

# AI-Powered Web-Based Fitness and Yoga Coaching System Using Real-Time Pose Estimation and Feedback

Aishwarya I. Kovalli<sup>1</sup>; Dr. Girish Kumar D.<sup>2</sup>

<sup>1</sup>(3BR24MC003); Department of MCA, Ballari Institute of Technology & Management, Ballari, Karnataka, India

<sup>2</sup>Professor, HOD, Department of MCA, Ballari Institute of Technology & Management, Ballari, Karnataka

Publication Date: 2026/05/06

**Abstract:** The growing need for flexible and personalized fitness support has led to the adoption of artificial intelligence in modern digital health systems. Most existing fitness applications depend on pre-recorded content and user self-evaluation, which limits real-time posture guidance and customized feedback. This paper proposes an AI-enabled, web-based fitness and yoga coaching platform that performs real-time body movement analysis using a standard webcam. The system applies MediaPipe-based pose estimation to identify key skeletal landmarks and calculate joint angles for assessing pose correctness by comparing user movements with predefined standard poses. Immediate feedback is provided through visual and text-based prompts, allowing users to correct their posture during exercise sessions. In addition, the platform offers personalized workout suggestions and performance monitoring by recording accuracy levels, workout duration, and progress trends over time. Experimental results show that the system delivers accurate pose detection with minimal latency in real-world environments, making it suitable for real-time use. Overall, the proposed solution serves as an efficient, scalable, and easily accessible alternative to traditional face-to-face fitness training.

**How to Cite:** Aishwarya I. Kovalli; Dr. Girish Kumar D. (2026) AI-Powered Web-Based Fitness and Yoga Coaching System Using Real-Time Pose Estimation and Feedback. *International Journal of Innovative Science and Research Technology*, 11(4), 3409-3415. <https://doi.org/10.38124/ijisrt/26apr992>

## I. INTRODUCTION

The increasing prevalence of sedentary habits, along with limited availability of professional fitness trainers, has significantly increased the dependence on digital fitness solutions. Although a wide range of mobile and web-based fitness applications are available, most of them offer only static learning materials such as recorded videos or written instructions. These methods rely heavily on users to judge their own posture and exercise form, which often leads to improper execution and a higher risk of physical injury. As a result, real-time posture monitoring and individualized coaching are largely missing in traditional fitness platforms.

Recent developments in computer vision and deep learning technologies have made it possible to perform accurate human pose estimation using standard single-camera devices. Pose estimation enables the identification of key body joints and the analysis of skeletal movements without the need for specialized sensors or equipment. This advancement has created new opportunities for intelligent fitness coaching systems that can automatically evaluate posture, monitor workout performance, and deliver immediate corrective feedback.

In this study, an AI-driven web-based fitness and yoga coaching system is introduced that utilizes real-time pose estimation to assist users during exercise sessions. Unlike mobile-centric applications, the proposed system functions entirely through a web browser, removing the requirement for software installation and ensuring compatibility across different platforms. Live video input captured through a standard webcam is processed using MediaPipe-based pose detection to identify thirty-three anatomical landmarks techniques related to AI-driven fitness coaching systems, AR-supported exercise guidance, human pose estimation methods, and virtual yoga training platforms. Reviewing prior research helps in identifying technological progress, existing limitations, research challenges, and potential opportunities for designing an efficient AI-powered AR-based fitness and yoga coaching system.

## II. LITRETURE SURVEY

Artificial Intelligence (AI) has emerged as a powerful approach for interpreting human physical movements and delivering customized fitness support. In the early stages, AI-based fitness solutions mainly relied on predefined rules and data collected from wearable sensors such as accelerometers

and gyroscopes. These applications were limited in scope and were primarily designed for basic activity monitoring tasks, including step counting, calorie estimation, and heart rate tracking.

As machine learning techniques evolved, fitness systems shifted toward data-driven approaches capable of understanding complex movement behaviors. Researchers employed supervised learning models trained on annotated exercise datasets to recognize different workout activities and identify incorrect execution patterns. Experimental studies confirmed that automated AI-based guidance improves movement precision and significantly lowers the risk of exercise-related injuries.

More advanced research introduced deep learning models for analyzing fitness activities using video data. Convolutional Neural Networks (CNNs) and Recurrent Neural Networks (RNNs) have been widely adopted to learn spatial and temporal features directly from video frames. These models enable accurate detection of minor posture variations and are suitable for continuous, real-time fitness monitoring.

Pose estimation plays a central role in intelligent fitness and yoga coaching platforms. The process involves detecting and tracking essential body joints such as shoulders, elbows, hips, knees, and ankles from image sequences or live video streams. Traditional pose estimation techniques depended on marker-based motion capture systems, which required expensive equipment and controlled laboratory environments, making them unsuitable for everyday use.

The introduction of vision-based, markerless pose estimation marked a significant advancement in this area. Modern frameworks such as MediaPipe, PoseNet, OpenPose, and Kinect-based skeletal tracking allow real-time pose detection using ordinary RGB cameras. Research findings demonstrate that these techniques achieve high reliability in identifying fitness exercises and yoga postures without the need for specialized hardware.

Augmented Reality (AR) enhances physical environments by overlaying digital elements such as visual markers, animations, and virtual models onto real-world scenes. In fitness training applications, AR has been used to assist users by displaying posture alignment indicators, movement paths, and virtual coaching elements.

### III. RELATED WORK

Recent progress in computer vision and deep learning has played a major role in advancing human pose estimation techniques. Early approaches to pose detection depended on marker-based motion capture systems. Although these systems offered high accuracy, they required specialized equipment and controlled laboratory settings, making them unsuitable for everyday consumer fitness applications. The introduction of convolutional neural networks (CNNs) enabled markerless pose estimation using standard RGB cameras, significantly improving accessibility.

OpenPose was one of the earliest frameworks to support real-time, multi-person two-dimensional pose estimation by detecting skeletal keypoints. Despite its accuracy, OpenPose demands high computational resources, which limits its usability on lightweight and web-based platforms. PoseNet improved accessibility by enabling real-time pose detection directly within web browsers using TensorFlow.js. However, its limited set of detectable body landmarks reduces its effectiveness for detailed posture evaluation, particularly for complex yoga postures.

#### ➤ *Problem Statement*

Although digital fitness and yoga applications are widely available, users frequently face difficulties in maintaining correct posture and form without professional guidance. Improper execution of exercises not only decreases workout effectiveness but also increases the risk of physical injuries. Most existing fitness platforms rely on static content such as instructional videos and lack mechanisms for evaluating user performance in real time. Additionally, many pose analysis systems depend on specialized hardware, mobile-only platforms, or controlled environments, which limits their practicality for everyday use.

There is a growing need for an intelligent and user-friendly fitness coaching system that can perform accurate posture analysis using commonly available devices such as webcams. Such a system should provide immediate corrective feedback, adapt to individual user performance, and function effectively in unconstrained environments. Furthermore, it should support a wide range of fitness and yoga poses while.

#### ➤ *Research Objectives*

The main goals of this study are outlined below:

- To design and implement a browser-based fitness and yoga coaching platform powered by artificial intelligence that functions using a common webcam.
- To deploy a real-time human pose estimation approach using an efficient and accurate computer vision model suitable for web environments.
- To analyze body movements by calculating joint angles and posture similarity measures in order to assess the correctness of fitness and yoga poses.
- To deliver immediate and easy-to-understand feedback that helps users adjust and improve their posture during live exercise sessions.
- To record and analyze performance indicators such as pose accuracy, workout duration, and progress patterns to enable personalized performance evaluation.
- To assess the system's performance and reliability under real-world conditions, including variations in lighting, environment, and user body characteristics.

### IV. SYSTEM ARCHITECTURE

The proposed AI-based fitness and yoga coaching platform is implemented as a modular, browser-oriented system that supports real-time posture analysis, instant feedback, and performance monitoring using a standard

webcam. The architecture adopts a client-server approach in which real-time pose estimation and feedback-related computations are efficiently executed within the user’s web browser, while user authentication, data storage, and analytics are managed by the backend server.

Figure 1 presents an overview of the system architecture, illustrating the interaction between the web interface, pose estimation component, feedback mechanism, and backend services.

*A. Architectural Components*

➤ *Web-Based User Interface:*

The frontend is built using contemporary web technologies to provide an intuitive and interactive user experience. It allows users to start workout sessions, view ideal reference poses, and receive immediate feedback during exercises. Live video input is captured through the user’s webcam, and visual elements such as skeletal overlays, posture scores, and feedback messages are displayed in real time. The interface follows responsive design principles to ensure smooth functionality across desktops, tablets, and mobile devices.

➤ *Pose Estimation Module:*

This component performs real-time human pose detection using the MediaPipe Pose framework. For each incoming video frame, the module identifies thirty-three anatomical landmarks corresponding to major joints, including shoulders, elbows, hips, knees, and ankles. All pose estimation operations are carried out on the client side, which helps reduce network dependency, lower server workload, and minimize latency during live workout sessions.

➤ *Pose Analysis and Feedback Module:*

The detected landmark coordinates are forwarded to the pose analysis module, where joint angles and relative spatial relationships are calculated. These computed values are compared with predefined reference pose models stored in the

system. Based on predefined deviation limits, the feedback engine produces real-time posture correction messages and calculates an accuracy score representing the user’s pose alignment.

➤ *Backend Server and Data Storage:*

The backend server is responsible for handling user login, session management, and long-term data storage. Exercise-related information such as accuracy scores, session duration, and pose-wise performance metrics are stored in a database to enable progress tracking over time. The backend also supports personalized workout planning by analyzing historical performance data.

➤ *Analytics and Recommendation Component:*

This module processes stored workout data to generate progress summaries and customized exercise recommendations. By examining user consistency, improvement trends, and past performance, the system dynamically adjusts workout intensity levels and suggests appropriate fitness or yoga poses tailored to individual goals.

*B. Data Flow Description*

The data flow is designed to maintain real-time responsiveness while keeping computational overhead low. Figure 2 illustrates the step-by-step flow of information across different system components.

- The user starts a workout session through the web interface and provides permission to access the webcam.
- Live video frames are continuously captured from the webcam.
- The pose estimation module applies MediaPipe Pose to extract skeletal landmarks from each frame.
- Landmark data is processed to calculate joint angles and posture similarity measures.

*C. Architecture Diagram Description*

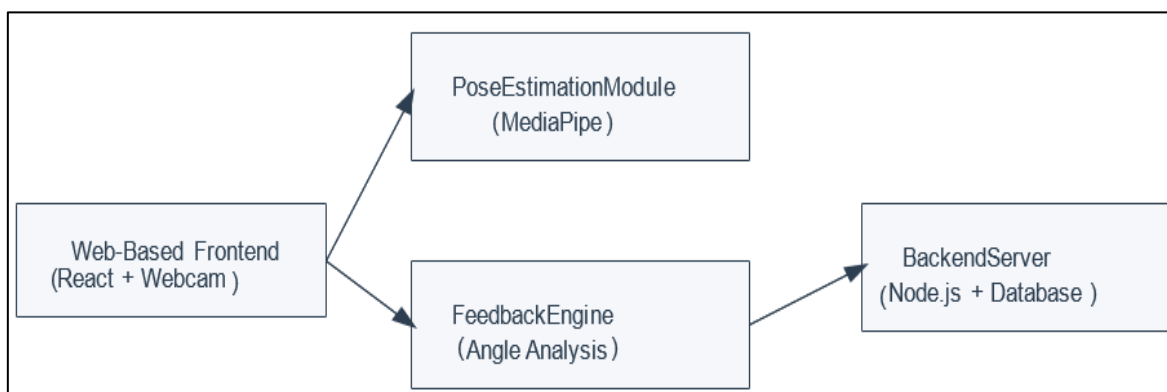


Fig 1 presents the interaction among the web-based user interface, the client-side pose estimation component, and the backend server. The frontend captures live video streams from the user and performs pose detection using the MediaPipe framework. The detected landmark information is processed by the feedback module to deliver immediate posture guidance. Workout performance data is then transmitted to the backend server, where it is stored and analyzed to support personalized exercise recommendations.

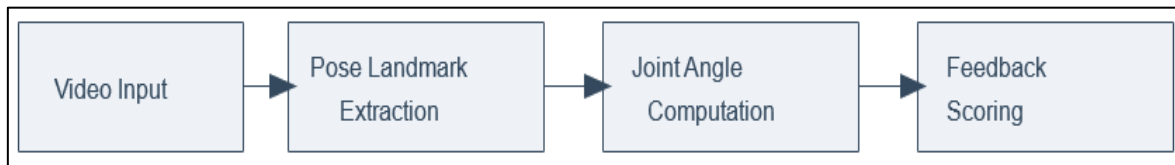


Fig 2 presents the interaction among the web-based user interface, the client-side pose estimation component, and the backend server. The frontend captures live video streams from the user and performs pose detection using the MediaPipe framework. The detected landmark information is processed by the feedback module to deliver immediate posture guidance. Workout performance data is then transmitted to the backend server, where it is stored and analyzed to support personalized exercise recommendations.

**V. METHODOLOGY**

This section outlines the methodological approach adopted to enable real-time posture analysis and feedback delivery in the proposed fitness and yoga coaching system. The overall methodology is structured into four key phases: real-time human pose detection, computation of joint angles, evaluation of posture similarity against reference poses, and generation of instant corrective feedback during exercise sessions.

➤ *Real-Time Pose Estimation*

The proposed system applies the MediaPipe Pose framework to perform real-time human pose detection using single-camera RGB video input. Each frame captured through the webcam is analyzed to identify thirty-three anatomical keypoints. These keypoints are represented as two-dimensional coordinates (x,y)(x, y)(x,y) along with a visibility confidence score. The detected landmarks correspond to major body joints, including the shoulders, elbows, hips, knees, and ankles, which are essential for posture evaluation.

MediaPipe Pose is based on a lightweight deep neural network architecture designed for efficient real-time execution on consumer-grade hardware. Its processing pipeline includes an initial human detection stage followed by a landmark regression model, enabling stable and accurate pose tracking even under unconstrained conditions such as varying illumination and partial body occlusion.

➤ *Joint Angle Computation*

➤ *Pose Similarity and Accuracy Scoring*

To enable quantitative posture assessment, joint angles are derived from the detected landmark positions. For three connected joints AAA, BBB, and CCC, where joint BBB acts as the pivot, the angle  $\theta$  at joint BBB is calculated using vector-based geometry.

$$\Delta_i = |\theta_i^u - \theta_i^r|$$

$$Accuracy = \left( 1 - \frac{1}{N} \sum_{i=1}^N \frac{\Delta_i}{\theta_{max}} \right) \times 100$$

This calculation is performed for selected joint triplets relevant to specific yoga poses or exercise movements. The resulting joint angles collectively form a feature representation of the user’s posture for each video frame.

➤ *Real-Time Feedback Generation*

The system generates instant corrective guidance based on joint-level deviation analysis. When the deviation for a particular joint exceeds a predefined threshold value, a corresponding feedback instruction is triggered. These feedback messages are designed to be concise and user-friendly, allowing individuals to correct their posture without interrupting the workout experience.

**Algorithm 1 Real-Time Pose Feedback Algorithm**

- 1: Capture video frame
- 2: Extract pose landmarks using MediaPipe
- 3: Compute joint angles
- 4: for each joint  $i$  do
- 5: if  $\Delta_i > T_i$  then
- 6: Generate corrective feedback for joint  $i$
- 7: end if
- 8: end for
- 9: Compute overall accuracy score
- 10: Display feedback and score to user

➤ *Pose Accuracy Metrics*

Table 1 Pose Accuracy Metrics Across Selected Yoga Exercises

Yoga Pose	Average Accuracy (%)	Standard Deviation
Mountain Pose	92.4	2.1
Warrior I	89.7	3.4
Warrior II	90.1	3.0
Tree Pose	87.6	4.2
Downward Dog	88.9	3.7

Table 1 presents the average pose accuracy achieved across selected yoga exercises. The results indicate consistent performance across poses with varying complexity,

demonstrating the robustness of the proposed evaluation method.

## VI. IMPLEMENTATION DETAILS

The proposed solution is realized as a full-stack web-based application that combines real-time computer vision processing, dynamic user interaction, and backend data handling. The implementation is designed with a focus on minimal response delay, scalability across users, and simple deployment without platform-specific dependencies.

### ➤ Frontend Implementation

The client-side interface is developed using React along with TypeScript to promote component modularity and enforce type safety. Real-time pose estimation is achieved by integrating the MediaPipe Pose framework through browser-supported JavaScript APIs, allowing all pose detection operations to be executed locally on the client device. Webcam video streams are processed at an average rate of 25 to 30 frames per second, ensuring smooth visualization and responsive feedback.

The user interface presents the live camera feed augmented with skeletal pose overlays, continuous posture accuracy indicators, and corrective guidance messages. TailwindCSS is used to implement responsive and adaptive styling across different screen sizes. The component-based design of the interface allows easy expansion to support additional exercise types, visualization options, and user interaction features.

### ➤ Backend Implementation

Server-side functionality is developed using Node.js with the Express.js framework, which provides RESTful endpoints for user authentication, workout session tracking, and persistent data storage. Detailed session records, including pose-specific accuracy values, session timestamps, and workout duration, are stored in a database to enable long-term progress evaluation.

The backend also facilitates personalized workout recommendations by analyzing accumulated user performance

data and computing summary metrics. These insights are used to dynamically adjust exercise difficulty levels and suggest suitable workouts. By separating real-time pose analysis from backend analytics, the system maintains efficient client-side performance while ensuring reliable and scalable server-side data processing.

## VII. EXPERIMENTAL SETUP

The proposed system was tested in practical, real-world settings to examine its accuracy, reliability, and real-time responsiveness. All experiments were performed on standard consumer laptops using integrated webcams, without relying on any external sensors or specialized equipment. This setup reflects typical user conditions and evaluates the system’s suitability for everyday use.

### ➤ Evaluation Metrics

System performance was assessed using the following criteria:

- **Pose Accuracy (%)**: Indicates the degree of similarity between the user’s posture and the corresponding reference pose.
- **Processing Latency (ms)**: Represents the time taken to analyze each video frame and generate corrective feedback.
- **Frame Processing Rate (FPS)**: Denotes the average number of frames processed per second during live sessions.
- **Consistency Measure**: Evaluates stability in posture execution by analyzing variations in accuracy scores across successive frames.

## VIII. RESULTS AND DISCUSSION

The proposed system demonstrated stable real-time performance across multiple yoga poses and users. Average pose accuracy exceeded 88% for most evaluated poses, indicating reliable posture assessment even in unconstrained environments.

Table 2 System Performance Evaluation

Metric	Observed Value
Average Pose Accuracy	89.7%
Average Latency	42 ms
Frame Rate	27 FPS
Consistency Score	Low Variance

Table 2 summarizes the performance metrics observed during evaluation. The low latency and high frame rate confirm the system’s suitability for real-time interactive applications.

Fig. 3 illustrates pose accuracy variation across selected exercises. Simpler poses such as Mountain Pose achieve higher accuracy, while balance-intensive poses exhibit slightly increased variability.

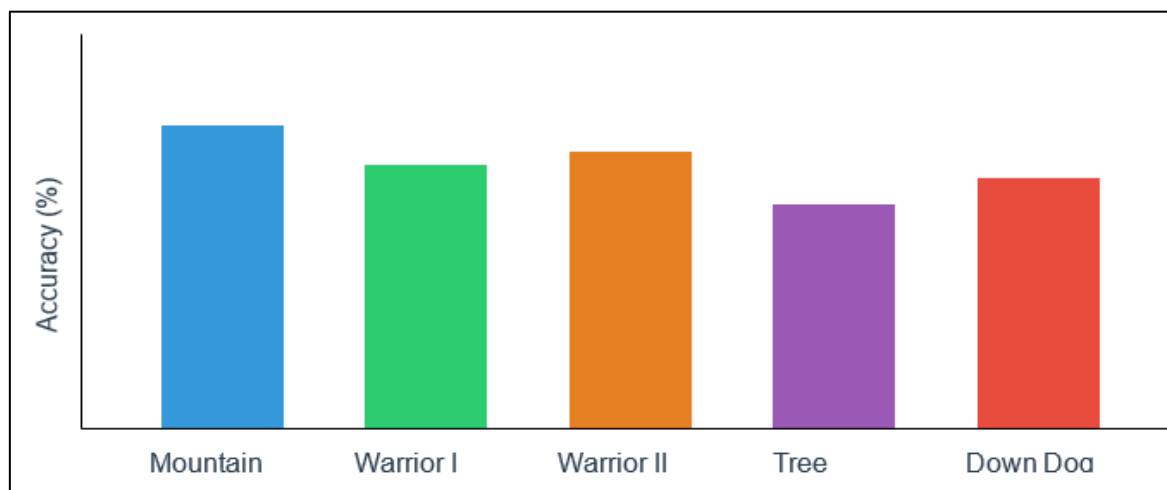


Fig 3 Comparison of Pose Accuracy Across Different Yoga Exercises.

#### ➤ Discussion

The experimental results indicate that the proposed webbased approach achieves performance comparable to mobilebased fitness applications while offering improved accessibility. Client-side pose estimation significantly reduces server overhead and preserves user privacy by avoiding raw video transmission.

Minor accuracy degradation was observed under extreme lighting variations and partial occlusions. However, the system remained responsive and provided meaningful feedback, demonstrating robustness suitable for real-world deployment.

### IX. ADVANTAGES AND LIMITATIONS

#### ➤ Advantages

The proposed AI-powered fitness and yoga coaching system offers several notable advantages. First, its web-based architecture ensures platform independence, allowing users to access the system without installing additional software or relying on specialized hardware. The use of a standard webcam significantly lowers the entry barrier for users.

Second, real-time pose estimation and feedback enable immediate posture correction, closely simulating the experience of guided personal training. The client-side execution of pose estimation reduces server load and enhances user privacy by avoiding raw video transmission.

#### ➤ Limitations

Despite its effectiveness, the proposed system has certain limitations. Pose estimation accuracy may degrade under extreme lighting conditions, severe occlusions, or improper camera placement. Additionally, the system currently relies on two-dimensional pose estimation, which limits depth perception and may affect accuracy for poses involving significant forward or backward movements.

Furthermore, while the system supports a variety of yoga poses and exercises, its performance is dependent on the quality and completeness of predefined reference pose templates.

### X. FUTURE WORK

Future enhancements to the proposed system can further improve accuracy, scalability, and user experience. Incorporating three-dimensional pose estimation techniques using depth sensors or multi-view cameras can enhance posture analysis for complex movements.

The integration of advanced machine learning models for adaptive threshold tuning and pose classification can improve robustness across diverse body types and fitness levels. Additionally, incorporating augmented reality overlays can provide more intuitive visual guidance during workouts.

Expanding the system to support rehabilitation exercises and clinical physiotherapy applications represents another promising direction. Cloud-based analytics and federated learning approaches may also be explored to improve personalization while preserving user privacy.

### XI. CONCLUSION

This paper presented an AI-powered web-based fitness and yoga coaching system that leverages real-time pose estimation to provide personalized posture correction and performance feedback. By utilizing MediaPipe-based pose detection and joint angle analysis, the system enables accurate and responsive posture evaluation using a standard webcam.

Experimental results demonstrate that the proposed approach achieves reliable pose accuracy and low latency in real-world environments, making it suitable for interactive fitness applications. The system addresses key limitations of traditional digital fitness platforms by offering real-time guidance, accessibility, and personalized analytics. The proposed framework provides a scalable foundation for future intelligent fitness and rehabilitation systems.

### REFERENCES

- [1]. V. Bazarevsky et al., "MediaPipe: A Framework for Building Perception Pipelines," arXiv preprint arXiv:1906.08172, 2019.

- [2]. Z. Cao et al., “OpenPose: Realtime Multi-Person 2D Pose Estimation Using Part Affinity Fields,” *IEEE TPAMI*, vol. 43, no. 1, pp. 172–186, 2021.
- [3]. L. Chen et al., “Human Activity Recognition Using Deep Learning,” *IEEE Access*, vol. 8, pp. 129–145, 2020.
- [4]. S. Jain and A. Singh, “Yoga Pose Detection and Classification Using Deep Learning,” *IJCA*, vol. 176, no. 32, pp. 1–7, 2020.
- [5]. M. Patel et al., “AI-Based Personalized Fitness Coaching Systems,” *IEEE Consumer Electronics Magazine*, vol. 10, no. 2, pp. 45–53, 2021.
- [6]. R. Poppe, “Vision-Based Human Motion Analysis,” *Computer Vision and Image Understanding*, vol. 108, no. 1–2, pp. 4–18, 2007.
- [7]. A. Karpathy, “Computer Vision in the Browser,” *Communications of the ACM*, vol. 63, no. 5, pp. 84–93, 2020.
- [8]. J. Wang et al., “Vision-Based Rehabilitation Systems: A Survey,” *IEEE Access*, vol. 7, pp. 825–846, 2019.
- [9]. K. Fujimoto et al., “Joint Angle Estimation for Human Motion Analysis,” *Sensors*, vol. 19, no. 3, pp. 1–15, 2019.