A Web-Based System for Real-Time Human Pose Estimation Using Live Video Stream and Keypoint Analysis

Vaishnavi Bhat[1], Shivani Shet[2], Faiyaz M[3], Chinmay Kulkarni[4], Dr. Pushpalatha S. Nikkam[5], Prof. Varsha S Jadhav[6], Ravikiran Reddy[7]

[1,2,3,4,5,6] *Department of Information Science and Engineering, SDM college of Engineering and Technology Dharwad, Karnataka*

[7] *Sr. Technical Product Manager, Amazon 410 Terry Avenue North Seattle, Washington – 98109*

Abstract- **Human pose estimation (HPE) plays an essential role in applications such as fitness tracking, yoga guidance, and rehabilitation. This paper presents a web-based system that uses real-time video input from a webcam to detect and evaluate human poses during workout and yoga sessions. The system identifies key body landmarks using pose estimation algorithms and analyzes posture accuracy while automatically counting exercise repetitions. Users interact with the platform through a simple browser interface, receiving instant visual feedback on their movements. Built using MediaPipe, OpenCV, and Streamlit, the application is designed for ease of use, remote accessibility, and scalability in personal fitness monitoring.**

Keywords: **Human Pose Estimation, Real-Time Video, Exercise Monitoring, Yoga Posture Detection, Repetition Counter, Web Application.**

I. INTRODUCTION

Human Pose Estimation (HPE) is a rapidly evolving field within computer vision that focuses on detecting and analysing key body joints to understand human posture and movement. It plays a vital role in applications such as fitness training, yoga, physical rehabilitation, and interactive gaming. Traditional methods for posture correction and exercise monitoring rely on in-person supervision from trainers or therapists, which may not be accessible to all users and can lack scalability or consistency.

Recent advancements in deep learning and real-time video processing have made it possible to perform accurate pose estimation using only standard webcams, removing the need for specialized sensors or wearable devices. Frameworks like MediaPipe and OpenCV have enabled robust detection of body keypoints, facilitating applications in posture tracking and movement analysis.

In this work, we present a web-based system for real-time human pose estimation aimed at supporting users during workout and yoga sessions. The platform captures live webcam input, processes it to extract pose landmarks, and evaluates the user's posture using pose analysis techniques. The system also includes an automated repetition counter and provides immediate feedback to users. With a secure login system and responsive interface built using Streamlit, this platform offers a scalable and accessible solution for remote fitness monitoring and form correction.

II. METHODOLOGY

The proposed Human Pose Estimation System integrates real-time pose detection through webcam input with feedback for yoga and exercise poses, providing users with the ability to track progress, adjust form, and receive personalized feedback through a unified web platform.

The process begins when the user selects a pose from either the yoga or exercise categories. Upon selecting a pose, the webcam is activated to capture the user's live video feed. The system preprocesses this video using OpenCV to convert the frames from BGR to RGB format and normalizes them for consistency across different devices, ensuring accurate pose analysis.

The video feed is then passed to MediaPipe, which detects keypoints on the user's body, such as shoulders, elbows, knees, and wrists. From these keypoints, the system extracts pose landmarks that define the user's body position. These landmarks are

used to assess the user's posture and determine whether the pose is being performed correctly.
To evaluate the pose's accuracy, the system calculates joint angles based on the positions of the detected landmarks. If any misalignments are detected, real-time feedback is provided to the user, such as suggestions to “straighten your back” or “lower your knees.” Additionally, for exercise poses, the system tracks repetitions, recognizing movement patterns like squats or push-ups and counting the number of repetitions completed. At the end of the session, users have the option to download the video of their session for personal review, allowing them to track progress and make improvements based on the recorded footage.

The system is designed to be accessible and user-friendly, making it suitable for users of all fitness levels—from beginners to advanced practitioners. By providing real-time guidance and performance analytics, the platform promotes consistent improvement and helps prevent common injuries caused by incorrect posture. It offers clear instructions and live feedback, which makes it easier for users to understand and correct their movements. The easy-to-use interface ensures that users can navigate through exercises without any confusion. Whether someone is starting their fitness journey or looking to perfect their form, the system supports them with personalized tips, repetition tracking, and progress monitoring, all within a smooth and interactive web experience.

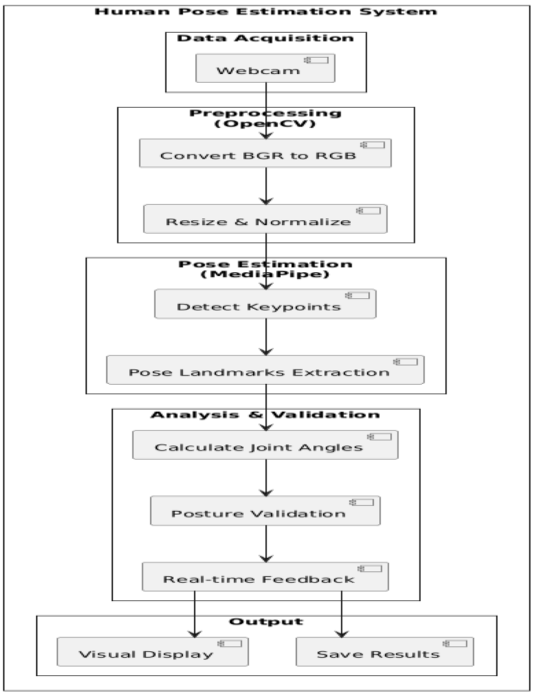

Figure 2 Architectural Diagram of HPE

III. ALGORITHMS USED

Two key technologies and algorithms were implemented in this system to perform real-time human pose estimation and provide interactive feedback for exercises, including yoga poses. The system combines MediaPipe Pose Estimation, which uses deep learning models based on Convolutional Neural Networks (CNNs), with OpenCV for video processing. These technologies are essential for tracking the user's movements and analyzing their exercise performance based on body keypoints.

MediaPipe Pose Estimation employs CNNs to accurately extract 33 key body landmarks (such as shoulders, elbows, knees, and hips) from the webcam feed, enabling real-time analysis of user poses. CNNs are particularly effective for image-based tasks like pose recognition due to their ability to capture spatial features and patterns from video frames. OpenCV complements this by capturing and processing video input, ensuring smooth and responsive pose tracking.

Additionally, for user authentication, SQLite is employed to securely store user credentials and session data, while SHA-256 is used for password hashing, ensuring that sensitive user information remains protected. These technologies work together to provide a secure, interactive, and efficient application for users to track and improve their exercise and fitness routines.

IV. TECHNOLOGIES USED

The proposed Human Pose Estimation System leverages a combination of web technologies and machine learning frameworks to provide an interactive and efficient user experience. The frontend is built using Streamlit, with custom HTML, CSS, and JavaScript to create an intuitive, responsive interface. Interactive features like user login, exercise selection, and pose estimation are easily accessible. CSS is used to provide a modern and polished design, ensuring a visually appealing interface for users. The user interface seamlessly integrates with backend functionalities, ensuring that users can interact with the pose estimation and feedback features smoothly.

On the backend, Python serves as the primary language, utilizing the Streamlit framework for the web interface. The application's backend relies heavily on SQLite for secure user authentication, where passwords are hashed using SHA-256 for security. For real-time pose estimation, the app integrates OpenCV and MediaPipe to detect and track human body poses through video input. The app provides feedback on exercises and tracks repetitions. The backend also handles data processing for exercise tracking and stores user data efficiently. User actions, like login, registration, and pose feedback, are processed through the Flask framework, which connects with the machine learning models and the user interface to ensure smooth functionality.

The system is designed to be platform-independent, running on any modern web browser. During the development process, it was thoroughly tested to ensure accuracy and reliability across all modules, from user registration and login to real-time pose detection and feedback. The pose detection functionality was evaluated and fine-tuned using cross-validation on real user data to ensure precise feedback and exercise tracking. Additionally, the entire pipeline—ranging from user input to feedback display—was integration-tested to ensure smooth and consistent performance on various devices. The modular nature of the system allows for easy updates and scalability, enabling future enhancements such as adding more exercises or improving the feedback algorithms.

V. RESULTS AND DISCUSSIONS

The proposed real-time fitness monitoring system was evaluated using webcam-based input, allowing users to perform exercises in front of their device's camera. Upon selecting a specific workout or yoga pose, the system activates the webcam and utilizes MediaPipe's pose estimation model to extract 33 key body landmarks. These landmarks capture the spatial orientation of joints such as shoulders, elbows, knees, and hips in each video frame. Once the keypoints are extracted, the system calculates joint angles and body alignment to assess whether the user is performing the pose correctly. Based on predefined thresholds and angle ranges, it provides real-time guidance, such as "Correct Posture" or "Raise your hands," to help users improve their form. Figure 5.1 illustrates this functionality during a yoga session, where the model overlays pose skeletons and corrective feedback directly onto the webcam feed to guide the user into accurate alignment.

The system also integrates a repetition counter, which increments each time a complete and correct movement cycle—such as an arm raise or full squat—is detected. This functionality relies on tracking key joint motions over time to identify exercise-specific patterns. As shown in Figure 5.2, after selecting an exercise from the list, the user interface loads the detection screen where pose tracking begins, providing a seamless transition into workout mode. This intuitive interaction encourages continued usage and simplifies session initiation.

To support user engagement and accountability, the interface offers visual displays of repetition counts, feedback messages, and control buttons like "Stop Webcam" and "Reset Counter." Figure 5.3 captures a live workout session where the system actively tracks repetitions, displays pose lines, and issues real-time cues such as "RAISE YOUR HANDS," ensuring accurate execution. These visual prompts, combined with progress indicators, enhance the overall effectiveness of the workout experience.

The system's responsiveness across varying lighting conditions and body types confirms its robustness and adaptability. Its combination of pose estimation, motion analysis, and real-time feedback delivers a comprehensive virtual fitness assistant. Compared to traditional mirror-based workouts, this system offers a smarter, data-driven approach that empowers users to improve their technique, reduce injury risk, and stay motivated over time.

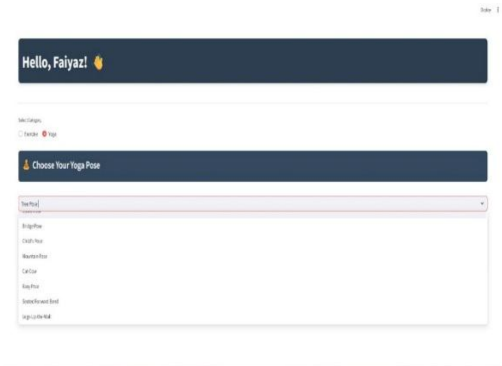

Figure 5.1 Output of yoga pose detection

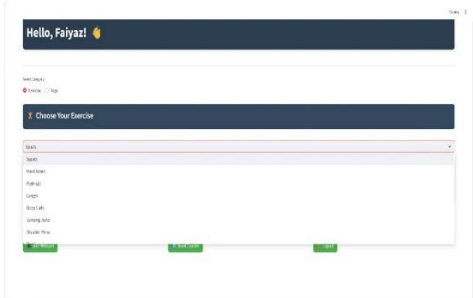

Figure 5.2 Exercise selection

Figure 5.3 Output of exercise detection with repetition counter

VI. CONCLUSION

This study introduces a web-based system that utilizes real-time human pose estimation and exercise feedback through the integration of MediaPipe and OpenCV. The application effectively tracks key body landmarks to analyze user movements during yoga and fitness routines, helping users maintain proper form and improve their workout experience. Built with the Streamlit framework, the system offers an intuitive and interactive interface that allows users to select exercises, receive real-time feedback, and monitor their progress with ease.

By leveraging pose estimation technology and machine learning, the system enhances the accuracy of movement tracking and sets a foundation for technology-driven physical wellness solutions. Its ability to analyze and correct movements in real time presents valuable applications not only in fitness but also in rehabilitation and injury prevention. This approach highlights the powerful role of computer vision in creating engaging, effective, and health-oriented user experiences.

REFERENCES

[1] D. Rai, Anjali, A. Kumar and A. Baghel, "Pose Detection Using OpenCV and Media Pipe," *2024 International Conference on Integrated Circuits, Communication, and Computing Systems (ICIC3S)*, Una, India, 2024

[2] H. Dhakate, S. Anasane, S. Shah, R. Thakare and S. G. Rawat, "Enhancing Yoga Practice: Real-time Pose Analysis and Personalized Feedback," *2024 International Conference on Emerging Systems and Intelligent Computing (ESIC)*, Bhubaneswar, India, 2024

[3] K. Aarthy and A. A. Nithys, "Yoga Pose Detection and Identification Using MediaPipe and OpenPose Model," *2023 International Conference on Computer Science and Emerging Technologies (CSET)*, Bangalore, India, 2023,

[4] "Gesture Recognition for Media Interaction: A Streamlit Implementation with OpenCV and MediaPipe.", Vaibhav Patil , Dr. Sanjay Sutar, Sanskruti Ghadage , Shubham Palkar, IJRASET, Volume 11 Issue IX Sep 2023.

[5] M. Rane, S. Mohole, Y. Pahade, G. Kurandale and A. Munjale, "SmartPosture: Real-Time Posture Detection and Feedback System Integrating IoT and Machine Learning," *2025 1st International Conference on AIML-Applications for Engineering & Technology (ICAET)*, Pune, India, 2025