yoga Trainer Using Human Pose Estimation" for Web Discovery." Journal of Electrical Systems 20.2 (2024): 2167-2176.
17) Long, C., Jo, E.& Nam, Y. Development of a yoga posture coaching system using an interactive display based on transfer learning. J Supercomput 78, 5269–5284 (2022).
18) Kishore, D Mohan; Bindu, S1; Manjunath, Nandi Krishnamurthy. Smart Yoga Instructor for Guiding and Correcting Yoga Postures in Real Time. International Journal of Yoga 15(3):p 254-261, Sep–Dec 2022.
19) A. Gupta and A. Jangid, "Yoga Pose Detection and Validation," 2021 International Symposium of Asian Control Association on Intelligent Robotics and Industrial Automation (IRIA), Goa, India, 2021
20) Sunney, Jothika (2022) Real-Time Yoga Pose Detection using Machine Learning Algorithm.

Masters thesis, Dublin, National College of Ireland.
21) A. Gupta and H. P. Gupta, ”YogaHelp: Leveraging Motion Sensors for Learning Correct Execution of Yoga With Feedback,” in IEEE Transactions on Artificial Intelligence, vol. 2, no. 4, pp. 362-371, Aug. 2021