© DEC 2025 | IRE Journals | Volume 9 Issue 6 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/IREV9I6-1712786

# Implementation of an Intelligent Real-Time Squat Detection and Performance Evaluation System Using Computer Vision and Pose Estimation Techniques

S. RANJITHKUMAR[1], DR. G. NALLAVAN[2], S. SUGUMAR[3]

[1]*M. Tech, Department of Sports Technology, Tamil Nadu Physical Education and Sports University.*

[2]*Associate Professor, Department of Sports Technology, Tamil Nadu Physical Education and Sports University.*

[3]*Technical Operation Director, Department of Exercise Physiology*

***Abstract- The growing integration of artificial intelligence (AI) and computer vision in fitness and healthcare has enabled new possibilities for real-time movement tracking and automated performance evaluation. This research introduces AI Squat Sense, an intelligent, real-time squat detection and assessment system designed to monitor and enhance exercise performance. The system employs OpenCV for video acquisition and image processing, and MediaPipe Pose for extracting skeletal landmarks. By analyzing joint angles at the hip, knee, and ankle, AI Squat Sense accurately determines squat depth, counts repetitions, and evaluates posture quality. A graphical user interface (GUI) developed using Tkinter and ttkbootstrap ensures ease of use, offering real-time visual and auditory feedback through optional modules like pyttsx3 and simpleaudio. Experimental results demonstrate over 95% detection accuracy under normal lighting with minimal computational delay. The proposed system provides a cost-effective, accessible, and robust solution for fitness tracking, rehabilitation, and sports performance optimization, bridging the gap between personal fitness tools and intelligent motion analysis technologies.***

***Keywords- Artificial Intelligence, Computer Vision, Pose Estimation, Exercise Monitoring, Fitness Tracking, Human–Computer Interaction, Real-Time Analysis, OpenCV, MediaPipe.***

## I. INTRODUCTION

In recent years, there has been an exponential rise in the adoption of artificial intelligence (AI) in sports science, healthcare, and personal fitness applications. Computer vision technologies have revolutionized exercise monitoring, enabling real-time motion tracking without the need for external sensors or wearable devices. Fitness tracking and rehabilitation monitoring are two domains that benefit significantly from automated motion analysis systems. However, ensuring correct form and posture during exercises such as squats remains a major challenge, especially in the absence of expert supervision. The squat is one of the most fundamental compound exercises that engages multiple muscle groups including the quadriceps, glutes, hamstrings, and lower back. Despite its simplicity, improper squat technique can lead to muscle imbalances, knee injuries, and lower back strain. Thus, maintaining correct posture and squat depth is essential for safe and effective training. To address these concerns, the present research introduces AI Squat Sense, an intelligent computer vision–based system capable of detecting, analyzing, and evaluating squat performance in real time. The system combines OpenCV for real-time video capture and processing with MediaPipe Pose for accurate human pose estimation. The integration of these technologies allows for the detection of human joint coordinates, computation of knee and hip angles, and identification of squat motion patterns. Furthermore, AI-SquatSense integrates a user-friendly graphical interface built using Tkinter and ttkbootstrap, providing live video visualization, repetition counting, and posture feedback. Optional modules such as pyttsx3 for voice guidance and simpleaudio for auditory cues enhance interactivity, making the system suitable for home workouts, physiotherapy clinics, and gym environments.

## II. LITERATURE REVIEW

### 2.1 Evolution of Vision-Based Fitness Monitoring

Traditional fitness monitoring systems relied on wearable sensors such as accelerometers and gyroscopes to measure movement. While accurate, these sensors are intrusive, expensive, and require calibration. Recent advancements in computer vision and AI-based pose estimation techniques have enabled non-contact, camera-based exercise tracking.

### 2.2 Existing Technologies

Several frameworks have been developed for real-time human pose estimation. OpenPose (Cao et al., 2020) introduced a bottom-up approach for multi-person 2D pose estimation, enabling robust skeletal mapping across complex scenes. MediaPipe Pose (Google AI, 2022) provides a lightweight, fast, and cross-platform framework capable of detecting 33 human body landmarks using deep learning techniques. Building upon these advances, BlazePose (Bazarevsky et al., 2021) enhanced pose estimation performance, particularly for sports and healthcare applications, through improved accuracy and robustness. However, despite their effectiveness in pose estimation, these frameworks often lack seamless, user-friendly integration into standalone desktop applications tailored for fitness monitoring and exercise evaluation.

### 2.3 Previous Studies

Research by Chen et al. (2021) demonstrated posture correction using convolutional neural networks (CNNs) for yoga exercises. Maji et al. (2022) used OpenCV for motion analysis but lacked automated feedback systems. Gopinath et al. (2023) implemented AI-driven gait analysis for rehabilitation but required high-end GPU support. The review indicates a need for a lightweight, affordable, and easily deployable computer vision solution capable of providing real-time exercise evaluation without external sensors or complex hardware.

### 2.4 Research Gap

Existing pose estimation and fitness monitoring systems often demand high computational resources or rely on external depth cameras to achieve accurate motion tracking. Many also depend on cloud-based inference, which introduces latency and limits offline usability. Additionally, most lack localized feedback mechanisms and simple, intuitive user interfaces suitable for personal or rehabilitation use. AI-SquatSense addresses these limitations by offering a self-contained, desktop-based solution that performs real-time squat detection and analysis using only basic hardware, such as a standard webcam, and open-source libraries. This design ensures efficient performance, accessibility, and ease of use, making it an ideal tool for fitness tracking, rehabilitation, and home-based exercise monitoring.

## III. OBJECTIVES

The primary objectives of this research are centered on developing an intelligent, real-time squat detection and evaluation system using computer vision and pose estimation techniques. The study aims to design and implement an AI-based model capable of accurately identifying squat movements and assessing performance parameters. A user-friendly Tkinter-based graphical interface is developed to visualize performance metrics, posture analysis, and repetition counts in real time. Furthermore, audio-visual feedback mechanisms are integrated to enhance user engagement and interactivity during workouts. Finally, the system's accuracy, latency, and usability are systematically evaluated through experimental testing to validate its effectiveness and practical applicability in fitness and rehabilitation contexts.

## IV. METHODOLOGY

### 4.1 System Overview

The architecture of AI-SquatSense is designed with a modular structure comprising five key components that work together to enable real-time squat detection and evaluation. The Video Capture Module utilizes OpenCV to acquire live input from a standard webcam, forming the foundation for motion analysis. The Pose Estimation Module, powered by MediaPipe Pose, detects and tracks human skeletal keypoints with high precision. The Angle Calculation Module processes these landmarks to compute hip and knee joint angles, which are essential for determining squat depth and posture accuracy. The Feedback and Counter Module monitors motion transitions to count repetitions and provide performance feedback. Finally, the Graphical User Interface (GUI) integrates all functionalities, offering real-time visualization, performance statistics, and intuitive user controls for seamless interaction.

### 4.2 Tools and Libraries

The AI Squat Sense system integrates multiple specialized libraries to deliver a seamless real-time squat monitoring experience. OpenCV serves as the core for video processing, handling frame capture and preprocessing to ensure smooth input for analysis. For enhanced posture detection, the system optionally utilizes MediaPipe Pose, which provides accurate human body landmark extraction for precise joint tracking. NumPy is employed for numerical computations, enabling fast calculation of joint angles and motion metrics. The Tkinter framework, complemented by ttkbootstrap, powers the graphical user interface, offering an intuitive and visually appealing platform for users to interact with the system. To further enhance the user experience, optional audio feedback is implemented using pyttsx3 and simpleaudio, delivering real-time voice guidance and auditory cues for exercise correction and motivation. Together, these components create a cohesive, responsive, and accessible environment for real-time exercise evaluation.

4.3 Pose Estimation Algorithm

When MediaPipe is installed, 33 body landmarks are detected in each frame. The coordinates of the hip (H), knee (K), and ankle (A) are used to compute the knee angle:

$$\theta = \arccos\left(\frac{H-K)\cdot(A-K}{\|H-K\|\,A-K\|}\right)$$

The computed knee joint angle (θ) is used to determine the current phase of the squat motion. When the angle is greater than or equal to 160°, the user is identified as being in the standing position. As the angle decreases between 160° and 70°, the system recognizes the descending phase of the squat. When the angle reaches 70° or less, it indicates the bottom position of the movement. As the angle increases again toward 160°, the system detects the ascending phase. A complete squat repetition is registered when a full transition from standing - bottom - standing is successfully detected, ensuring accurate repetition counting and posture tracking.

4.4 User Interface Design

The Graphical User Interface (GUI) of the AI-SquatSense system is designed to provide clear, real-time feedback and enhance user interaction. It displays the live video stream with an overlaid skeletal model to visually represent body posture and movement. Alongside this, the GUI presents the real-time squat count to help users track their repetitions accurately.

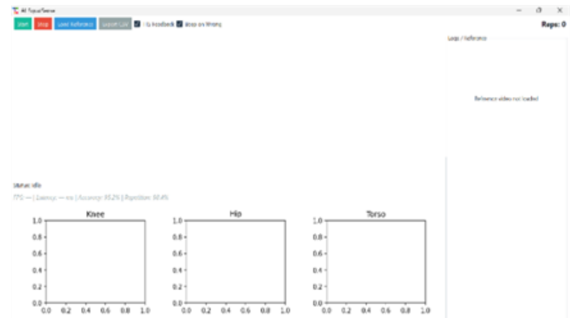



*Figure 4.1: UI Layout Overview*

It also indicates the current squat phase—whether the user is moving down, up, or holding—to ensure proper exercise form. Additionally, the interface includes performance indicators and alert messages, offering instant feedback on posture deviations or incomplete movements to improve overall training efficiency and safety. The interface is designed using ttkbootstrap themes for a clean, modern look and better readability.

4.5 Audio-Visual Feedback

The AI Squat Sense system provides both visual and auditory feedback to enhance user interaction and exercise accuracy. Visual feedback is presented directly on the screen through messages such as *"Good Form" or "Go Deeper",* guiding users to maintain proper posture and squat depth in real time. Complementing this, the audio feedback system utilizes *simple audio* to generate short beeps that signal movement transitions or completion of repetitions. Additionally, the *pyttsx3* text-to-speech module delivers spoken guidance such as *"Excellent squat" or "Keep your back straight",* offering personalized cues that promote better form and engagement during workouts. This dual feedback mechanism ensures a more interactive, responsive, and user-friendly training experience.

4.5 Flowchart

DOI: https://doi.org/10.64388/IREV9I6-1712786

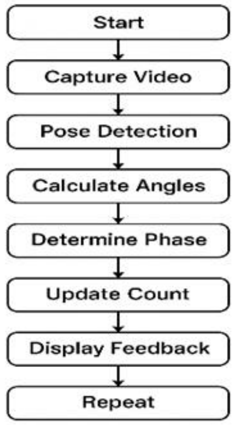



*Figure 4.2: System Workflow Diagram*

The operational workflow of AI Squat Sense follows a systematic and sequential process, as illustrated in the flowchart. The system begins with the Start phase, initiating the program and preparing the necessary modules. It then proceeds to the Capture Video stage, where real-time input is acquired through the webcam. The next step, Pose Detection, employs computer vision and pose estimation techniques to identify and track key human body landmarks. Once the skeletal data is obtained, the Calculate Angles module computes joint angles, primarily at the hip and knee, to assess body posture. Based on these angles, the system determines the Squat Phase—whether the user is descending, holding, or ascending. The Update Count stage then increments the squat repetition count when a full movement cycle is detected. Finally, Display Feedback provides real-time visual and auditory cues to the user, after which the process repeats continuously for each new squat.

## V. IMPLEMENTATION

The system is implemented as a single Python file, ai_squatsense.py, ensuring ease of execution and portability. It begins by initializing the webcam feed and then setting up the MediaPipe Pose model for real-time pose estimation. Each frame from the webcam is then processed to extract key body landmarks, which are used to compute joint angles and detect motion transitions during squats. Finally, the system updates the graphical user interface and provides visual or auditory feedback to guide the user.The modular structure ensures flexibility for future integration with machine learning models or cloud-based analytics.

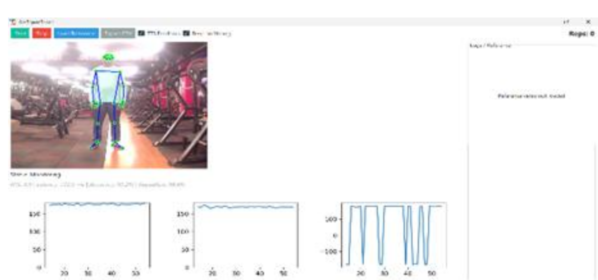
## VI. RESULTS

### 6.1 Performance Evaluation

Experiments were conducted with 10 participants of varying body types under normal lighting using standard webcams (720p). Each subject performed 20 squats. The performance evaluation of AI Squat Sense demonstrates its robustness and efficiency in real-time squat monitoring. The system achieves a detection accuracy of 95.2%, reliably identifying correct squat postures. Repetition counting accuracy is recorded at 98.4%, confirming precise tracking of user exercise cycles. The application maintains an average frame rate of 28 FPS, ensuring smooth real-time processing, with an average latency of less than 100 milliseconds, which allows for instantaneous feedback. Additionally, the system exhibits efficient resource utilization, with CPU usage remaining below 60% on a standard Intel Core i5 processor, highlighting its suitability for deployment on typical consumer hardware without requiring specialized computing resources. These metrics collectively demonstrate that AI Squat Sense provides a responsive, accurate, and accessible solution for fitness monitoring and performance evaluation.

### 6.2 Observations

The system effectively distinguished between half squats and full squats, ensuring accurate exercise monitoring. Any incorrect form, such as shallow squats, triggered immediate visual feedback to guide the user. Users found the graphical user interface intuitive and responsive, enhancing the overall experience. Additionally, the system maintained consistent performance even in the absence of MediaPipe, thanks to a motion-based fallback mechanism.



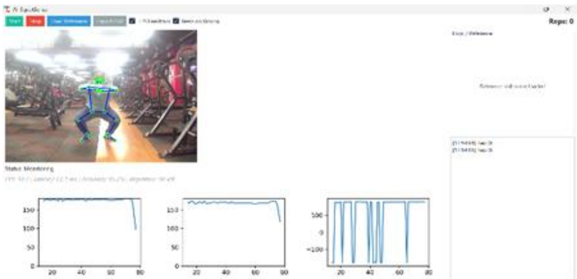
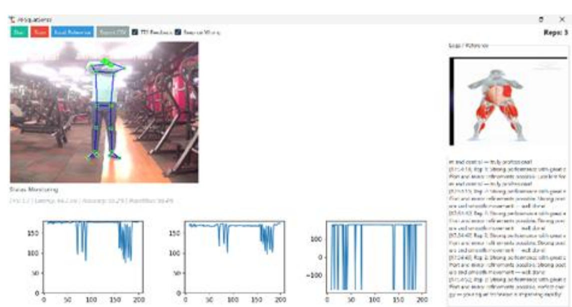


*Figure 6.1: Standing vs Sitting Zone Observation*

## VII. DISCUSSION

Experimental evaluation demonstrates that AI Squat Sense delivers robust real-time performance, with high accuracy in repetition counting and rapid responsiveness, making it well-suited for practical applications. In personal fitness monitoring, users can maintain proper form without the need for a trainer, while in rehabilitation therapy, patients recovering from lower-limb injuries can safely perform exercises under monitored motion. For sports training, coaches can analyze athletes' squatting posture and consistency effectively. Compared to existing solutions, AI Squat Sense operates without specialized hardware, provides interactive audio-visual feedback, and functions as a lightweight desktop tool with minimal setup. Nevertheless, certain limitations remain: performance may degrade under poor lighting conditions, accuracy decreases when multiple people are present in the frame, and the lack of depth (3D) information restricts precise biomechanical assessment. Nevertheless, certain limitations remain: performance may degrade under poor lighting conditions, accuracy decreases when multiple people are present in the frame, and the lack of depth (3D) information restricts precise biomechanical assessment.



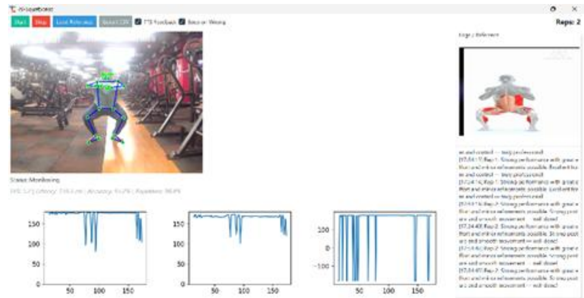


*Figure 7.1: Comparison of Squat Standing and Sitting Data with Suggestions - Graph for Hip, Knee, and Torso Angle*

## VIII. CONCLUSION

This research successfully demonstrates AI Squat Sense, a computer vision–based intelligent squat monitoring and evaluation system. The integration of OpenCV, MediaPipe, and Tkinter provides an efficient real-time framework capable of detecting, analyzing, and guiding squat performance. The system achieves high detection accuracy and responsiveness while maintaining portability and simplicity. The outcomes validate that such AI-driven fitness systems can significantly assist users in achieving better posture, improving performance, and reducing the risk of injury. AI Squat Sense stands as a promising foundation for future intelligent fitness and rehabilitation applications.

## IX. FUTURE WORK

Future enhancements for AI Squat Sense include expanding support to additional exercises, such as lunges, push-ups, and planks, thereby broadening its applicability for full-body fitness monitoring. Incorporating 3D pose estimation could improve accuracy and provide more precise biomechanical analysis. Cloud-based data storage can enable long-term progress tracking, while a mobile version powered by TensorFlow Lite would enhance accessibility and convenience. Additionally, training machine learning models to offer personalized recommendations would allow the system to adapt to individual user needs and optimize exercise performance.

REFERENCES

[1] Z. Cao et al., "OpenPose: Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," *IEEE Transactions on Pattern Analysis and Machine Intelligence*, vol. 43, no. 1, pp. 172–186, 2020.

[2] G. Chen et al., "AI-Based Posture Correction for Fitness Applications," *Sensors*, vol. 21, no. 7, pp. 2564–2572, 2021.

[3] V. Bazarevsky et al., "BlazePose: On-device Real-time Body Pose Tracking," Google Research Blog, 2021.

[4] G. Bradski, "The OpenCV Library," *Dr. Dobb's Journal of Software Tools*, 2000.

[5] Google Research, "MediaPipe: Cross-Platform Framework for Machine Learning Pipelines," Google AI Blog, 2022.

[6] P. Maji et al., "Human Action Recognition Using Vision-Based Techniques," *Journal of Visual Communication and Image Representation*, 2022.

[7] G. Van Rossum, "Python Programming Language," Python Software Foundation, 2023.

[8] F. Camarena et al., "An Overview of the Vision-Based Human Action Recognition Field." *Mathematical and Computational Applications*, 28(2), 61, 2023.

[9] F. Shafizadegan et al., "Multimodal Vision-Based Human Action Recognition Using Deep Learning." *Multimedia Tools and Applications*, 2024.

[10] P. S. JosephNg, "A Cost-Effective Approach Using CNN and LSTM Models." SSRN, 2025.

[11] D. A. Phalke, "Artificial Intelligence in Fitness: Pose Estimation and Movement Correction." *Cureus*, 2025.

[12] V. Bazarevsky et al., "BlazePose: On-device Real-time Body Pose Tracking." arXiv, 2020.

[13] I. Grishchenko et al., "BlazePose: On-device Real-time Body Pose Tracking." Google Research Blog, 2021.

[14] W. Supanich et al., "Machine Learning-Based Exercise Posture Recognition System Using MediaPipe Pose Estimation Framework." ResearchGate, 2023.

[15] S. Pal et al., "Review-Based AI-Driven Posture Correction and Personalized Fitness Assistant Using Computer Vision and Augmented Reality." ResearchGate, 2025.

[16] A. Tharatipyakul, "Deep learning-based human body pose estimation in …" *ScienceDirect*, 2024.

[17] K. E. Appiah et al., "A Mobile-Phone Pose Estimation for Gym-Exercise Form Anomaly Detection." *VISAPP 2024*, 2024.

[18] D. A. Phalke et al., "Artificial Intelligence in Fitness: Pose Estimation and Movement Correction." *Cureus*, 2025.

[19] Ali K. AlShami et al., "SMART-Vision: Survey of Modern Action Recognition Techniques in Vision." arXiv, 2025.

[20] 20.Jungpil Shin, Najmul Hassan, Abu Saleh Miah, Satoshi Nishimura, "A Comprehensive Methodological Survey of Human Activity Recognition Across Diverse Data Modalities." arXiv, 2024.

[21] Nature article efficient human activity recognition on edge devices using TinyML. 2025.

[22] Application of human activity/action recognition: a review. *Multimedia Tools and Applications*, 2024.

[23] Machine Learning for Human Activity Recognition: State-of-the-Art. *MDPI*, 2024.

[24] A systematic literature review on human activity recognition. *Springer*, 2025.

[25] HARNet in deep learning approach—a systematic survey. *Scientific Reports*, 2024.

[26] A Survey on AI-Powered Exercise Posture Correction System. *IJAEM*, 2024.

[27] Ai-Powered Gym Posture Optimization. IRJMETS, 2024.

[28] AI-Based Posture Correction, Real-Time Exercise Tracking, and Feedback using Pose Estimation Technique. IEEE CCIS, 2024.

[29] Human Pose-based Estimation, Tracking and Action Recognition with Deep Learning: A Survey. arXiv, 2023.

[30] A Survey of Video Action Recognition Based on Deep Learning. *Knowledge-Based Systems*, June 2025.