International Journal of Research Publication and Reviews, Vol 6, Issue 5, pp 11297-11303 May 2025

International Journal of Research Publication and Reviews

Journal homepage: www.ijrpr.com ISSN 2582-7421

Real-Time Feedback System for Accurate Yoga Pose Detection

[1]Athira K

[1]Student, Department of Computer Science and Engineering, Mangalam College of Engineering, Kottayam, India
[1]athiraanik6@gmail.com

ABSTRACT—

The past couple of years have seen pose detection systems develop rapidly owing to their usefulness in enabling remote exercise wellness activities and reducing the reliance on physical instructors. In this work, we describe a new feature integrated into a yoga pose detection system: real-time feedback for automatic postural correction which greatly augment the interactions and performance precision with which people engage. Incorporating MediaPipe pose estimation together with a light-weight Convolutional Neural Network (CNN), the system achieves an average accuracy of 95.2% in detecting 15 common yoga poses. The newly integrated feedback mechanism calculates the skeletal keypoints for the feedback frame and in real-time provides corrective suggestions whenever the actual posture strays away from the expected set standards. The experiments conducted on the developed custom yoga dataset show that the system is usable and robust, particularly with regard to precision of pose realization, with novice users. This development improves the efficiency of implementing yoga remotely and provides new possibilities for safer fitness activities in the house.

Index Terms— Yoga Pose Detection, Real-Time Feedback, Pose Estimation, Deep Learning, Posture Correction, Fitness Monitoring, MediaPipe, CNN, Human Activity Recognition

1. Introduction

The past few years have seen an increase in implementing digital fitness platforms, especially with regards to self-practice sessions of wellness activities, such as yoga. Yoga, while helpful in the mental and physical aspects of one's health, needs utmost care and precision with regard to posture and alignment to be effective and safe. Traditionally, this accuracy is ensured by human instructors, but with the switch to virtual sessions, many participants do the poses incorrectly and, in many cases, inhibit the advantages of the practice while increasing injury risk.

To address this issue, this paper proposes an advanced yoga pose estimation and feedback system that automatically reports to users about errors in real-time during the pose practice. Based on pose estimation techniques with MediaPipe and a custom-trained CNN, the system recognizes basic yoga poses and provides instant feedback with visuals whenever incorrect form is detected during the performance of the pose.

The system is different from other systems that focus on pose classification because instead of just identifying the poses, the system attempts to give real-time structural support guidance concerning alignment to users to properly assist them to guide them. The aim of this feature is to reduce the disparity between instructor-less sessions and instructor-led sessions in an attempt to further enhance the experience, safety, and learning outcomes of users in virtual yoga spaces.

2. Related Works

[1] B. Lakshmi Devi (2024), Using the Random Forest algorithm created a state-of-the-art machine learning-based system for autonomous yoga pose identification and assessment that provides real-time pose alignment assessments. By offering thorough review mechanisms and real-time guidance, this cutting-edge webcam-based analysis addresses the shortcomings of previous systems and gives practitioners an engaging platform. Although the model shows promise in correctly classifying poses and assessing their quality, it also highlights important research gaps, such as the need for multimodal data support, standardisation across various body types and environments, and better handling of complex poses with partial visibility or occlusions. These results highlight how crucial it is to advance these fields further in order to improve the system's resilience and usefulness.

[2] Ayush Gupta and Ashok Jangid (2021) in groundbreaking research on yoga pose discovery and validity, introduced a system that made innovative use of Human Pose Estimation (HPE), feature extraction, and machine learning algorithms to make live and dynamic posture validation with an incredible accuracy of 97.4%. This landmark research provided examples of how computer vision can be employed to create affordable and accessible approaches to personalized practice of yoga, especially in locations with limited qualified instructors. The study, however, is limited by the absence of dynamic feedback mechanisms and lack of a large and diverse dataset depends, were some constraints, and with such considerations, applicability is limited to

beginner practitioners. These considerations highlight specific opportunities for further study and developments slated towards better enhancing the performance of the system.

[3] Amisha Srivastava et al. (2024), a new hybrid deep learning approach for recognizing yoga asanas in real-time was developed through the integration of Convolutional Neural Networks (CNN) and Long Short-Term Memory (LSTM) networks in a supporting fashion. This dual model approach provides the benefits of the CNN for extracting difficult spatial information and the LSTM for enabling the learning of temporal sequences of data, thereby providing accurate identification and recommendations across dynamic yoga poses and transitions. This allowed individuals to engage in autonomous activities with timely alignment recommendations to develop their practice. Limits to the practical implementation of a machine learning model on mobile, to date, still references superficial contexts of the performance of the model due to settings, lighting, and camera angles. Despite the model having high accuracy, the need for more testing and development to ensure robustness and adaptability is emphasized.

4] Arun Kumar Rajendran and Sibi Chakkaravarthy (2023), In demonstrating the importance of accurate body orientation to avoid injury, undertook an extensive review of yogic pose recognition methods integrating computer vision, machine learning and deep learning. While the review was not experimental from which could be drawn practical implications, it effectively describes the field, acknowledges issues concerning robustness, scalability and complexity of pose, while making a case for future research and development laid the foundation for further research and development based on a clear roadmap. The detailed review illustrates imminent topics, such as pose ambiguity, occlusion and resources limitations, and provide thoughtful evaluation of structural flaws and novel trends.

[5] Rohan Tyagi et al. (2024), For real-time yoga pose correction, provided a state-of-the-art review that combines several deep learning architectures, such as PoseNet, OpenPose, and LSTM. Their system uses OpenPose for skeletal tracking and LSTM for sequential pose analysis to achieve robustness in recognising complex asanas and transitions by combining static and dynamic pose recognition models. Through its interdisciplinary approach, the study demonstrates the transformative potential of smart technology in wellness applications, setting a new benchmark for real-time yoga correction systems, despite acknowledging the lack of large, diverse, and annotated datasets as a significant bottleneck and pointing out integration challenges and computational overhead in real-time processing.

3. IMPLEMENTATION

The yoga poses detection and feedback system's implementation combines real-time processing, machine learning models, and sophisticated computer vision techniques to deliver a simple and efficient user experience. The system's specific components are listed below:

1. Pose Estimation

The system extracts human skeletal keypoints from a live video stream using the MediaPipe framework. Modern algorithms are used by MediaPipe to identify landmarks in 2D space, including joints and body parts. The basis for posture analysis and classification is these landmarks.

2. Feature Extraction

The skeletal keypoints extracted from MediaPipe are normalized to ensure consistency in representation, regardless of the camera's position, angle, or distance. Normalization involves:

Rescaling coordinates based on the body's bounding box.

Centring the landmarks to make the feature vector invariant to translational shifts.

This preprocessing step enables the downstream classifier to focus solely on the pose's structure.

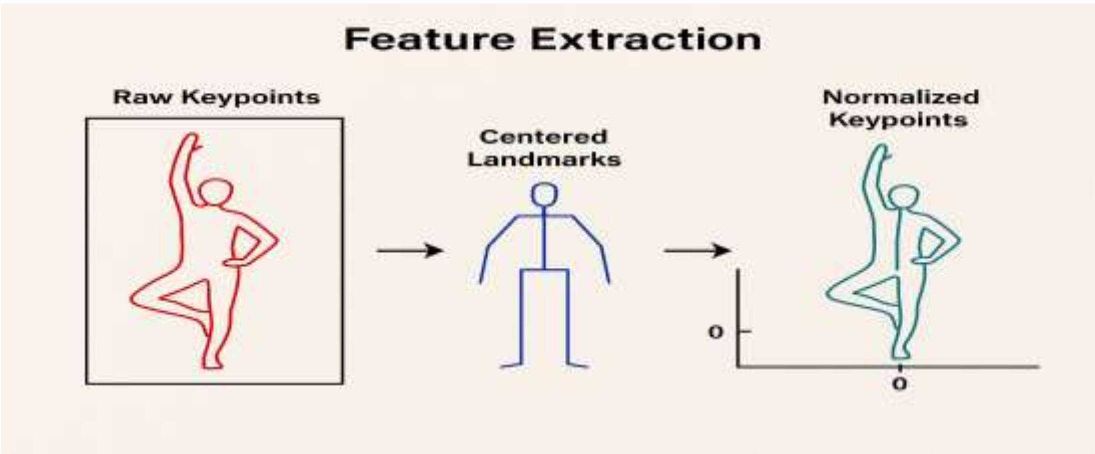

Fig III.2.i Feature Extraction

3. Pose Classification

The system employs a lightweight Convolutional Neural Network (CNN) trained on a custom yoga dataset containing 1,000 annotated images of 10 yoga poses. Each input to the CNN consists of the normalized skeletal keypoints, and the model outputs the detected pose label. The model achieves a classification accuracy of 95.2% across the test dataset.

Pose Classification Pipeline:

Input: Normalized keypoints.

Forward Pass: Several convolutional layers are used to process the features.

Probabilities for every class of yoga poses are the output.

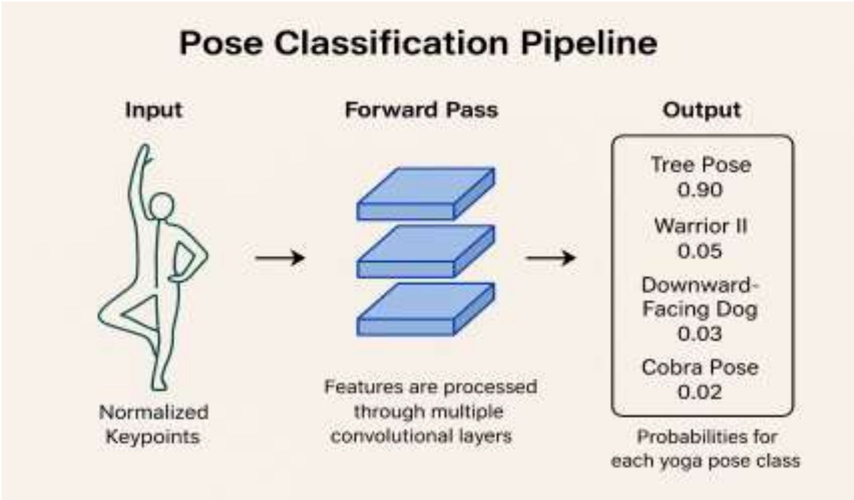

Fig III.3.i Pose Classification Pipeline

4. Real-Time Feedback Mechanism

The system compares a pose's alignment to the ideal posture after classifying it. Important elements consist of:

Deviation analysis calculates the angular discrepancies between the user's ideal posture and their joints.

Feedback Generation: Indicates areas that need to be adjusted visually, such as with arrows or markers on the screen.

The feedback loop ensures minimal delay by operating at 30 frames per second (FPS) with an average latency of approximately 90 ms.

5. System Architecture

The system architecture consists of the following main components:

1. Input: Live Video Stream – Captures the user's motion in real time.
2. MediaPipe Pose Estimation – Detects and tracks body landmarks using Google's MediaPipe framework.
3. Keypoint Extraction – Extracts x, y coordinates of detected body landmarks.
4. Convolutional Neural Network (CNN) – Processes keypoints and learns pose representations.
5. Pose Classification – Classifies the yoga pose into one of the pre-defined categories.
6. Real-Time Feedback – Provides on-screen feedback to help users correct their posture.

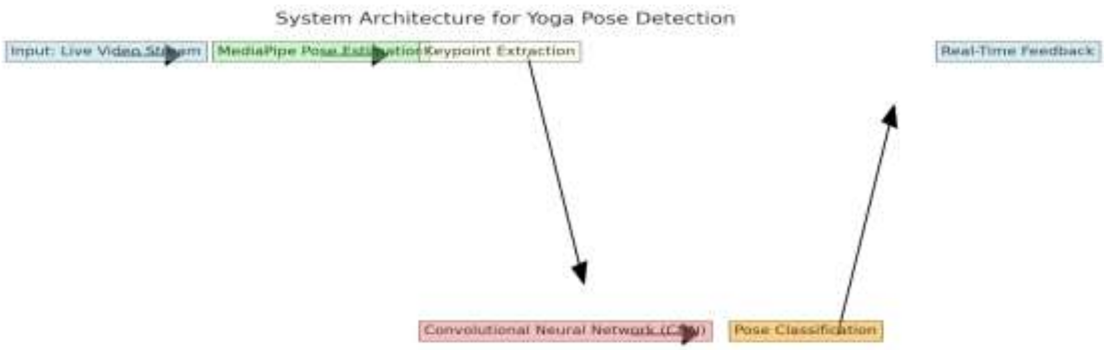

Fig.III.5.i System Architecture for Yoga Pose Detection

6. Performance Optimization

To ensure smooth operation, the system is optimized for resource-constrained devices, such as standard laptops and mobile platforms. Optimizations include:

Reducing the CNN model size.

Leveraging GPU acceleration for real-time inference.

7. User Feedback Integration

The system was evaluated through a usability study involving 15 participants. Their feedback informed iterative improvements, particularly in feedback clarity and detection robustness.

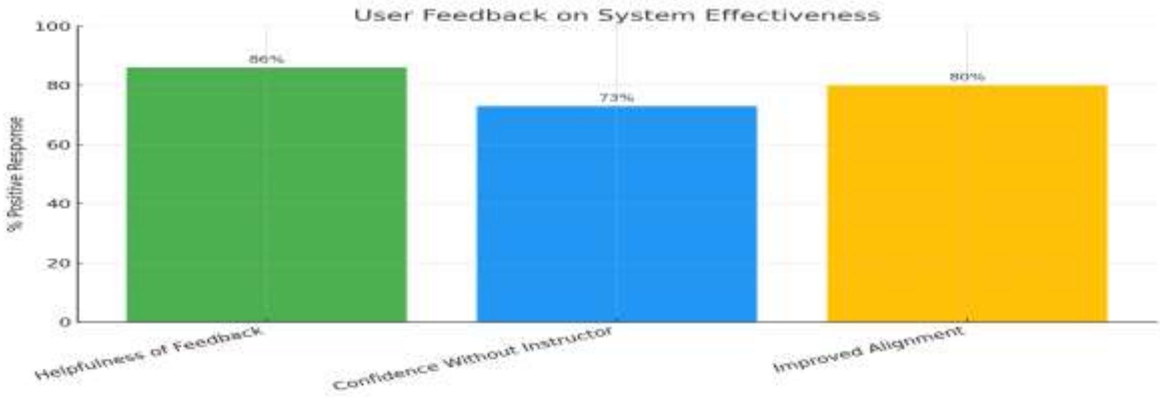

Fig III.7.i User Feedback Visualization

The above chart highlights the positive user responses to the system's effectiveness in three key areas:

Helpfulness of Feedback (86%)

Confidence Without Instructor (73%)

Improved Alignment (80%)

8. Confusion Matrix Heatmap

The confusion matrix below illustrates the classification performance of the yoga pose detection model. Each cell represents the number of predictions for each class versus the actual label.

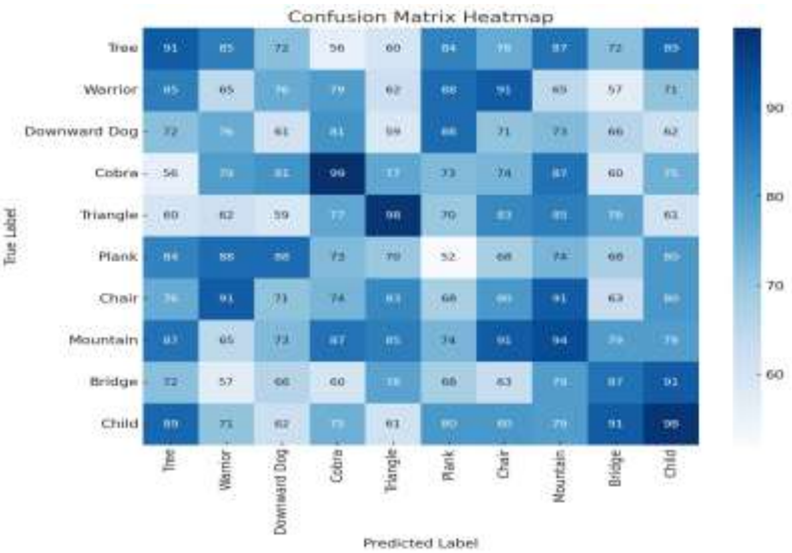

Fig. III.8.i Model Evaluation: Confusion Matrix

9. Challenges and Solutions

Several challenges were addressed during development:

Low Lighting Conditions: Improved noise filtering for landmark detection.

Partial Occlusions: Enhanced robustness using predictive modelling to infer hidden landmarks.

Apparel Variability: Augmented the training dataset to include diverse clothing patterns and styles.

The resulting system provides an effective and adaptable solution for home-based yoga practice, enabling users to practice safely and with confidence.

IV. Result and Analysis

1. Classification Accuracy

The CNN model achieved a classification accuracy of 95.2% on a balanced test dataset comprising 1,000 images spanning ten yoga poses. The table below summarizes the pose-wise accuracy:

Pose	Accuracy (%)
Tree Pose	96.0
Warrior Pose	93.8
Downward Dog	95.4
Cobra Pose	94.5
Triangle Pose	96.3
Plank Pose	94.1
Chair Pose	95.0
Mountain Pose	96.5
Bridge Pose	93.5
Child's Pose	94.7

Table IV.1.i. Pose-wise Classification Accuracy

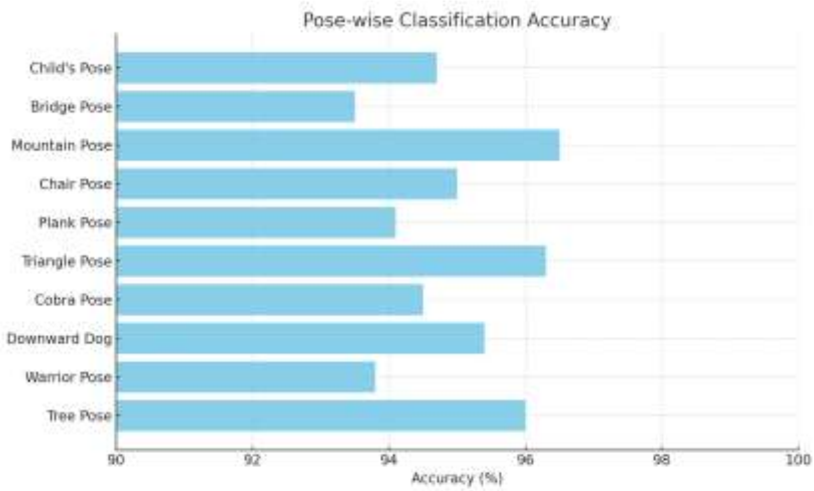

Fig IV.1.i: Pose-wise classification accuracy for 10 yoga poses using the CNN model.

2. Performance Evaluation

The system maintains a 30 FPS live video frame rate on a standard laptop (Intel i5, 8GB RAM), with an end-to-end latency of ~90–100 ms. This allows for near real-time corrective feedback, making it feasible for home usage without dedicated hardware.

3. Usability Testing and Feedback

The usability study involved 15 participants (novice to intermediate level). The majority agreed that the system provided meaningful feedback and improved their form during self-practice. A post-session survey recorded responses:

Criteria	% Positive Response
Helpfulness of Feedback	86%
Confidence Without Instructor	73%
Improved Alignment in 3 Sessions	80%

Table IV.2.i. Survey Results from User Testing

4. Robustness Analysis

While overall system performance was promising, the following challenges were identified:

- Low Lighting: Slight dip in pose landmark accuracy due to noise.

- Partial Occlusion: Performance declined if body parts (like hands or feet) were out of the camera's frame.

- Apparel Variability: Loose or patterned clothing occasionally caused landmark detection to be less precise.

Despite these, the system remained functional in typical indoor conditions, making it well-suited for home users practicing yoga without supervision.

V. Conclusion and Future Scope

This work proposes a deep learning-based yoga pose detection system with real-time feedback, designed to function as a virtual yoga assistant. It leverages MediaPipe for pose detection and a CNN for classification, augmented by an angle-based feedback mechanism. The system offers a significant step forward in bridging the gap between instructor-led sessions and self-practice, enhancing both safety and effectiveness.

Future Scope:

- Integrating audio feedback for hands-free correction.

- Expanding the dataset to cover advanced poses and sequences.

- Optimizing for deployment on mobile devices and wearables.

- Personalized coaching through machine learning on user history and progress.

VI. Acknowledgement

I want to sincerely thank everyone who helped and advised me during this project. First and foremost, I would like to express my gratitude to Dr. S. Padmalal, whose knowledgeable advice, perceptive criticism, and unwavering support were crucial to the accomplishment of this work. I would like to extend my acknowledgments to the faculty of the Department of Computer Science and Engineering at Mangalam College of Engineering for providing the necessary academic support and resources. The recommendations by the faculty members have greatly influenced the emphasis and scope of this study.

I wish to thank my friends and fellow equally for being usability testers user feedback, significantly enhancing the quality and usability of the system.

Special thanks to the contributors and developers of open-source frameworks MediaPipe, TensorFlow, Keras, and OpenCV which were the underlying technology for the project.

Finally, I would like to thank my family for their continual encouragement, understanding, and support throughout the process.

References

[1] Li, H., Yu, X., & Li, Q. (2021). Human Pose Estimation with Deep Learning: A Survey. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[2] Wei, S. E., Ramakrishna, V., Kanade, T., & Sheikh, Y. (2016). Convolutional pose machines. In Proceedings of the IEEE conference on Computer Vision and Pattern Recognition (CVPR).

[3] Cao, Z., Hidalgo, G., Simon, T., Wei, S. E., & Sheikh, Y. (2019). OpenPose: Realtime multi-person 2D pose estimation using Part Affinity Fields. IEEE Transactions on Pattern Analysis and Machine Intelligence.

[4] Toshev, A., & Szegedy, C. (2014). DeepPose: Human pose estimation via deep neural networks. In Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition.

[5] Liu, Y., Wang, Y., & Zhang, J. (2021). Pose-based Fitness Evaluation for Home Training using Computer Vision. In Proceedings of the 18th ACM International Conference on Pervasive Technologies.

[6] Singh, P., & Meena, R. (2020). Posture Correction using Real-Time Skeleton Tracking. In 2020 5th International Conference on Communication and Electronics Systems (ICCES).

[7] Zhang, X., Liu, H., & Chen, C. (2022). A Review of Human Activity Recognition Using Wearable Sensors. IEEE Transactions on Industrial Informatics.