

Browser-Based AI Coach for Real-Time Fitness and Dance Training Using Human Pose Estimation

R. C. Mahajan¹; Bhupesh Cholake²; Chinmayee More³;
Aryan Varkhede⁴; Anushka Shirsath⁵

¹Professor

^{1,2,3,4}Department of Artificial Intelligence and Data Science PVGs College of Engineering & S.S.D. IOM
Nashik, Maharashtra, India

Publication Date: 2026/06/13

Abstract: This project proposes a practical, web-based intelligent coaching framework that integrates browser-side human pose estimation, rule-based biomechanical reasoning, natural-language coaching through large language models, adaptive workout generation, and nutrition planning into a unified digital health platform. The system is designed to provide live exercise assessment and dance guidance without training custom deep posture-classification models. Instead, it leverages pretrained pose landmark detectors and a layered decision architecture consisting of pose extraction, movement interpretation, safety filtering, scoring, and personalized coaching. The project sits at the intersection of human-computer interaction, computer vision, intelligent tutoring systems, digital therapeutics, sports analytics, and personalized health informatics.

Keywords: AI Fitness Coach; Human Pose Estimation; Posture Correction; Dance Analysis; Explainable AI.

How to Cite: R. C. Mahajan; Bhupesh Cholake; Chinmayee More; Aryan Varkhede; Anushka Shirsath (2026) Browser-Based AI Coach for Real-Time Fitness and Dance Training Using Human Pose Estimation. *International Journal of Innovative Science and Research Technology*, 11(5), 4212-4218. <https://doi.org/10.38124/ijisrt/26may2242>

I. INTRODUCTION

The advancement of browser-based machine learning and large language models has enabled new possibilities for real-time, personalized digital health systems. However, existing fitness applications often lack real-time posture correction, explainability, and integrated coaching features, limiting their effectiveness and accessibility.

This paper proposes an AI Fitness and Dance Coach, a browser-based platform that uses live camera input to perform real-time pose detection, movement analysis, repetition counting, and posture correction. The system provides actionable feedback, performance scoring, and personalized workout and diet recommendations through a combination of rule-based biomechanics and LLM-assisted coaching.

Unlike traditional approaches, the system avoids custom CNN model training and instead relies on pretrained pose-estimation models, making it efficient, interpretable, and deployable on consumer devices. The platform supports both exercise and beginner-level dance coaching by analyzing pose similarity and timing alignment.

The proposed system addresses key gaps in digital fitness solutions by enabling real-time analysis, explainable feedback, and low-cost deployment, while improving user engagement through conversational AI. It also provides a foundation for future research in browser-based computer vision, digital health, and intelligent coaching systems.

II. RELATED WORK

In recent years, the integration of artificial intelligence in fitness and dance training has gained significant attention. Early systems focused on basic activity recognition, while recent advancements emphasize real-time feedback, personalization, and intelligent coaching.

Zhang and Chen proposed an AI-driven personalized training system for dance movements using ontology models and neural networks to recommend customized training programs based on user similarity. Their work demonstrated improvements in dance performance but lacked real-time interactive feedback and explainable movement analysis [1].

Ko et al. proposed a real-time posture correction system using YOLOv5 and MediaPipe for powerlifting exercises.

The system provided detailed feedback and improved safety but focused mainly on specific exercises and required structured datasets [3].

Xiang and Gao introduced a deep learning-based framework for sports dance scene detection using perceptual feature integration and active learning techniques. Their approach focused on scene understanding and action optimization but did not provide real-time coaching or user interaction [4].

Nair et al. developed an AI-based dance coaching system using pose estimation, Vision Transformers, and Dynamic Time Warping (DTW) for segment-wise feedback. While the system achieved high accuracy and improved user performance, it relied on deep learning models and required structured datasets, limiting accessibility [5].

Gurbuxani et al. presented a virtual fitness trainer using OpenCV and MediaPipe to monitor exercise posture and provide real-time feedback. Although effective for fitness guidance, the system lacked advanced personalization, explainability, and integration with conversational AI [6].

A systematic review by Jubair et al. highlighted that machine learning-based systems can achieve high accuracy in exercise correction and injury prevention, often exceeding 90% in movement analysis. However, challenges such as data variability, lack of standardization, and privacy concerns remain significant [7].

III. PROPOSED SYSTEM

➤ System Overview

The AI-Powered Personalized Coach is a browser-based intelligent fitness and dance training system that provides real-time posture correction, movement analysis, and personalized coaching. The system integrates computer vision, YOLOv5-based human detection, MediaPipe BlazePose pose estimation, rule-based biomechanical analysis, and LLM-assisted feedback to create an interactive digital coaching environment. The architecture is divided into three major layers: (1) Input Layer — captures live video streams from the user; (2) Processing Layer — performs pose estimation, landmark analysis, and movement interpretation; and (3) Output Layer — delivers corrective feedback, performance scoring, and personalized recommendations. The modular architecture enables scalability and easy integration of additional exercises, dance routines, and coaching features without requiring custom model retraining.

➤ Input Acquisition

Live video input is captured from webcams or smartphone cameras using browser-based APIs such as WebRTC and OpenCV [6]. The captured frames undergo preprocessing operations including resizing, normalization, and noise reduction before being passed to the pose estimation pipeline. The system is designed for low-latency processing to support smooth real-time feedback during exercise and dance sessions. The YOLOv5 detection model is executed through a browser-compatible deployment pipeline,

enabling real-time human detection without requiring custom model training or dedicated server-side inference. All pose analysis is performed using pretrained models without custom training, enabling efficient execution on standard consumer devices.

➤ Human Pose Estimation

The proposed system employs a two-stage computer vision pipeline consisting of YOLOv5 for human detection and MediaPipe BlazePose for landmark extraction. YOLOv5 first identifies and localizes the user within each video frame by generating a bounding box around the detected person. Although MediaPipe BlazePose includes a lightweight built-in person detector, YOLOv5 is used as an additional localization stage to improve robustness in complex environments, reduce background interference, and accurately identify the primary user. The detected region of interest is then passed to MediaPipe BlazePose, which extracts 33 body landmarks corresponding to major joints and key body points. These landmarks form the basis for posture assessment, repetition counting, movement scoring, and dance pose evaluation. This hybrid approach improves detection stability while maintaining real-time performance on consumer hardware and supports accurate movement analysis for both fitness and dance training.

➤ Feature Extraction and Movement Analysis

The extracted landmarks are processed to compute biomechanical and kinematic features such as joint angles, range of motion, posture alignment, movement stability, and repetition phases [1], [2]. Joint angles are calculated using vector geometry and cosine similarity between adjacent limb segments. The system compares observed movement patterns against predefined biomechanical rules and expert-defined thresholds to identify posture deviations and unsafe motion patterns [3]. For dance and sequential movements, temporal alignment techniques such as Dynamic Time Warping (DTW) are used to compare user motion against reference routines and generate similarity-based feedback [5], [7].

➤ Adaptive Feedback Generation

The system delivers real-time corrective feedback using a hybrid rule-based coaching framework. Feedback is provided through multiple modalities:

- Visual cues using skeletal overlays and highlighted posture errors
- Text prompts such as “Straighten your back” or “Lower slightly deeper”
- Voice guidance using browser speech synthesis for auditory coaching.

Feedback prioritization focuses first on safety-related corrections, followed by posture improvement and motivational cues. The feedback frequency dynamically adapts based on user consistency and performance trends to reduce repetitive interruptions and improve learning efficiency.

➤ Personalization and Progress Tracking

The system continuously tracks user performance

metrics such as posture accuracy, range of motion, repetition quality, and correction frequency. These metrics are used to adapt workout intensity, exercise difficulty, and coaching recommendations over time [3]. Progress summaries, performance trends, and session statistics help users monitor long-term improvement while maintaining safe exercise practices.

➤ *Data Storage and Security*

User performance metrics, session summaries, and progress data are securely stored in an SQL-based database. The system follows a privacy-conscious design by avoiding raw video storage and retaining only processed metrics and landmark-based analytical data. Data transmission is protected through HTTPS and authenticated access mechanisms to ensure user privacy and security.

➤ *System Strengths*

The proposed system offers several advantages:

- Unified support for both fitness and dance coaching
- Real-time posture correction with low-latency feedback
- Browser-based deployment without custom CNN model training
- Explainable rule-based biomechanical analysis
- Privacy-focused architecture without raw video storage
- Personalized workout, feedback, and progress tracking capabilities
- AI-assisted personalized nutrition and diet planning based on user profile, goals, preferences, and performance history
- Compatibility with consumer-grade devices and webcams.

IV. MATHEMATICAL FORMULATION

The following are the mathematical formulations which are used in our system.

➤ *Landmark Representation*

For frame t , let the pose detector return a set of landmarks: $L_t = \{l_1, l_2, \dots, l_n\}$

Where each landmark l_i is represented as: $l_i = (x_i, y_i, z_i, c_i)$

Here:

- x_i, y_i, z_i are normalized coordinates
- c_i is confidence
- n is the number of landmarks

➤ *Joint Angle Computation*

Given three landmarks A, B, and C, the joint angle at B is computed using vector geometry:

$$\theta = \arccos \left(\frac{(A - B) \cdot (C - B)}{\|A - B\| \|C - B\|} \right)$$

This formulation can be used for elbows, knees, hips, and shoulders depending on the exercise.

➤ *Range of Motion*

Let θ_{max} and θ_{min} represent the maximum and minimum observed angle over one repetition.

Let θ_{target_max} and θ_{target_min} denote exercise-specific ideal boundaries. A normalized Range-of-motion score can be expressed as:

$$ROM_score = 100 \times \text{clamp} \left(\frac{(|\theta_{max} - \theta_{min}|)}{(|\theta_{target_max} - \theta_{target_min}|)}, 0, 1 \right)$$

➤ *Tempo Score*

If a repetition has observed phase durations $d_{eccentric}$, d_{pause} , and $d_{concentric}$, and the target durations are $\tau_{eccentric}$, τ_{pause} , and $\tau_{concentric}$, then a simple tempo error may be defined as:

$$Etempo = |d_{eccentric} - \tau_{eccentric}| + |d_{pause} - \tau_{pause}| + |d_{concentric} - \tau_{concentric}|$$

Where:

$d_{eccentric}$, d_{pause} , and $d_{concentric}$ represent the observed phase durations, while $\tau_{eccentric}$, τ_{pause} , and $\tau_{concentric}$ represent the target phase durations.

➤ *Stability Score*

Let p_t denote a landmark coordinate sequence during a rep or hold. Stability can be estimated using jitter Variance:

$$J = (1/T) \sum_{(t=1 \rightarrow T)} |p_t - \bar{p}|^2$$

A lower J implies higher stability. Stability score is inversely related to jitter magnitude.

➤ *Composite Score*

A composite exercise score can be defined as a weighted sum:

$$S = w_f F + w_r R + w_t T + w_{st} St + w_{sy} Sy + w_c C + w_{sa} Sa$$

Where:

- F = form score
- R = range-of-motion score
- T = tempo score
- St = stability score
- Sy = symmetry score
- C = consistency score
- Sa = safety score and the weights satisfy:

$$\sum w_i = 1$$

➤ *Dance Similarity Score*

For dance, pose similarity between user vector u_t and reference vector r_t at checkpoint t can be modeled by cosine similarity:

$$Simt = (u_t \cdot r_t) / (|u_t| \times |r_t|)$$

An overall dance score may combine similarity and timing alignment.

V. IMPLEMENTATION AND TOOLS

➤ Development Environment

The proposed system is developed using Next.js and TypeScript for the frontend and full-stack architecture, while MediaPipe BlazePose is used for real-time pose estimation [6]. The system integrates browser-based APIs for camera access, real-time processing, and interactive feedback generation. PostgreSQL is used for storing user profiles, session history, and progress data.

➤ Pose Estimation and Motion Analysis

The system utilizes YOLOv5 for human detection and MediaPipe BlazePose for real-time pose estimation. YOLOv5 identifies the user within the camera frame, while BlazePose extracts 33 body landmarks for skeletal tracking [6]. The extracted landmarks are processed using rule-based biomechanical analysis to compute joint angles, posture alignment, range of motion, and repetition phases. These features are further used for posture correction, movement evaluation, and dance pose comparison [7].

➤ Feedback and Coaching System

Real-time corrective feedback is generated using a hybrid rule-based and AI-assisted coaching framework. The system provides:

- Visual pose overlays
- Text-based posture corrections
- Voice guidance using speech synthesis

Feedback prioritizes safety and posture correction while adapting dynamically based on user performance and consistency [7].

➤ Frontend and User Interface

The frontend interface provides live camera streaming, pose visualization, exercise tracking, and performance dashboards. Users can access workout plans, diet recommendations, dance coaching, and progress reports through an interactive and responsive web interface [7].

➤ Personalization and Progress Tracking

The system stores workout history, posture scores, repetition counts, and performance trends to generate personalized workout and nutrition recommendations. Coaching feedback and difficulty levels are adjusted according to user goals, preferences, injuries, and historical performance data.

➤ Data Security and Privacy

Only processed analytical data such as landmarks, scores, and session summaries are stored, while raw video data is discarded to preserve user privacy [6]. Secure communication protocols and authenticated access mechanisms are used to protect user information.

➤ Runtime Performance

The system is optimized for real-time browser-based execution and maintains smooth pose tracking with low-latency feedback on consumer-grade devices [6]. The architecture is designed to support scalable deployment while maintaining efficient real-time performance.

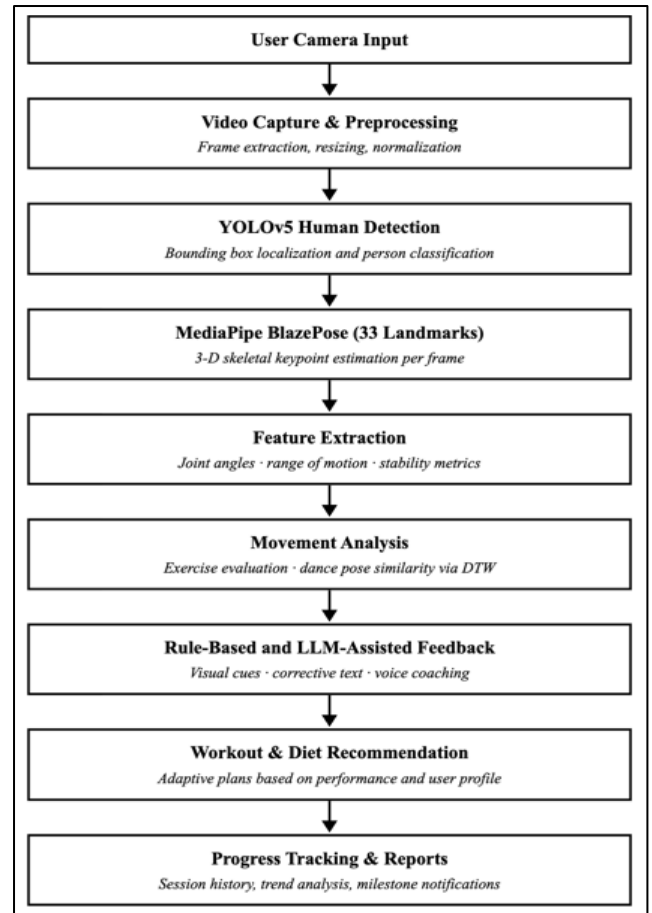


Fig 1 Overall Architecture of the Proposed AI-Powered Fitness and Dance Coaching System Integrating YOLOv5-Based Human Detection, MediaPipe BlazePose Landmark Extraction, Movement Analysis, Personalized Feedback, and Progress Tracking.

VI. EXPERIMENTS AND RESULTS

The proposed AI Fitness and Dance Coach was implemented as a browser-based web application and evaluated across multiple exercise and dance scenarios. Experimental testing demonstrated successful real-time pose estimation, posture analysis, repetition counting, and feedback generation using YOLOv5 and MediaPipe BlazePose. The system achieved approximately 95% repetition counting accuracy during internal testing for common exercises such as squats, jumping jacks, and bicep curls under normal lighting conditions. The dance coaching module successfully compared user poses against reference movements and generated similarity-based corrective feedback. The results indicate that the proposed architecture can provide responsive, explainable, and personalized fitness and dance guidance while remaining suitable for deployment on consumer-grade devices.

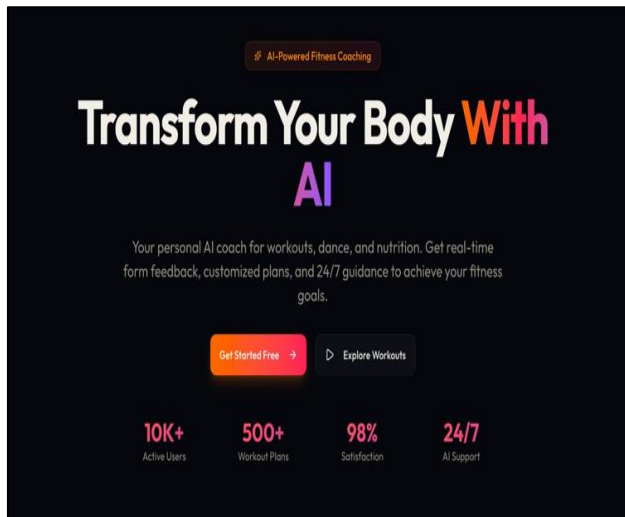


Fig 2 Home Page of the Proposed AI-Powered Fitness and Dance Coaching System.

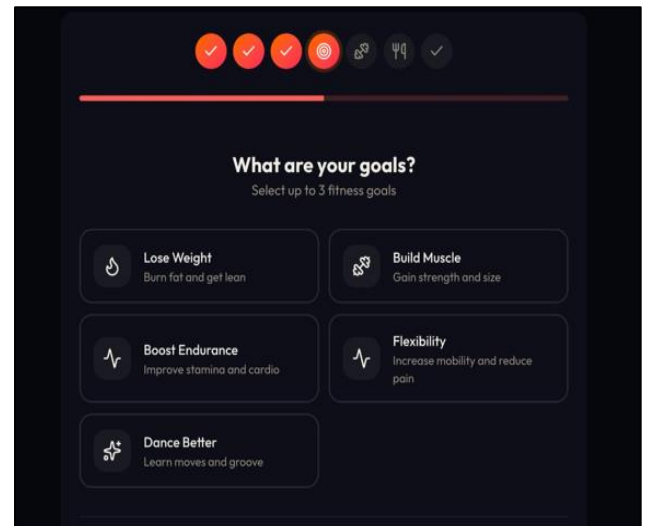


Fig 3 User Onboarding and Goal Selection Interface.

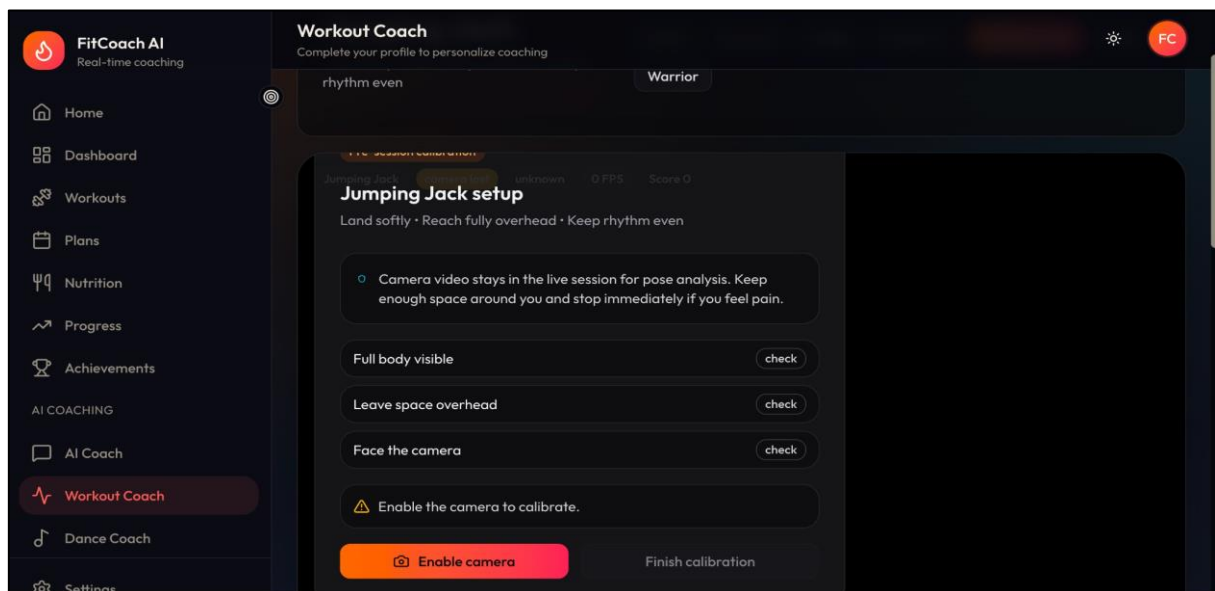


Fig 4 Real-Time Workout Coaching and Posture Analysis Module.

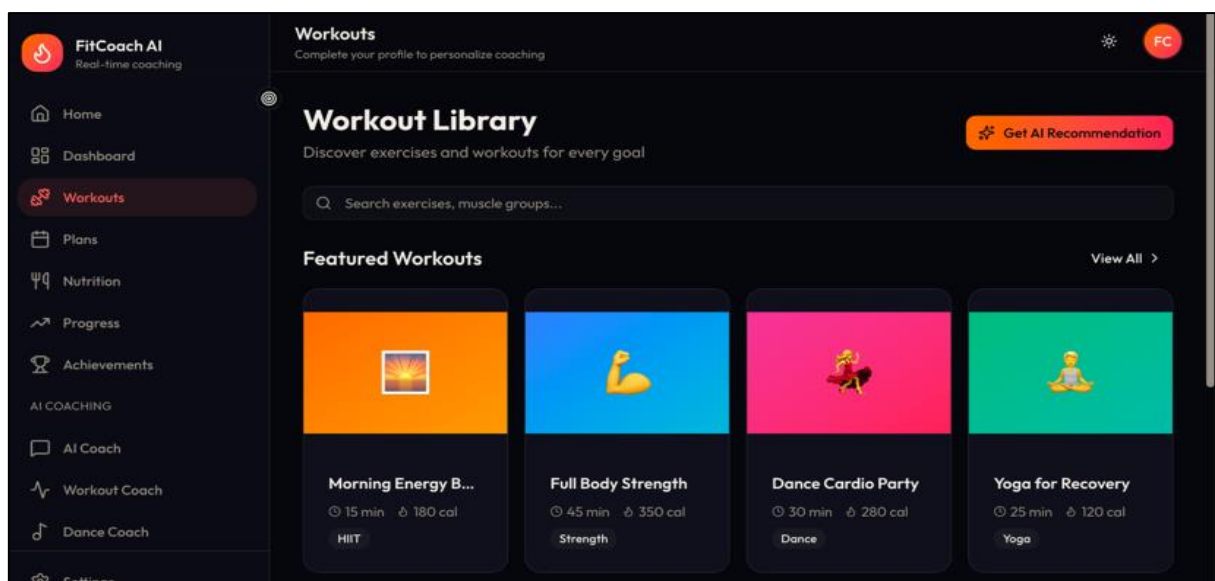


Fig 5 Personalized Workout Library and Recommendation Dashboard

VII. FUTURE SCOPE

The proposed system can be further enhanced by integrating 3D pose estimation to improve accuracy for depth-based movements and complex activities. The platform can be expanded to support a wider range of exercises and multiple dance styles with higher complexity. Future improvements may include adaptive learning techniques to

personalize feedback based on user progress over time. Additionally, integrating wearable devices can provide supplementary health metrics such as heart rate and calorie estimation. Dance evaluation can be improved by incorporating analysis of rhythm, style, and expression beyond basic pose similarity. Finally, deploying the system as a mobile application can enhance accessibility and usability in real-world scenarios.

Table 1 Module-Wise Functionality of the Proposed Fitness AI Coach System

Module	Input Used	Output Generated	Purpose
User Pose Detection	Live webcam/video frames	Body keypoints and joint positions	Tracks user movement during exercise or dance
Exercise Posture Analysis	Detected body landmarks	Correct/incorrect posture feedback	Helps users improve exercise form
Repetition Counting	Continuous body movement	Rep count and movement status	Automatically counts exercise repetitions
Injury Risk Detection	Posture angles and movement patterns	Warning for risky posture	Reduces chances of injury
Dance Pose Evaluation	User pose and reference pose	Similarity score or correction feedback	Evaluates dance accuracy
Real-Time Feedback	AI analysis result	Suggestions and corrections	Guides the user instantly
Performance Tracking	Session data	Accuracy score, reps, and improvement report	Helps users monitor fitness progress

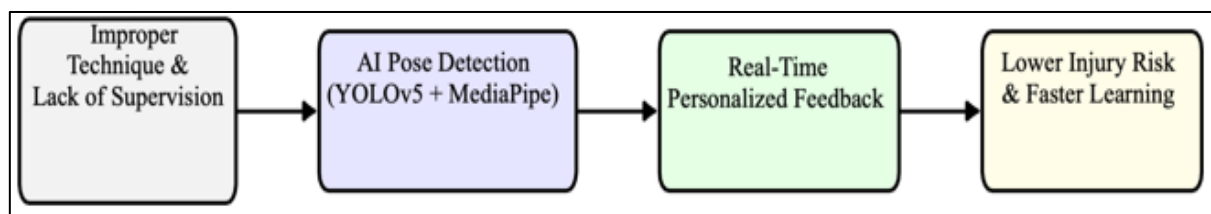


Fig 6 AI-Driven Posture Correction Pipeline from Risk Factors to Outcomes.

VIII. ACKNOWLEDGMENT

We would like to express our sincere gratitude to our project guide, faculty members, and department for their continuous guidance, encouragement, and valuable suggestions throughout the development of this project. Their technical insights and support played a significant role in helping us successfully design and implement the AI Fitness and Dance Coach system.

We also thank all team members for their dedication, collaboration, and contributions during the research, development, testing, and documentation phases of the project. Finally, we acknowledge the support of open-source technologies, research resources, and modern AI frameworks that contributed to the successful completion of this work.

IX. CONCLUSION

This paper presented an AI-powered personalized coaching system for fitness and dance training that combines real-time pose estimation, rule-based biomechanical analysis, and AI-assisted coaching within a browser-based environment. The proposed system enables posture correction, repetition counting, movement analysis, dance pose evaluation, personalized workout generation, nutrition planning, and progress tracking without requiring custom CNN model training. By leveraging pretrained pose

estimation models and explainable decision-making logic, the system provides accessible, low-cost, and interactive coaching on consumer devices. Experimental evaluation demonstrated the feasibility of delivering real-time feedback and personalized guidance while maintaining usability, interpretability, and scalability. The developed platform contributes toward the growing field of intelligent digital health and coaching systems by offering an integrated solution for both fitness and dance training.

REFERENCES

- [1]. International Journal of Creative Research Thoughts, "AI-Based Posture Correction Systems," IJCRT, vol. 13, no. 1, pp. 1–6, 2025.
- [2]. IOSR Journal of Computer Engineering, "Computer Vision for Human Pose Estimation," IOSR-JCE, vol. 27, no. 2, pp. 15–22, 2025.
- [3]. Y.-M. Ko, A. Nasridinov, and S.-H. Park, "Real-Time AI Posture Correction for Powerlifting Exercises Using YOLOv5 and MediaPipe," IEEE Access, vol. 12, pp. 195830–195842, 2024.
- [4]. L. Xiang and G. Gao, "Perceptual Feature Integration for Sports Dancing Action Scenery Detection and Optimization," IEEE Access, vol. 11, pp. 112345–112356, 2023.
- [5]. R. Nair, A. Yang, and H. Zhou, "AI-Powered Dance Coaching via Pose Estimation, Vision Transformers,

- and DTW,” in Proc. IEEE Int. Conf. on AI and Computer Vision, 2025, pp. 210–218.
- [6]. S. Gurbuxani, L. Gupta, and K. Madan, “Virtual Fitness Trainer Using Artificial Intelligence,” *Int. J. Eng. Res. Technol. (IJERT)*, vol. 13, no. 2, pp. 45–50, 2024.
- [7]. “Machine Learning for Real-Time Exercise Correction and Injury Prevention: A Systematic Review,” *Sensors*, vol. 24, no. 3, pp. 1120–1135, 2024.