Received:30/07/2024 Revised: 28/12/2024 Accepted:16/02/2025

Journal of Emerging Computer Research and Applications (ISSN: 3048-7773)
Volume 2/Issue 1/ Jan-June 2025
Research Article

AI Powered Interactive Workout Assistant for Form Detection & Correction

Siddhesh Dige[1], Abhishek Bapat[2], Tanisha Chitnis[3], Shlok Dalvi[4], Taruna Sharma[5]
Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning)
A. P. Shah Institute of Technology, Thane, Maharashtra
Email Id: siddheshsdige@gmail.com, bapatabhishek4@gmail.com, tanishachitnis27@gmail.com, shlokdalvi1013@gmail.com, tmsharma@apsit.edu.in
Corresponding Author ORCID: 0009-0009-5393-3430

Abstract

The importance of fitness and gym training has increased over the years. The need for personal trainers has also increased, but due to their high cost, it is not feasible to hire a trainer. To overcome this problem, we introduce the AI Powered Interactive Workout Assistant for Form Detection & Correction. The AI Powered Interactive Workout Assistant for Form Detection & Correction project is an innovative initiative that combines artificial intelligence and fitness technology. This project aims to revolutionize personal fitness journeys by applying advanced machine learning algorithms and leveraging the capabilities of OpenCV, a robust computer vision library. Building on the foundation of AI development, the project envisions a dynamic virtual assistant capable of delivering personalized exercise recommendations and providing immediate feedback. Integration with OpenCV enables the assistant to analyze real-time video data, allowing for precise evaluation of form and adjustment of exercises. Additionally, the AI companion extends its impact beyond the gym, offering comprehensive wellness advice based on individualized health data, dietary habits, and sleep patterns. We use machine learning algorithms such as logistic regression, ridge classifier, random forest classifier, and gradient boosting Classifier. This summary encapsulates the project's objective to enhance physical fitness and overall well-being, envisioning a future where AI-driven fitness companions serve as essential partners in our pursuit of a healthier lifestyle.

Keywords: OpenCV, Classifier, Gym Assistant, fitness Technology, AI-driven

Introduction

In recent years, the integration of artificial intelligence (AI) and fitness has ushered in a new era of personalized exercise guidance and training. This project, centered around the development of an AI gym assistant, represents a significant leap forward in leveraging machine learning algorithms to enhance the fitness journey for individuals across diverse demographics. Combining the power of AI with the familiarity of gym settings, this endeavor aims to revolutionize how we approach and optimize our workouts. The concept of a virtual AI gym assistant stems from a rich history of technological advancements in both AI and fitness-related applications. With roots in early AI assistants like ELIZA in the 1960s, which marked the inception of natural language interactions with computers, and the proliferation of personal digital assistants (PDAs) in the 1980s and 1990s, the stage was set for the convergence of AI and personal devices. The new millennium brought forth chatbots, paving the way for more interactive and intuitive user interfaces, while the 2010s witnessed the transformative integration of wearable technology and fitness apps, democratizing access to expert fitness guidance. This project endeavors to build upon this legacy, taking inspiration from the trajectory of AI's evolution. By fusing cutting-edge machine learning techniques with real-time data processing, the AI gym assistant aspires to become a dynamic, interactive companion for individuals seeking to optimize their fitness routines. Far surpassing static

[1] Corresponding Author: Siddhesh Dige, Department of Computer Science and Engineering (Artificial Intelligence and Machine Learning), A. P. Shah Institute of Technology, Thane, Maharashtra, India Email Id: siddheshsdige@gmail.com Orcid: 0009-0009-5393-3430

[2] Cite As: Siddhesh Dige , Abhishek Bapat, Tanisha Chitnis, Shlok Dalvi and Taruna Sharma (2025), AI Powered Interactive Workout Assistant for Form Detection & Correction, *Journal of Emerging Computer Research and Applications*, 2(1), 11-20.

exercise routines, this assistant will have the capability to adapt and evolve alongside the user, tailoring recommendations based on individual goals, fitness levels, and even real-time biometric feedback. Beyond the confines of the gym, this AI-driven companion has the potential to extend its influence into broader realms of wellness. This holistic approach reaffirms the project's commitment to not only enhancing physical fitness but also nurturing overall well-being. As we embark on this innovative endeavor, it is essential to recognize that this project builds on the foundation of extensive research and development in AI, machine learning, and fitness technology. With a forward-looking perspective, this AI gym assistant project seeks to redefine how we engage with our personal fitness goals, offering a glimpse into a future where AI-driven fitness companions become indispensable partners in our pursuit of a healthier, more vibrant life.

Literature Survey

In the dynamic realm of artificial intelligence, the evolution of a virtual AI gym assistant signifies a fusion of leading-edge technologies and a growing focus on personalized health and fitness. From early pioneers like ELIZA in the 1960s to the emergence of personal digital assistants (PDAs) in the 1980s and 1990s, AI companionship has made strides. The 2010s brought wearable tech and fitness apps, democratizing access to professional fitness guidance. Simultaneously, AI and natural language processing improved, enabling conversational fluency in virtual assistants like Amazon Echo and Google Home. In healthcare and fitness, AI algorithms play a prominent role in areas such as personalized dietary planning and real-time health monitoring. In a hypothetical post-2021 scenario, a virtual AI gym assistant integrates these trends, becoming an interactive, personalized coach. Wearable integration ensures real-time feedback, form correction, and dynamic adjustments, creating a responsive training environment. Beyond the gym, this virtual AI companion leverages individual health metrics, dietary habits, and sleep patterns for comprehensive wellness recommendations.

This paper explains how to install the software and gets you started with simple examples with single images and then onto video and working with the all-important graphical user interface (GUI). He addresses the education market by providing an extensive set of exercises at the end of each chapter. The degree of difficulty for these exercises is not entirely uniform, but it is evident that the authors have been thoughtful in their respective choices of exercises. OpenCV offers a free and easy way for people to get started in computer vision. It creates a way to grow the developer community and encourages innovation in a space where many of the algorithms and methods for computer vision systems are locked behind corporate and R&D laboratory doors. We utilized this technology, which would provide a real-time camera feed. [1]

Four different pose estimation difficulties are explained in this work. Solutions for the over-constrained 2D-2D and 3D-3D pose estimation problems are provided using closed-form least squares methods. For the 2D-perspective projection-3D posture estimate problem, a globally combined iterative approach is provided. Furthermore, a strong solution and a reorganized straight solution to the 2D-perspective-projection pose-estimation problem are provided. When the signal-to-noise ratio (SNR) is less than 40 dB, precise revolution deduction and interpretation with loud data may necessitate comparing point information sets with hundreds of comparing point sets, according to data from recreational tests consisting of millions of trials with varying numbers of sets of comparing focuses and changing SNRs with either Gaussian or uniform commotion. The test also indicates that the vigorous method can stifle the botched information that comes from exceptions or jumbled focuses. We utilized this 2D pose estimation technique to extract body landmarks. [2]

This paper explains a benchmark suite for robots using reinforcement learning. Over time, learning-based control and support learning have gained substantial popularity, along with the security considerations that are important for real-world robot arrangements. In any case, to sufficiently gauge the advance and appropriateness of modern learning, we require the apparatuses to evenhandedly compare the approaches proposed by the control and fortification learning communities. We demonstrate how to use a safe-control gym to quantitatively compare the information productivity, security, and control execution of various approaches from the domains of traditional control, learning-based control, and reinforcement learning in order to support our proposal and foster greater collaboration among the research communities. We utilized his learnings in our model, which efficiently detects a moving human body while performing exercises. [3]

This paper explains the development and study of statistical algorithms that can effectively generalize and thus perform tasks without explicit instructions. The fundamental purpose of the ML strategy is to empower computers to learn without human assistance. ML is primarily separated into three categories: specific, administered, unsupervised, and semi-supervised learning approaches. Administered calculations require people to grant input and the required yield; in addition to giving input, almost all forecast precision is prepared. Unsupervised learning approaches are differentiated from directed learning approaches, which do not require any preparation. But directed learning approaches are more difficult than unsupervised learning approaches. This paper surveys the administered learning approaches that are broadly utilized in the information classification process. The methods are surveyed on the premise of point, technique, focal points, and

Frequency: Half Yearly

impediments. At last, the perusers can get an outline of administered ML approaches in terms of information classification. We utilized these classification algorithms on top of the body landmarks obtained from the Mediapipe landmarker model in order to detect the human body performing exercises. [4]

This paper explains an AI-based workout assistant and fitness guide system. Health and fitness play a vital role in our day-to-day lives. This can be attained in several different ways, of which exercising is one. Performing exercise can help us maintain very good health, but only if carried out properly and in a defined manner; otherwise, the repercussions may have adverse effects on our body. To handle this issue, we have made a framework that keeps track of body developments and gives us the number of redundancies performed, if performed inside the establishment of the model. The system also provides audio instruction to the user when performing the exercise inappropriately, and with the assistance of the user's physical measurements and his or her diet, the system is able to keep track of the user's calorie intake and recommend a certain amount of calorie intake to be followed in order to achieve a normal body mass index and stay fit. The proposed system uses the Mediapipe Pose Estimation Model to track body movements while performing the exercise.

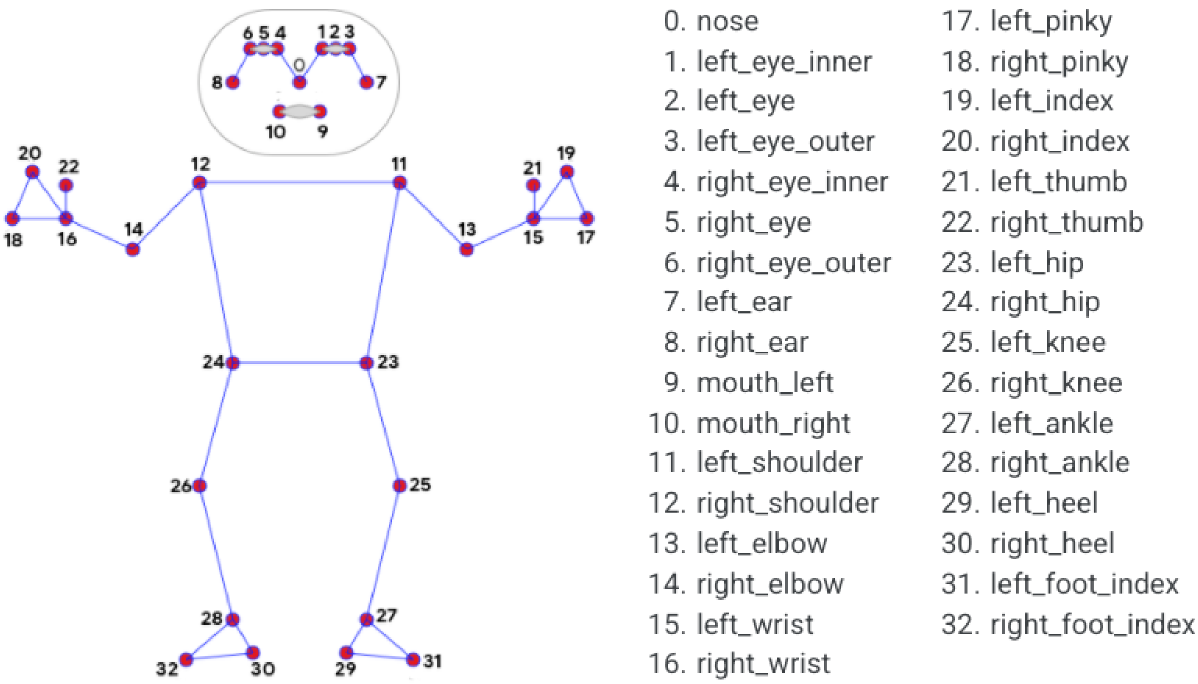

Fig 1: Body Landmarks

Source: Project source code (Original)

We utilized the above body points to calculate angles between joints, and these angles helped us build the logic of our repetition counter. [5]

This paper explains a posture estimation system using Mediapipe and OpenCV for precise workout analysis and injury prevention. It combines Mediapipe's models with OpenCV's processing functions to analyze body posture in real-time, offering feedback and corrective suggestions. While initially focusing on bicep curls, the system is adaptable to other exercises and performs effectively in varied conditions, making it suitable for deployment as an AI gym trainer to enhance form and technique and reduce injury risks. [6]

The whisper model, a popular method of communication for exchanging confidential information or avoiding disturbing others, is explained in this essay. Whisper's voice design would work incredibly well in open-office or open-run circumstances as a human-handheld/computer combination. Unfortunately, the way communication is developing now is transcendentally focused on private (fair) conversation, and it collapses totally when whispering is exposed. The need for publicly available massively decoded whispered chat corpora is one of the main obstacles to the development of productive whisper affirmation engines. This study offers nearly two approaches that rely on a small sample size of untranscribed whisper tests to approximate the beat of whisper-like (pseudo-whisper) utterances from commercially feasible private conversation recordings. [7]

This research argues that typical automated speech recognition (ASR) systems trained on neutral speech perform much worse when whispered due to the large changes in acoustic features between neutral and hushed speech. This study first analyzes the acoustic characteristics of whispered speech, addresses the issues of whispered discourse acknowledgment in bungled conditions, and then proposes a modern vigorous cepstral highlights and preprocessing approach based on a profound denoising autoencoder (DDAE) that upgrades whisper acknowledgment. This approach aims to deeply analyze this mismatched train/test situation and develop an efficient way for whisper recognition. The experimental results demonstrate that Teager-energy-based cepstral highlights—TECCs in particular—are more powerful and effective whisper descriptors than traditional Mel-frequency cepstral coefficients (MFCC). [8]

This essay provides information on whispered speech. Whispering is an essential kind of human communication, but there isn't any end-to-end recognition for it that has been documented, most likely because there isn't enough easily available data on whispered communication. In light of the unique qualities of whispered speech and the paucity of available data, we offer various methods for end-to-end (E2E) whispered voice identification in this study. In order to significantly better capture the high-frequency structures of whispered discourse, this combines a frequency-weighted SpecAugment arrangement with a frequency-divided CNN highlight extractor. Additionally, it employs a layer-wise exchange learning approach to pre-train a model with normal or normal-to-whispered converted speech, then fine-tune it with whispered speech to close the gap between whispered and normal speech. [9]

This paper explains how automatic speech recognition (ASR) systems are trained on acoustic representations of neutral speech. In this work, we investigate the strength of articulatory highlights in the ASR of unbiased and whispered speech. We use acoustic, articulatory, and integrated acoustic and articulatory feature vectors in matched and mismatched train-test cases. The authors recommend that the articulatory information is valuable in ASR of both impartial and whispered discourse, particularly in the bungled train-test cases. When we concatenate acoustic and articulatory include vectors and convey them to the bungled train-test case where the demonstration is prepared with impartial discourse and tried with whispered discourse, a relative improvement in the phone error rate of 27.2% is observed compared to when only acoustic features are used. [10]

This paper clarifies the plan and execution of a smaller-scale watt-level control utilization human body capacitance-based sensor for recognizing and tallying exercise center workouts. The concept also works when the gadget is joined to a body portion that is not specifically included in the activity's movement. To distinguish, the majority of widely used movement-detecting methods necessitate placing the sensor on the moving body part (e.g., the leg in the case of leg-based gym exercises). We illustrated the basic principle underlying the ubiquitous electric connection between the human body and surroundings and examined the applicability of this detection technique in exercises at fitness centers. We evaluated our sensor on 11 participants who completed 7 common workouts from fitness centers every day for 5 days. Our sensor was positioned in three different body positions, including a non-contact position where the subject held the sensor in their pocket. The average counting accuracy that our sensing technique attained, according to the results, was 91%, which is quite competitive with commercial equipment.[11]

This paper explains a human pose estimation project using computer vision. The human posture estimation can be created utilizing manufactured insights or machine learning, in which the framework is encouraged with test information or prepared models and consequently can limit joints in a photograph or video of the human body. Now that the human body's joints are restricted, we can use them for a variety of purposes, like determining a person's gait cycle or tracking the advancements of an experienced competitor to determine the physical strategies and tactics required to secure their victory. Hence, one of the applications of human posture estimation may be creating a shrewd exercise center coach computer program that may offer assistance to bodybuilders in achieving their goals. [12]

This paper explains human pose detection and recognition using MediaPipe. With the advent of numerous new technologies and their widespread use in the fields of gaming and public security, among others, the importance of human action recognition has grown significantly in the modern day. We suggest a system that can distinguish between human activity in different environments and visual cues that enable the identification of distinct designs based on different spatiotemporal orientations. Using state-of-the-art technology, such as MediaPipe all-encompassing, which provides posture, confront, and hand point of interest discovery models, we present this paper. Our MediaPipe all-encompassing demonstration parses the outlines obtained through real-time apparatus nourishment using OpenCV, yielding an additional 501 points of interest that are sent out as arrangements to a CSV record, upon which we prepare a custom multi-class classification show to determine the relationship between the lesson and arrangements to classify and identify custom body dialect posture. This study uses the following machine learning classification algorithms: gradient boosting classifier, ridge classifier, random forest, and linear regression. [13]

This paper explains the MediaPipe framework. Building an application that forms perceptual inputs includes more than running an ML demonstration. Designers have to saddle the capabilities of a wide range of gadgets, adjust asset utilization and quality, run numerous operations in parallel and with pipelining, and ensure that time-series data is properly synchronized. The MediaPipe framework addresses these challenges. A developer can use MediaPipe to easily and rapidly combine existing and new perception components into prototypes and advance them to polished cross-platform applications. [14]

This paper explains that AI fitness has become a new and practical way of exercising, but most of the mainstream fitness apps focus on guiding and planning fitness activities, ignoring the detection and evaluation of users' fitness movements. Aiming at this wonder, this paper proposes a strategy to classify and tally fundamental wellness developments based on the Google Mediapipe system. The strategy comprises three steps: To begin with, a single wellness activity is isolated into two discovery states: up and down, and the comparing picture tests are collected and prepared. Besides, based on the created preparation set (csv record), KNN calculations were utilized to distinguish and classify diverse wellness activities. Finally, the classification results are processed, and the fitness actions are counted. [15]

System Architecture

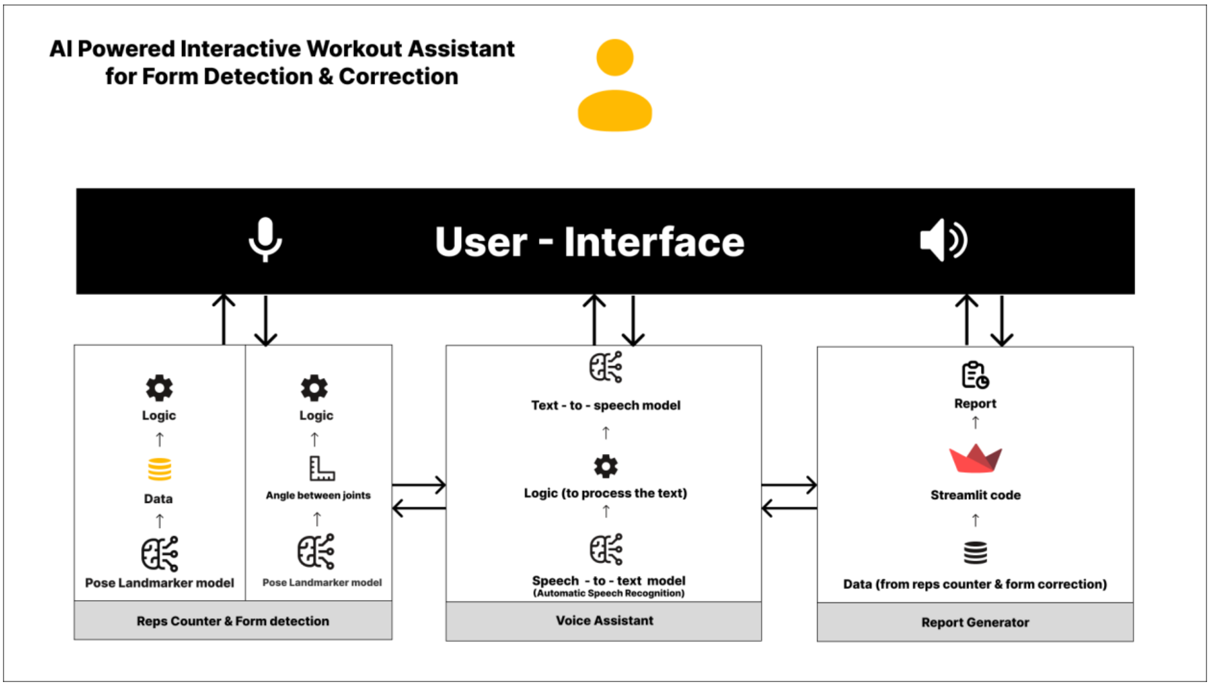

Fig 2: System Architecture

Source**:** Project Source Files (Original)

Description of the System Architecture

I. User Interface:

A user-friendly interface that provides steps for performing exercises properly and also provides a live video feed while performing exercise. We have implemented Streamlit, which is a Python framework for creating GUIs.

II. Reps Counter and Form Detection:

The reps counter and form detection feature in the gym assistant app utilizes a pose landmarker model, angle between joints, data, and logic to track and analyze the user's movements during exercises.

- Pose Landmarker Model: This is a pre-trained model in the mediapipe library that has a record of all joint angles and poses of the human body. This model proves helpful in extracting accurate joint angles, thus extracting valuable data points.

- Angle between joints: The application determines the user's body posture and movement by computing the angles between particular joints (such as the elbow, knee, and hip) once the pose landmarker model has identified the important spots on the user's body.
- Logic: The application interprets the angle measurements and determines if the user is completing the exercise correctly based on predetermined logic or guidelines.
- Data: In real-time, the posture landmarker model retrieves information from the user's body, including the coordinates of important joints and limbs. The user's body posture and movement during exercises are described in this data.

III. VOICE ASSISTANT:

The gym assistant app includes a voice assistant feature that allows users to interact with the app using voice commands. A user can start or stop his desired exercise by giving voice commands.

- Whisper-large (Automatic Speech Recognition): A pre-trained model for speech translation and automated speech recognition (ASR) is called Whisper. conditioned using 680k hours of annotated data. Alec Radford et al.'s Robust Speech Recognition via Large-Scale Weak Supervision paper from OpenAI is where Whisper first appeared.

IV. REPORT GENERATOR:

The system also features a report generator that automatically generates workout reports based on user activity. The report generator analyzes the user's workout data, such as repetitions and form quality, and generates a report that can help users track their progress, identify areas for improvement, and stay motivated in their fitness journey.

- Data from Reps Counter and Form Correction: The report generator collects data from the reps counter and form correction features, including information on the number of repetitions performed, exercise form quality, and any corrective measures suggested. This data is used to provide insights into the user's workout performance.
- Streamlit Code: The report generator uses Streamlit, a Python library for building web applications, to generate the user interface for the workout reports. Streamlit allows for the creation of interactive and customizable reports, making it easy for users to view and analyze their workout data.
- Report Generation: Based on the collected data and Streamlit code, the report generator creates a detailed workout report for the user. The report includes information on the user's workout performance, including metrics such as the number of repetitions and set exercise-specific details. The report is designed to be easy to understand and provides actionable insights for improving workout effectiveness.

Technical Description

I. DATASET:

We have manually collected videos of different exercises including Deadlift, Push-ups, Squats of our fellow member.

Fig 3: Push Ups

Fig 4: Deadlift

Source: Project Dataset (Original)

Then these videos were passed through a custom function, which further passes these videos through the pose landmarker model and extracts relevant data points needed to train another machine learning model to detect the form of the exercise. The output of the custom function is in CSV format.

II. Bicep Curl:

We have introduced an additional technique for bicep curls that makes use of the angle formed by the wrist, elbow, and shoulder, utilizing a pose landmarker model. To calculate angles, the following body points are employed, with reference to Fig. 1, Body Landmarks.

- Left shoulder
- Right shoulder
- Left elbow
- Right elbow
- Left wrist
- Right wrist

To find the angle between the shoulder, elbow, and wrist, use this formula:

- An array of the desired body part's x, y, and z coordinates is provided by the pose landmarker model.
- The angle has only been calculated using x and y variables.
- Stored in a numpy array, let the x and y coordinates of the wrist, elbow, and shoulder be a, b, and c, respectively.
- Then the angle between a, b, and c is determined by:

$$angle = |\, arctan2(c_1 - b_1, c_0 - b_0) - arctan2(a_1 - b_1, a_0 - b_0) \times \frac{180}{\pi} \,| \text{ ---- (1)}$$

Here,
arctan2 = $tan^{-1}(x)$
a = $[a_0,\ a_1]$
b = $[b_0,\ b_1]$
c = $[c_0,\ c_1]$

III. Algorithms:

We have used the following algorithms to train the exercise form detection model:

- Logistic Regression: For the purpose of categorizing body coordinates and determining the user's "up" and "down" positions, logistic regression has been employed. Given that the body coordinates dataset is linearly separable with limited features, logistic regression was shown to be a suitable fit for the data.
- Ridge Classifier: The Ridge classifier is utilized for classifying body coordinates and detecting the 'up' and 'down' positions of the user's body. It proves beneficial for high-dimensional data, which leads to a risk of multicollinearity and overfitting; hence, opting for a ridge classifier is a wise decision.

- Random Forest Classifier: In the context of data points being classified, we have high-dimensional data featuring different body coordinates, which may lead to overfitting issues due to unwanted noise and outliers that are not representative of the true underlying relationships in the data; hence, the random forest classifier is adept at handling such situations.
- Gradient Boosting Classifier: The data points under consideration are intricate. The body coordinate features showcase complex interactions and patterns; hence, gradient boosting tends to achieve high predictive accuracy and can proficiently handle our data; hence, gradient boosting classifiers play an important role.

Table 1: Accuracy of Algorithms

Algorithms	Accuracy	Precision	Recall
Logistic Regression	0.91	0.94	0.96
Ridge Classifier	0.89	0.88	0.87
Random Forest Classifier	0.98	0.96	0.97
Gradient Boosting Classifier	0.97	0.95	0.94

Source: Project Source Code (Original)

Conclusion

In conclusion, the AI Powered Interactive Workout Assistant for Form Detection & Correction, leveraging technologies like MediaPipe and OpenCV, represents a significant leap forward in revolutionizing fitness training. By seamlessly integrating computer vision and machine learning, this innovative system empowers users with real-time, accurate feedback on their exercise form. The precise detection of exercises such as bicep curls, push-ups, deadlifts, and squats not only enhance safety but also facilitates more effective workouts. Through the utilization of state-of-the-art algorithms, this gym assistant ensures that users receive personalized guidance, helping them optimize their fitness routines. The visual feedback provided by the system serves as a valuable tool for both beginners and seasoned fitness enthusiasts, promoting proper technique and minimizing the risk of injury. Moreover, the adaptability of this AI-powered assistant allows for a dynamic range of exercises to be incorporated into its repertoire, ensuring a comprehensive fitness experience. Its potential for continuous improvement and expansion of exercise recognition capabilities further cements its position as a cutting-edge solution in the fitness industry. Ultimately, this AI-powered gym assistant not only signifies a technological milestone but also embodies a commitment to enhancing the well-being and performance of individuals pursuing their fitness goals. As this technology continues to evolve, it holds the promise of transforming the way we approach fitness training, making it more accessible, precise, and personalized than ever before.

Acknowledgement

We extend our deepest gratitude to our team of four team members, alongside our esteemed project guide, for their helpful and unwavering dedication to the development of an AI Powered Interactive Workout Assistant for Form Detection & Correction. Each member has played an integral role in the project, contributing expertise in diverse areas such as model training for Bicep Curl, Squat, Deadlift, and Push Ups, form detection and correction for the exercises mentioned above, a speech recognition system, and a report generator. Together, we have leveraged various machine learning algorithms and mediapipe pose landmarker model to tackle the nuances of existing gym exercise systems. Furthermore, the collective effort in building our own datasets through videos and relevant technical stack thus integrating all the individual components of the application to portray a stupendous representation of the developed system, is commendable.

Conflict of Interest

The authors declare no conflict of interest concerning the research, authorship, and publication of this article.

Author Contribution Statement

SD[1] developed the form detection and progress tracking of bicep curls, squats, push-ups and deadlift exercises.

AB[2] fostered the form correction of bicep Curls, squats, push-ups and deadlift exercises as a successor stage to form detection.

TC[3] developed the voice assistant feature using the Whisper large language model for automatic speech recognition for the entire system.

SD[4] developed the exercise report generation feature using the Streamlit framework of the Python programming language.

Sub-author (Project Guide), TS[5] provided invaluable guidance throughout all stages of the project. They guided each stage of the project with critical feedback and suggestions, which fostered the development of a successful exercise form correction system.

References

[1]. Bradski, G., & Kaehler, A. (2008). *Learning OpenCV: Computer vision with the OpenCV library*. " O'Reilly Media, Inc.".

[2]. Haralick, R. M., Joo, H., Lee, C. N., Zhuang, X., Vaidya, V. G., & Kim, M. B. (1989). Pose estimation from corresponding point data. *IEEE Transactions on Systems, Man, and Cybernetics*, *19*(6), 1426-1446.

[3]. Yuan, Z., Hall, A. W., Zhou, S., Brunke, L., Greeff, M., Panerati, J., & Schoellig, A. P. (2022). Safe-control-gym: A unified benchmark suite for safe learning-based control and reinforcement learning in robotics. *IEEE Robotcs and Automation Letters*, *7*(4), 11142-11149.

[4]. Saravanan, R., & Sujatha, P. (2018, June). A state of art techniques on machine learning algorithms: a perspective of supervised learning approaches in data classification. In *2018 Second international conference on intelligent computing and control systems (ICICCS)* (pp. 945-949). IEEE.

[5]. Taware, G., Kharat, R., Dhende, P., Jondhalekar, P., & Agrawal, R. (2022, August). AI-based Workout Assistant and Fitness guide. In *2022 6th International Conference On Computing, Communication, Control And Automation (ICCUBEA* (pp. 1-4). IEEE.

[6]. Bhamidipati, V. S. P., Saxena, I., Saisanthiya, D., & Retnadhas, M. (2023, April). Robust intelligent posture estimation for an ai gym trainer using mediapipe and opencv. In *2023 International Conference on Networking and Communications (ICNWC)* (pp. 1-7). IEEE.

[7]. Ghaffarzadegan, S., Bořil, H., & Hansen, J. H. (2016). Generative modeling of pseudo-whisper for robust whispered speech recognition. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *24*(10), 1705-1720.

[8]. Grozdić, Đ. T., & Jovičić, S. T. (2017). Whispered speech recognition using deep denoising autoencoder and inverse filtering. *IEEE/ACM Transactions on Audio, Speech, and Language Processing*, *25*(12), 2313-2322.

[9]. Chang, H. J., Liu, A. H., Lee, H. Y., & Lee, L. S. (2021, January). End-to-end whispered speech recognition with frequency-weighted approaches and pseudo whisper pre-training. In *2021 IEEE Spoken Language Technology Workshop (SLT)* (pp. 186-193). IEEE.

[10]. Srinivasan, G., Illa, A., & Ghosh, P. K. (2019, May). A study on robustness of articulatory features for automatic speech recognition of neutral and whispered speech. In *ICASSP 2019-2019 IEEE International Conference on Acoustics, Speech and Signal Processing (ICASSP)* (pp. 5936-5940). IEEE.

[11]. Bian, S., Rey, V. F., Hevesi, P., & Lukowicz, P. (2019, March). Passive capacitive based approach for full body gym workout recognition and counting. In *2019 IEEE International Conference on Pervasive Computing and Communications (PerCom* (pp. 1-10). IEEE.

[12]. Dsouza, G., Maurya, D., & Patel, A. (2020, November). Smart gym trainer using Human pose estimation. In *2020 IEEE International conference for innovation in technology (INOCON)* (pp. 1-4). IEEE.

[13]. Singh, A. K., Kumbhare, V. A., & Arthi, K. (2021, June). Real-time human pose detection and recognition using mediapipe. In *International Conference on Soft Computing and Signal Processing* (pp. 145-154). Singapore: Springer Nature Singapore.

[14]. Lugaresi, C., Tang, J., Nash, H., McClanahan, C., Uboweja, E., Hays, M., ... & Grundmann, M. (2019, June). Mediapipe: A framework for perceiving and processing reality. In *Third workshop on computer vision for AR/VR at IEEE computer vision and pattern recognition (CVPR)* (Vol. 2019).

[15]. [15] Li, X., Zhang, M., Gu, J., & Zhang, Z. (2022, November). Fitness Action Counting Based on MediaPipe. In *2022 15th International Congress on Image and Signal Processing, BioMedical Engineering and Informatics (CISP-BMEI)* (pp. 1-7). IEEE.

[16]. Martinez-Martin, E., & Fernández-Caballero, A. (2025). Improved human emotion recognition from body and hand pose landmarks on the GEMEP dataset using machine learning. Expert Systems with Applications, 269, 126427.

[17]. Kwon, Y., & Kim, D. (2022). Real-Time workout posture correction using OpenCV and MediaPipe. *한국정보기술학회논문지*, *20*(1), 199-208.

[18]. Chaudhary, I., Singh, N. T., Chaudhary, M., & Yadav, K. (2023, May). Real-time yoga pose detection using opencv and mediapipe. In *2023 4th International conference for emerging technology (INCET)* (pp. 1-5). IEEE.

[19]. Gulati, S., Rastogi, A. K., Virmani, M., Jana, R., Pradhan, R., & Gupta, C. (2022, February). Paint/Writing Application through WebCam using MediaPipe and OpenCV. In *2022 2nd International Conference on Innovative Practices in Technology and Management (ICIPTM)* (Vol. 2, pp. 287-291). IEEE.

[20]. Patil, V., Sutar, S., Ghadage, S., & Palkar, S. (2023). Gesture Recognition for Media Interaction: A Streamlit Implementation with OpenCV and MediaPipe. *International Journal for Research in Applied Science & Engineering Technology (IJRASET)*.

[21]. Tomar, Y., Devi, S., & Kaur, H. (2023, December). Human Motion Tracker using OpenCv and Mediapipe. In *2023 3rd International Conference on Innovative Mechanisms for Industry Applications (ICIMIA)* (pp. 1199-1204). IEEE.
[22]. Kale, S., Kulkarni, N., Kumbhkarn, S., Khuspe, A., & Kharde, S. (2023). Posture Detection and Comparison of Different Physical Exercises Based on Deep Learning Using Media Pipe, Opencv. *International Journal of Scientific Research in Engineering and Management*, *7*(04), 1-29.
[23]. Rai, D., Kumar, A., & Baghel, A. (2024, June). Pose Detection Using OpenCV and Media Pipe. In *2024 International Conference on Integrated Circuits, Communication, and Computing Systems (ICIC3S)* (Vol. 1, pp. 1-6). IEEE.
[24]. Kumar, A., Kumar, A., Maggu, S., Kapoor, B., Kaushik, A. K., & Kalonia, S. (2024, May). Human Posture Detection and Correction Using MediaPipe and OpenCV. In *International Conference on Computing and Communication Systems for Industrial Applications* (pp. 135-147). Singapore: Springer Nature Singapore.
[25]. Gupta, A., Chawla, N., Jain, R., Thakur, N., & Devi, A. (2023). Gesture-Based Touchless Operations: Leveraging MediaPipe and OpenCV. *NEU Journal for Artificial Intelligence and Internet of Things*, *1*(2).
[26]. Ram, B. G., Murgod, T. R., Chandan, H., & Kumar, D. S. (2024). Real-time exercise assessment: MediaPipe & OpenCV approach. In *Computer Science Engineering* (pp. 676-685). CRC Press.
[27]. Li, L., Huang, H., Zeng, S., Cao, H., Zheng, R., & Lin, S. (2022). AI body detection and teaching system based on MediaPipe machine learning platform and OpenCV computer vision library. *Learning & Education*, *10*(8), 183-184.
[28]. Howse, J. (2013). *OpenCV computer vision with python* (Vol. 27). Birmingham, UK: Packt Publishing.
[29]. Hasan, R. T., & Sallow, A. B. (2021). Face detection and recognition using opencv. *Journal of Soft Computing and Data Mining*, *2*(2), 86-97.
[30]. Singh, G., Gupta, I., Singh, J., & Kaur, N. (2022, October). Face recognition using open source computer vision library (OpenCV) with Python. In *2022 10th International Conference on Reliability, Infocom Technologies and Optimization (Trends and Future Directions)(ICRITO)* (pp. 1-6). IEEE.
[31]. Mishra, S., Verma, V., Akhtar, N., Chaturvedi, S., & Perwej, Y. (2022). An intelligent motion detection using OpenCV. *International Journal of Scientific Research in Science, Engineering, and Technology*, *9*(2), 51-63.
[32]. Marengoni, M., & Stringhini, D. (2011, August). High level computer vision using opencv. In *2011 24th SIBGRAPI Conference on Graphics, Patterns, and Images Tutorials* (pp. 11-24). IEEE.
[33]. Rosenfeld, A. (1988). Computer vision: basic principles. *Proceedings of the IEEE*, *76*(8), 863-868.
[34]. SIMOES, W., REIS, L., ARAUJO, C., & MAIA JR, J. (2024). Accuracy Assessment of 2D Pose Estimation with MediaPipe for Physiotherapy Exercises. *Procedia Computer Science*, *251*, 446-453.
[35]. Supanich, W., Kulkarineetham, S., Sukphokha, P., & Wisarnsart, P. (2023, March). Machine learning-based exercise posture recognition system using mediapipe pose estimation framework. In *2023 9th International Conference on Advanced Computing and Communication Systems (ICACCS)* (Vol. 1, pp. 2003-2007). IEEE.[i]