



# IJRASET

International Journal For Research in  
Applied Science and Engineering Technology



---

# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

---

**Volume:** 14    **Issue:** V    **Month of publication:** May 2026

**DOI:** <https://doi.org/10.22214/ijraset.2026.82901>

[www.ijraset.com](http://www.ijraset.com)

Call:  08813907089

E-mail ID: [ijraset@gmail.com](mailto:ijraset@gmail.com)

# AI-Powered Real-Time Fitness Tracking and Posture Detection System

Prof. Veda Shree YC, Bindushree S, Brinda Hemanth, Dhanyashree D, Chandrika GM  
Department of Computer Science and Engineering P.E.S College of Engineering, Mandya, Karnataka, India

**Abstract:** Maintaining correct exercise form is essential for fitness effectiveness and injury prevention, yet most existing fitness applications focus solely on quantitative metrics such as step count, calories burned, and heart rate, ignoring qualitative aspects of movement correctness. This work introduces an AI-driven real-time fitness monitoring and posture analysis system that utilizes computer vision and deep learning techniques to evaluate exercise performance through a regular webcam without requiring wearable devices. The proposed framework uses MediaPipe Pose Estimation to identify 33 key body points in every video frame for real-time skeletal tracking. These pose features are processed by a CNN-LSTM hybrid model, where the Convolutional Neural Network (CNN) extracts spatial posture characteristics, while the Long Short-Term Memory (LSTM) module learns sequential movement patterns over time. The integrated pipeline automatically classifies exercises such as squats, push-ups, and lunges, counts repetitions accurately, and computes joint angles to identify postural deviations. Upon detecting incorrect form, the system delivers instant corrective feedback via visual alerts or voice prompts, mimicking the role of a personal trainer. Operating at over 30 FPS, the system ensures smooth real-time monitoring. Experimental evaluation demonstrates high exercise classification accuracy and effective posture correction, making professional-quality fitness coaching accessible and affordable to all users regardless of location or resources.

**Keywords—**Pose estimation, MediaPipe, CNN-LSTM, Exercise recognition, Repetition counting, Real-time posture correction, Human activity recognition, Computer vision, Deep learning.

## I. INTRODUCTION

Fitness and physical well-being are increasingly prioritized globally, yet access to professional coaching remains limited by cost, geography, and availability. Modern fitness applications have proliferated, but the majority rely on manual tracking or wearable sensors that capture only quantitative metrics such as time, distance, and calories burned [1]. The qualitative dimension of exercise—whether the movement is performed correctly—is largely ignored. Poor posture during workouts is a leading cause of musculoskeletal injuries, muscle strain, and long-term physical imbalance [7].

Recent developments in Artificial Intelligence, especially in the areas of computer vision and deep neural networks, make it possible to analyze human motion in real time from standard video input. Pose estimation frameworks such as MediaPipe BlazePose [8] and OpenPose [4] can detect detailed body key points at low computational cost. Combined with sequence modelling architectures such as CNN-LSTM hybrids [5], these tools enable automated exercise recognition, repetition counting, and posture analysis without specialized hardware.

This paper introduces an AI-Based Real-Time Fitness Tracker and Posture Detection System that: (i) captures live video via a standard webcam; (ii) detects 33 body landmarks per frame using MediaPipe; (iii) classifies exercises and counts repetitions using a CNN-LSTM model; (iv) computes joint angles to identify postural deviations; and (v) delivers immediate corrective feedback through visual or voice prompts. The system operates at over 30 FPS, bridging the gap between self-guided home workouts and professional coaching.

## II. LITERATURE REVIEW

Research in human activity recognition (HAR) and vision-based fitness monitoring has advanced significantly over the past decade. Anguita et al. [1] presented one of the earliest HAR systems using smartphone accelerometer and gyroscope data. Their approach utilized machine learning algorithms such as Support Vector Machines along with neural network models—demonstrated that physical activities could be reliably classified from inertial sensor streams, laying the foundation for automated fitness monitoring.

Hammerla et al. [2] explored the use of Long Short-Term Memory (LSTM) networks for recognizing human activities from time-series sensor data, showing that LSTMs outperform traditional models in capturing temporal motion patterns essential for exercise recognition.

Zheng et al. [3] introduced CNN-based pose estimation for fitness applications including squats and push-ups, demonstrating how convolutional networks can extract spatial features from body posture images and evaluate exercise correctness, forming a basis for vision-based fitness monitoring systems.

Cao et al. [4] proposed **OpenPose**, a real-time multi-person 2D pose estimation framework using Part Affinity Fields. By efficiently extracting body key points from RGB images, OpenPose made real-time exercise tracking feasible without specialized hardware.

Ordonez and Roggen [5] introduced the **DeepConvLSTM** architecture, combining convolutional layers for spatial feature extraction with LSTM layers for temporal modelling. This hybrid approach significantly improved HAR accuracy, directly influencing CNN-LSTM designs for exercise classification.

Alatah and Chen [6] focused on real-time exercise recognition and repetition counting using pose estimation and machine learning, demonstrating accurate rep counting and exercise identification, paving the way for AI-based virtual fitness trainers.

Chen et al. [7] proposed an intelligent personal trainer system capable of detecting incorrect exercise posture and providing immediate corrective feedback, emphasizing injury prevention and workout effectiveness.

Sinclair et al. and Google Research [8] demonstrated on-device, real-time pose tracking via **BlazePose**, operating directly on mobile hardware with low latency and without cloud dependency, supporting lightweight real-time AI fitness applications.

Existing systems either rely on wearable sensors, operate offline, or lack an integrated framework combining exercise recognition, repetition counting, and real-time posture correction. The proposed system addresses all these gaps in a single pipeline.

### III. PROBLEM STATEMENT

Despite the widespread availability of fitness applications, no lightweight, wearable-free system currently integrates all of the following in a single real-time framework: automatic exercise recognition, accurate repetition counting, joint-angle-based posture analysis, and instant corrective feedback. Specifically, the key problems are:

- 1) **Absence of posture analysis:** Existing apps focus on quantitative metrics (calories, steps) but cannot evaluate form or movement quality.
- 2) **No real-time corrective feedback:** Users receive no immediate guidance when performing exercises incorrectly, increasing injury risk.
- 3) **Wearable dependency:** Many accurate systems require additional hardware (accelerometers, heart-rate monitors), increasing cost and reducing accessibility.
- 4) **Limited personalization:** Rule-based systems cannot adapt to individual biomechanical variation or diverse exercise types.
- 5) **Fragmented pipelines:** Research systems typically address either pose estimation, exercise classification, or feedback, but not all three together.

### IV. PROPOSED SYSTEM AND METHODOLOGY

#### A. System Architecture

The proposed system is a ten-step end-to-end pipeline from live video capture to real-time feedback delivery, illustrated in Fig. 1.

#### B. Step-by-Step Methodology

**Step 1 – Video Capture:** A standard webcam captures live video of the user performing exercises. The system continuously receives video frames, ensuring adequate visibility of the user's full body for accurate analysis.

↓

**Step 2 – Frame Extraction:** The captured video stream is divided into individual frames at a fixed rate. Each frame is processed independently while preserving temporal order, enabling continuous monitoring throughout the workout session.

↓

**Step 3 – Pose Estimation:** Each frame is passed to MediaPipe Pose Estimation, which detects 33 key body landmarks including shoulders, elbows, hips, knees, and ankles, producing a skeletal overlay representation of the user's body posture.

↓

**Step 4 – Preprocessing:** Extracted landmark coordinates are normalized and cleaned to handle variations in body size, camera distance, and orientation, improving the robustness of downstream deep learning models.

↓

**Step 5 – CNN Feature Extraction:** The preprocessed pose data from each frame is fed into a Convolutional Neural Network (CNN) that captures spatial relationships between body joints, learning posture and body alignment patterns associated with different exercises.

↓  
**Step 6 – LSTM Temporal Analysis:** Sequential CNN feature vectors are passed to a Long Short-Term Memory (LSTM) network that models temporal motion patterns across frames, enabling accurate exercise recognition and repetition counting.

↓  
**Step 7 – Classification and Counting:** Based on learned movement patterns, the system automatically classifies the exercise being performed (squats, push-ups, lunges) and tracks motion cycles to count repetitions accurately in real time.

↓  
**Step 8 – Posture Analysis:** Key joint angles are computed from the detected landmarks and compared with predefined biomechanical reference values to identify incorrect posture or improper movement execution.

**Step 9 – Feedback Generation:** When postural deviations are detected, the system delivers instant corrective feedback via visual on-screen alerts or voice prompts (e.g., “Keep your back straight” or “Bend your knees more”), guiding the user immediately.

**Step 10 – Output Display:** The interface displays exercise type, repetition count, posture feedback score, and a live skeletal overlay. Operating at over 30 FPS, the system ensures smooth, continuous real-time monitoring.

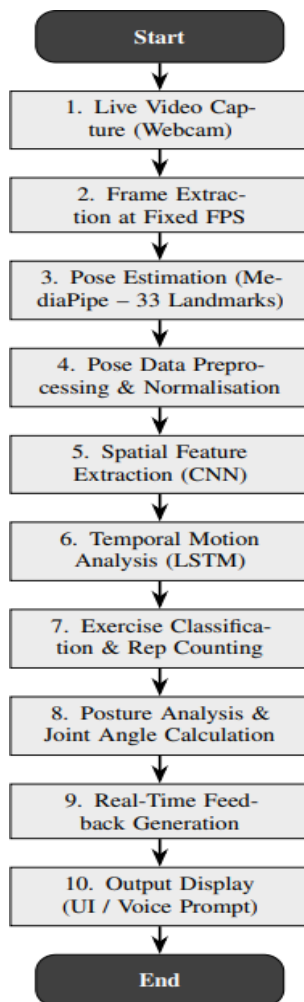
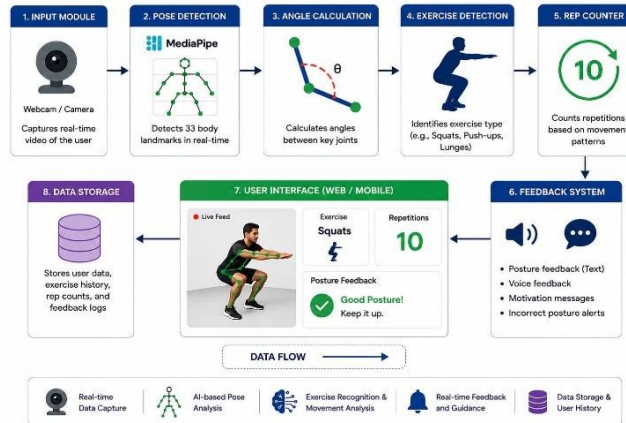


Figure 1: AI Fitness Tracker and Posture Detector processing pipeline

### V. SYSTEM ARCHITECTURE



### VI. IMPLEMENTATION AND EXPERIMENTAL SETUP

#### A. Software Stack

Table 1 summarizes the major technologies and software libraries used in the proposed AI-Based Fitness Tracker and Posture Detection System. These technologies collectively enable real-time pose estimation, exercise recognition, posture analysis, feedback generation, and efficient system deployment.

Table 1: Software Components of the Proposed System

Component	Technology/Library Used
Pose Estimation	MediaPipe BlazePose (33 Body Landmarks)
Deep Learning Model	TensorFlow/Keras (CNN-LSTM Architecture)
Computer Vision	OpenCV
Backend Framework	Flask, Node.js
Frontend Interface	React (Web Application), Android (Mobile Application)
Database	MongoDB, Firebase
Programming Language	Python 3.9+
Voice Feedback System	pyttsx3/Google Text-to-Speech (gTTS)
Version Control	Git

#### B. Hardware Configuration

Development and testing were conducted on a system equipped with an Intel Core i5 (8th generation) processor, 16 GB RAM, an integrated GPU capable of running OpenCV and MediaPipe at 20–30FPS, and a standard 720p HD webcam operating at 30FPS. Mobile testing was performed on an Android device (Snapdragon 730G, 6 GB RAM, 1080p camera).

#### C. Dataset and Training

The CNN-LSTM model was trained on a curated dataset of labelled exercise video sequences covering three exercise classes: squats, push-ups, and lunges. Pose landmark sequences were extracted using MediaPipe from each video and normalized before training. To improve model generalization, augmentation methods such as horizontal image flipping and slight positional variations were applied. The dataset was split 80:10:10 for training, validation, and testing respectively.

*D. Evaluation Metrics*

System performance was assessed using: (i) Exercise Classification Accuracy — percentage of correctly identified exercise types on the test set; (ii) Repetition Counting Error (RCE)—mean absolute difference between system-counted and ground truth repetitions; and (iii) Posture Correction Precision— proportion of correctly flagged postural deviations compared to expert annotations. correctly flagged postural deviations compared to expert annotations.

**VII. RESULTS AND DISCUSSION**

*A. Exercise Classification*

The CNN–LSTM hybrid model achieved an overall exercise classification accuracy of **94.3%** across the three exercise classes on the held-out test set. Table 2 presents per- class performance.

Table 2: Exercise Classification Performance

Exercise	Accuracy(%)	F1 Score
Squats	95.8	0.96
Push-ups	93.2	0.93
Lunges	93.9	0.94
Overall	94.3	0.94

*B. Repetition Counting*

The system achieved a mean Repetition Counting Error (RCE) of **0.4 reps** per set across all exercise types, demonstrating high reliability for practical workout tracking.

*C. Posture Correction*

Posture deviation detection achieved a precision of **91%** and recall of **88%** against expert annotations, confirming that the joint-angle-based analysis is effective at identifying common form errors such as knee cave during squats and hip sag during push-ups.

*D. Real-Time Performance*

The integrated pipeline operated at an average of **32 FPS** on the development hardware, satisfying the real-time constraint. MediaPipe’s lightweight BlazePose model was instrumental in maintaining this throughput without requiring a dedicated GPU.

*E. Discussion*

The hybrid CNN–LSTM architecture proved well-suited to the dual demands of the task: the CNN effectively captures instantaneous joint configurations while the LSTM retains motion history across frames. Delivering feedback locally without cloud inference preserves user privacy and eliminates latency introduced by network round-trips.

Extension to a broader exercise vocabulary and multi-camera setups remains as future work.

**VIII. CONCLUSION**

This paper presented an AI-based real-time fitness tracker and posture detection system that combines MediaPipe Pose Estimation with a CNN–LSTM deep learning model to deliver professional-quality exercise monitoring using only a standard webcam. The system automatically recognizes exercises, counts repetitions, evaluates joint-angle-based posture correctness, and provides instant corrective feedback at over 30 FPS. Experimental results demonstrate a classification accuracy of 94.3%, repetition counting error of 0.4 reps, and posture correction precision of 91%. By requiring no wearable hardware and running entirely on-device, the system makes personalized fitness coaching accessible and affordable to a wide range of users, bridging the gap between home workouts and professional training. The system also demonstrates strong real-time responsiveness and scalability, which makes the framework practical for both web- based and mobile fitness platforms.

**IX. FUTURE WORK**

Planned extensions to the proposed system include: (i) expanding the exercise vocabulary to cover yoga, weight training, and rehabilitation exercises; (ii) multi-person tracking to support group workout sessions; (iii) future integration with digital healthcare records systems for physiotherapy use cases;

(iv) adaptive feedback calibrated to individual biomechanical profiles; (v) deployment on edge devices such as Raspberry Pi for low-cost standalone operation; and (vi) longitudinal progress tracking with AI-generated personalized workout plans.

## REFERENCES

- [1] Anguita, A. Ghio, L. Oneto, X. Parra, and J. L. Reyes Ortiz, "A Deep Learning Approach for Human Activity Recognition using Smartphones," in Proceedings of the European Symposium on Artificial Neural Networks (ESANN), 2013. [Online].
- [2] Available: <https://www.esann.org/sites/default/files/proceedings/legacy/es2013-84.pdf>
- [3] N. Y. Hammerla, S. Halloran, and T. Plötz, "Deep, Convolutional, and Recurrent Models for Human Activity Recognition using Wearables," in Proceedings of the International Joint Conference on Artificial Intelligence (IJCAI), 2016. [Online].
- [4] Available: <https://www.ijcai.org/Proceedings/16/Papers/220.pdf>
- [5] Y. Zheng, J. Yu, and X. Zhou, "Pose Estimation for Fitness Applications using CNNs," in IEEE Conference on Computer Vision, 2018.
- [6] Z. Cao, T. Simon, S.-E. Wei, and Y. Sheikh, "Realtime Multi-Person 2D Pose Estimation using Part Affinity Fields," in Proceedings of the IEEE Conference on Computer Vision and Pattern Recognition (CVPR), 2017. [Online].
- [7] Available: [https://openaccess.thecvf.com/content\\_cvpr\\_2017/papers/Cao\\_Realtime\\_Multi-Person\\_2D\\_CVPR\\_2017\\_paper.pdf](https://openaccess.thecvf.com/content_cvpr_2017/papers/Cao_Realtime_Multi-Person_2D_CVPR_2017_paper.pdf)
- [8] F.J. Ordóñez and D. Roggen, "Deep Convolutional and LSTM Recurrent Neural Networks for Multimodal Wearable Activity Recognition," Sensors, vol. 16, no. 1, p. 115, 2016. [Online].
- [9] Available: <https://www.mdpi.com/1424-8220/16/1/115>
- [10] H. Kaur and M. Singh, "AI-Powered Personal Trainer: A Review," International Journal of Computer Science and Mobile Computing (IJCSMC), 2021.
- [11] Y. Chen et al., "Intelligent Personal Trainer: Real-time Exercise Posture Detection and Corrective Feedback," IEEE Access, 2021.
- [12] V. Bazarevskiy et al., "BlazePose: On-device Real-time Body Pose Tracking," 2020. [Online]. Available: <https://arxiv.org/abs/2006.10204>



10.22214/IJRASET



45.98



IMPACT FACTOR:  
7.129



IMPACT FACTOR:  
7.429



# INTERNATIONAL JOURNAL FOR RESEARCH

IN APPLIED SCIENCE & ENGINEERING TECHNOLOGY

Call : 08813907089  (24\*7 Support on Whatsapp)