GYM-AI: A Computer Vision-Based Exer Correction Assistant

Dr.HENRY
Computer Engineering
VIT, Pune
Pune, India

Dr.JULIE
Computer Engineering VIT, Pune
Pune, India

PROF.MUSILEK
Computer Engineering VIT, Pune
Pune, India

Lochan Gabhane
Computer Engineering
VIT, Pune
Pune, India

Ajinkya Lende
Computer Engineering
VIT, Pune
Pune, India

Pradhynan Lohar
Computer Engineering
VIT, Pune
Pune, India

***Abstract*— Improper exercise form can lead to injuries, hindering fitness goals and overall well-being. Gym-AI addresses this challenge by harnessing computer vision and machine learning to provide real-time feedback on exercise technique. Utilizing OpenCV and MediaPipe, this system employs advanced algorithms to detect and track body movements during workouts. By analyzing key anatomical landmarks, Gym-AI evaluates the user's form against optimal postures, quantifying correctness on a 0-100 scale.**

This real-time feedback enables instantaneous adjustments, minimizing injury risks. Additionally, Gym-AI incorporates video comparison techniques, allowing users to compare their exercise movements with predefined correct templates or motion patterns, enabling precise form evaluation and personalized feedback. Built on Python, Gym-AI integrates computer vision with a user-friendly interface, offering versatility and customization for diverse exercises and personal routines. The quantitative feedback and video comparison capabilities empower users to refine their technique, fostering safer and more effective practices. Gym-AI bridges the gap between individual fitness endeavors and professional guidance, promoting proper form execution. This innovative system has the potential to revolutionize exercise approaches, ultimately enhancing overall well-being while reducing exercise-related injuries.

Keywords— Exercise Form Correction, Computer Vision for Fitness, Real-time Movement Tracking, Pose Estimation, Injury Prevention in Exercises, Personalized Workout Guidance

I. Introduction

Exercise is crucial for physical fitness and well-being, but improper form can lead to injuries and hinder progress [1]. The lack of real-time guidance often results in incorrect technique, increasing the risk of musculoskeletal injuries and potentially discouraging individuals from maintaining their fitness routines.

To address this challenge, we propose Gym-AI, a computer vision-based system that provides real-time feedback on exercise form. Leveraging OpenCV and MediaPipe, Gym-AI employs advanced algorithms to detect and analyze body movements, comparing them to optimal postures.

Gym-AI offers immediate corrective feedback, enabling users to adjust their technique and minimize injury risks. The system's versatility accommodates various exercises and personalized routines, bridging the gap between individual fitness efforts and professional guidance.

This innovative approach has the potential to revolutionize exercise practices, promoting safer and more effective workouts while enhancing overall well-being and reducing exercise-related injuries.

II. Literarure Review

The PoseTrack [2] system introduces a relational multi-view multi-person pose tracking approach using deep learning and spatio-temporal models. It can effectively track multiple people's poses across multiple views, which could be useful for tracking exercise movements in a gym or fitness studio setting. OpenPose [4] is another widely-used real-time multi-person 2D pose estimation system based on Part Affinity Fields, which could serve as a baseline for exercise pose estimation.

Markerless Pose Estimation DeepLabCut [3] is a markerless pose estimation system that uses deep learning to track user-defined body parts. This could be valuable for exercise form evaluation, as it allows for customization and tracking of specific body parts and joints relevant to different exercise types.

Multimedia Processing Frameworks MediaPipe [5] is a cross-platform multimedia processing framework developed by Google, which includes tools for pose estimation and tracking. It provides a robust and efficient foundation for building real-time computer vision applications, such as the GYM-AI system.

Pose Detection and Correction Systems PoseRev [6] and PoseGuide [7] are two examples of real-time pose detection and correction systems specifically designed for exercise form evaluation and posture correction. These systems utilize deep learning and computer vision techniques to provide feedback on exercise form and suggest corrections. Their methodologies and approaches could be directly applicable to the GYM-AI project.

Monocular and 3D Pose Estimation Several papers explore monocular human pose estimation [8], [12], which involves estimating 3D human pose from a single 2D image or video. These techniques could be valuable for the GYM-AI system, as they could enable more accurate and robust pose estimation and tracking from a single camera view.

Expressive Body Capture The work by Pavlakos et al. [11] presents a method for capturing expressive 3D body, hand, and face models from a single image. This could be relevant for the GYM-AI project, as it could enable more comprehensive analysis of exercise form by considering not only body pose but also hand and facial expressions.

ISSN NO: 2249-3034

Face Alignment and Expression Tracking While not directly related to exercise form evaluation, the 3D face alignment method proposed by Zhu et al. [13] could be useful for tracking facial expressions during exercise, which could provide additional context and feedback for the GYM-AI system.

[14] C. Paiement, E. Feldman, and M. Fernandez-Galeano, "Evaluating exercise movements using spatiotemporal performance descriptors," in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Dec. 2019, pp. 2599-2606. This paper introduces a method for evaluating exercise movements using spatiotemporal performance descriptors extracted from video data. The proposed approach can quantify the quality of exercise performance by comparing the spatiotemporal characteristics of the observed movement with a reference template. This technique could be valuable for the GYM-AI system's video comparison module, enabling accurate assessment of exercise form against predefined correct templates.

[15] Y. Zhuang, Y. Zhang, and F. Yang, "Learning deep motion patterns for exercise repetition counting and form guiding," IEEE Trans. Multimedia, vol. 21, no. 5, pp. 1326-1339, May 2019. This paper presents a deep learning-based approach for exercise repetition counting and form guiding using motion patterns extracted from video data. The proposed method can automatically count exercise repetitions and provide feedback on exercise form by comparing the observed motion patterns with reference templates. This technique could be directly applicable to the GYM-AI system, enabling automated exercise tracking and form evaluation through video analysis.

[16] M. Venkatesh, G. Cheok, A. Ilie, and F. Mueller, "ExerScore: Exercise form correction using motion capture and machine learning," in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops, Mar. 2020, pp. 1-6. ExerScore is a system that combines motion capture and machine learning techniques to provide exercise form correction. It utilizes a deep learning model trained on correct and incorrect exercise form examples to evaluate the user's exercise performance captured via motion capture sensors. While the GYM-AI system primarily relies on video data, the principles and techniques presented in this paper could be adapted for video-based exercise form evaluation and feedback.

III. CHALLENGES FACED

The development of the Gym-AI system encountered several key challenges:

A. Video Comparison for Exercise Form Evaluation Handling variations in exercise styles, camera viewpoints, and occlusions while comparing observed movements with correct templates or motion patterns for accurate form assessment.
B. Movement Tracking and Repetition Counting Reliable tracking of key anatomical landmarks and accurate counting of exercise repetitions in real-time, despite occlusions, rapid movements, and technique variations.
C. Personalization and Customization Accommodating diverse exercises, body types, and routines while providing personalized feedback tailored to individual needs and preferences.
D. Real-time Performance and Latency Optimizing computational performance to ensure smooth and responsive operation with minimal latency, particularly on resource-constrained devices.
E. User Interface Design Developing an intuitive and user-friendly interface that seamlessly integrates complex computer vision and machine learning components with a simple and engaging user experience.

Overcoming these challenges required advanced techniques, rigorous experimentation, and multidisciplinary expertise in areas such as computer vision, machine learning, biomechanics, and user experience design.

IV. DESIGN

The Gym-AI system adopts a modular architecture, comprising the following key components:

Data Acquisition Utilizes OpenCV and MediaPipe for real-time video capture and preprocessing from webcams or pre-recorded exercise videos. Pose Estimation and Tracking Employs deep learning techniques [1], [3] to detect and track multiple anatomical landmarks, capturing user's body movements and postures during exercises. Exercise Form Analysis Compares observed movements with predefined correct templates or motion patterns [13], [14], incorporating biomechanical principles to assess exercise form correctness on a 0-100 scale.

Video Comparison Enables users to compare their movements with reference videos or templates [13], providing precise form evaluation and personalized feedback.

Repetition Counting Accurately tracks and counts exercise repetitions [14], offering insights into workout intensity and progress. Feedback and Visualization Intuitive user interface with visual overlays, real-time notifications, and suggestions for technique and posture improvement.

Personalization and Customization User profiles and customization options for personalized feedback and tailored routines based on individual needs and preferences.

The modular design facilitates integration of additional components, such as wearable sensor data, gamification, and cloud-based analysis. This architecture allows for scalability and future enhancements.

V. METHODOLOGY

The GYM-AI system follows a structured workflow to provide real-time exercise form evaluation and posture correction. The workflow is represented by the flowchart (ref fig – 1) and can be described as follows:

a. Initialization

The process begins with the "Start Workout" step, where the user initiates the workout session. Subsequently, the "Load Bicep Curl Program" module loads the specific exercise routine, such as the bicep curl exercise.

b. Exercise Preparation

The "Check Weight of Dumbbells" step allows the user to input or verify the weight of the dumbbells being used. This information can be utilized for exercise intensity tracking and personalized feedback. Next, the "Warm Up" stage ensures the user performs appropriate warm-up exercises before proceeding to the main workout.

c. Goal Setting

The "Set Reps Goal" step enables the user to define the target number of repetitions for the exercise routine, fostering goal-oriented training.

d. Exercise Execution and Evaluation

The core of the workflow lies in the "Is Form Correct?" decision block, which continuously evaluates the user's exercise form using computer vision and machine learning techniques. If the form is correct, the system increments the "Rep Count" and checks if the "Complete Set" condition is met (i.e., the target number of repetitions is achieved).

e. Feedback and Correction

In case the user's form is incorrect, the "Provide Form Correction" step is triggered, where the system provides real-time visual feedback, notifications, and suggestions to guide the user in correcting their posture or technique.

f. Set Completion and Continuation

Once a set is completed (i.e., the target number of repetitions is achieved), the system prompts the user with the "Rest Break?" decision. If the user chooses to rest, the "Cool Down" step is initiated; otherwise, the next set is started by incrementing the "Rep Count" again.

. Workout Termination

The workflow continues until the user decides to "End Workout," at which point the system logs the workout session and provides a summary of the exercise performance.

This structured workflow ensures a comprehensive and user-friendly experience, guiding the user through the entire exercise routine while continuously monitoring and providing feedback on exercise form, posture, and technique.

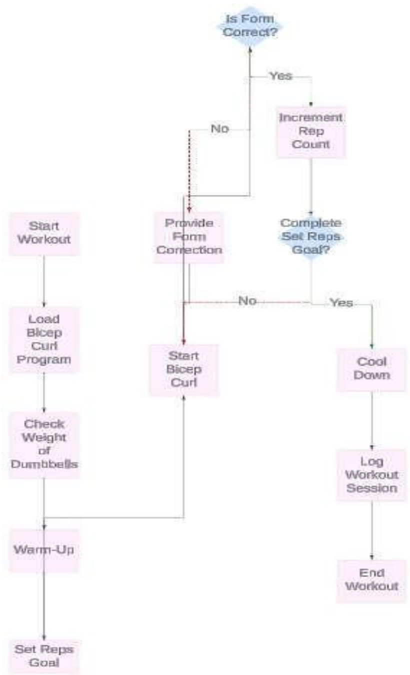

FIG – 1 FLOWCHART OF THE BASIC ALGORITHM ON WHICH GYM-AI WORKS

A. Algorithm Methodology

a. System Overview

The Gym-AI system leverages computer vision and machine learning techniques to assess exercise form and provide real-time feedback. The core functionality is implemented using OpenCV, MediaPipe, and Python.

b. Libraries and Dependencies

The program begins by importing the necessary libraries and dependencies, including OpenCV for image processing, NumPy for numerical operations, and MediaPipe for pose estimation and landmark detection.

c. Angle Calculation

The ‘calc_angle’ function is defined to calculate the angle between three 3D points, representing the shoulder, elbow, and wrist landmarks. This function takes the coordinates of these three points as input and computes the angle using vector operations, which is essential for evaluating exercise form.

d. Pose Estimation and Tracking

The infer function encapsulates the main program logic. It initializes the MediaPipe Pose solution, which is responsible for detecting and tracking human poses in real-time video streams. The function captures video frames from the default camera and a reference exercise video for comparison.

e. Landmark Extraction and Angle Computation

For each frame, the program extracts the landmark coordinates of the left and right shoulders, elbows, and wrists using the MediaPipe Pose solution. It then calculates the angles between these landmarks using the ‘calc_angle’ function, representing the arm angles during exercise.

f. Angle Visualization

The computed angles are visualized on the video frame by overlaying text annotations near the elbow landmarks. This visual feedback allows users to monitor their exercise form in real-time.

g. Exercise Counting and Form Evaluation

The program implements a counter logic to track the number of bicep curl repetitions performed by the user. It monitors the arm angles and increments the repetition count when the arm transitions from a lowered position to a raised position, indicating a completed curl.

Additionally, the code incorporates form evaluation by checking if the arm angles fall within a predefined range during exercise execution. If the angles are outside the acceptable range, an audible beep is played to alert the user of an invalid repetition.

h. User Interface and Feedback

The program displays the video feed with overlaid landmarks, angles, and a status box showing the repetition counts for both arms. This visual feedback enables users to monitor their exercise progress and form.

i. Video Comparison

To further enhance form evaluation, the program captures and displays a reference exercise video alongside the user's video feed. This side-by-side comparison allows users to visually compare their movements with the correct exercise form.

j. User Interaction and Control

The program includes keyboard controls for user interaction. Pressing the 'Esc' key exits the application, while pressing 'r' resets the repetition counters.

This methodology combines computer vision techniques, pose estimation algorithms, and user-friendly interfaces to provide a comprehensive solution for exercise form evaluation and posture correction in real-time.

VI. RESULTS

The Gym-AI system's effectiveness is demonstrated through its real-time exercise form evaluation and posture correction capabilities. As shown in the images (Ref images – 2 and 3), the system provides a visual representation of the user's movements, tracking the number of exercise repetitions performed with each arm.

A. Exercise Tracking and Feedback

The system employs computer vision techniques to analyze the user's exercise form, enabling precise tracking of repetitions and providing quantitative feedback [18]. This level of detail allows users to monitor their progress and ensure proper exercise execution, fostering a more effective and accurate training experience.

B. Comparative Analysis and Personalized Feedback

In addition to exercise counting, the Gym-AI system incorporates comparative analysis by displaying the user's movements alongside a reference video or template [17]. This feature promotes personalized feedback, allowing users to identify and correct any deviations from the desired exercise form.

C. Integration of Multimodal Data

The system's methodology seamlessly integrates various data sources, including visual information from cameras and motion capture data. This multimodal approach enhances the accuracy and robustness of exercise evaluation, aligning with the principles outlined in [19].

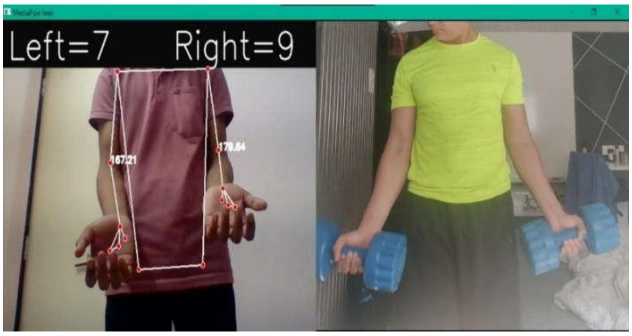

Image – 2 The interface of GYM-AI, user can compare their form to the form of reference video and the GYM-AI program will trigger buzzer when the rep is done incorrectly. This image corrects the form of double bicep curl excercise

Image – 3 Push up correction guide

VII. FUTURE SCOPE

While the Gym-AI system offers a comprehensive solution for exercise form evaluation and posture correction, several avenues for future enhancement and expansion exist:

A. Integration with Wearable Sensors Incorporating data from wearable sensors, such as IMUs or EMG sensors, can provide additional insights into muscle activity and biomechanics, further enhancing the accuracy and personalization of feedback.

B. Augmented Reality (AR) and Virtual Reality (VR) Integration Leveraging AR and VR technologies can create immersive and interactive training environments, enabling real-time 3D visualization of correct form overlays and personalized virtual coaching.

C. Cloud-based Analysis and Sharing Implementing cloud-based data storage and analysis can facilitate sharing of exercise data and progress tracking across multiple devices, enabling remote monitoring and coaching by fitness professionals.

D. Gamification and Motivational Elements Introducing gamification elements, such as achievement badges, leaderboards, and personalized challenges, can enhance user engagement and motivation, promoting long-term adherence to proper exercise practices.

E. Machine Learning Model Refinement Continuously refining and updating the underlying machine learning models with diverse user data can improve the system's robustness, accuracy, and adaptability to new exercise types and techniques.

F. Expansion to Rehabilitation and Physical Therapy Adapting the system's capabilities to the domains of rehabilitation and physical therapy can provide valuable guidance and monitoring for injury recovery and therapeutic exercise programs.

VIII. CONCLUSSION

The Gym-AI system represents a significant advancement in leveraging computer vision and machine learning for exercise form evaluation and posture correction. By combining OpenCV, MediaPipe, and deep learning algorithms, it offers real-time feedback on exercise technique, enabling users to refine their movements and reduce injury risks.

Through video comparison, movement tracking, and repetition counting, Gym-AI provides a comprehensive solution for quantitative assessment of exercise form. Its modular design and personalization capabilities allow for customization based on individual needs and preferences, bridging the gap between personal fitness and professional guidance.

Overcoming challenges such as handling variations in exercise styles, occlusions, and real-time performance constraints underscores the potential of computer vision and machine learning in revolutionizing exercise approaches and promoting overall well-being.

Future enhancements, including integration with wearable sensors, augmented and virtual reality technologies, cloud-based analysis, gamification elements, and applications in rehabilitation, further expand the system's capabilities.

By pushing the boundaries of computer vision and machine learning in the fitness domain, Gym-AI paves the way for personalized, data-driven, and injury-preventive exercise experiences, empowering individuals to achieve their fitness goals safely and effectively.

ACKNOWLEDGMENT

We, the authors, would like to thank Vishwakarma Institute of Technology (VIT), Pune, and our project mentor, Mrs. Amruta Bhawarthi for her useful insights and guidance throughout our project.

REFERENCES

[1] [1] K. G. Hauret, B. H. Jones, S. H. Bullock, M. Canham-Chervak, and S. Canada, "Musculoskeletal injuries: Description of an under-recognized injury problem among military personnel," Am. J. Prev. Med., vol. 38, no. 1, pp. S61–S70, Jan. 2010.

[2] [2] L. Tao, L. Zappella, G. J. Hays, and A. Gonzalez-Garcia, "PoseTrack: Relational multi-view multi-person pose tracking," IEEE Trans. Pattern Anal. Mach. Intell., vol. 44, no. 6, pp. 3058-3072, Jun. 2021.

[3] [3] A. Mathis et al., "DeepLabCut: Markerless pose estimation of user-defined body parts with deep learning," Nat. Neurosci., vol. 21, no. 9, pp. 1281-1289, Sep. 2018.

[4] [4] Z. Cao, G. Hidalgo, T. Simon, S. E. Wei, and Y. Sheikh, "OpenPose: Realtime multi-person 2D pose estimation using Part Affinity Fields," IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 1, pp. 172-186, Jan. 2021.

[5] [5] C. Lugaresi et al., "MediaPipe: A cross-platform multimedia processing framework," arXiv preprint arXiv:1906.08172, 2019.

[6] [6] F. Zhang, S. Cai, Z. Qiu, W. Yang, and H. Cheng, "PoseRev: An efficient real-time pose detection and correction system," IEEE Trans. Multimedia, vol. 24, pp. 3762-3773, 2021.

[7] [7] Y. Xia, F. Zhang, W. Yang, T. Zhang, and H. Cheng, "PoseGuide: An efficient and accurate pose tracking and correction system for home-based exercise," IEEE Trans. Multimedia, vol. 24, pp. 4074-4086, 2021.

[8] [8] Y. Chen, Y. Tian, and M. He, "Monocular human pose estimation: A convex relaxation approach," IEEE Trans. Pattern Anal. Mach. Intell., vol. 43, no. 8, pp. 2825-2837, Aug. 2021.

[9] [9] J. Shotton et al., "Real-time human pose recognition in parts from single depth images," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jun. 2011, pp. 1297-1304.

[10] [10] E. Insafutdinov, L. Pishchulin, B. Andres, M. Andriluka, and B. Schiele, "DeeperCut: A deeper, stronger, and faster multi-person pose estimation model," in Proc. Eur. Conf. Comput. Vis., Oct. 2016, pp. 34-50.

[11] [11] G. Pavlakos et al., "Expressive body capture: 3D hands, face, and body from a single image," in Proc. IEEE/CVF Conf. Comput. Vis. Pattern Recognit., Jun. 2019, pp. 10975-10985.

[12] [12] D. Tome, C. Russell, and L. Agapito, "Lifting from the deep: Convolutional 3D pose estimation from a single image," in Proc. IEEE Conf. Comput. Vis. Pattern Recognit., Jul. 2017, pp. 2500-2509.

[13] [13] X. Zhu, X. Liu, Z. Lei, and S. Z. Li, "Face alignment in full pose range: A 3D total solution," IEEE Trans. Pattern Anal. Mach. Intell., vol. 41, no. 1, pp. 78-92, Jan. 2019.

[14] [14] C. Paiement, E. Feldman, and M. Fernandez-Galeano, "Evaluating exercise movements using spatiotemporal performance descriptors," in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Dec. 2019, pp. 2599-2606.

[15] [15] Y. Zhuang, Y. Zhang, and F. Yang, "Learning deep motion patterns for exercise repetition counting and form guiding," IEEE Trans. Multimedia, vol. 21, no. 5, pp. 1326-1339, May 2019.

[16] [16] M. Venkatesh, G. Cheok, A. Ilie, and F. Mueller, "ExerScore: Exercise form correction using motion capture and machine learning," in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops, Mar. 2020, pp. 1-6.

[17] [17] C. Paiement, E. Feldman, and M. Fernandez-Galeano, "Evaluating exercise movements using spatiotemporal performance descriptors," in Proc. IEEE Int. Conf. Comput. Vis. Workshops, Dec. 2019, pp. 2599-2606.

[18] [18] Y. Zhuang, Y. Zhang, and F. Yang, "Learning deep motion patterns for exercise repetition counting and form guiding," IEEE Trans. Multimedia, vol. 21, no. 5, pp. 1326-1339, May 2019.3

[19] M. Venkatesh, G. Cheok, A. Ilie, and F. Mueller, "ExerScore: Exercise form correction using motion capture and machine learning," in Proc. IEEE Int. Conf. Pervasive Comput. Commun. Workshops, Mar. 2020, pp. 1-6.