FitSight: Tracking and Feedback Engine for Personalized Fitness Training

Hitesh Kotte*
German Research Center for Artificial Intelligence (DFKI)
Berlin, Germany
hitesh.kotte@dfki.de

Florian Daiber
German Research Center for Artificial Intelligence (DFKI)
Saarbrücken, Germany
email@dfki.de

Miloš Kravčík
German Research Center for Artificial Intelligence (DFKI)
Berlin, Germany
milos.kravcik@dfki.de

Nghia Duong-Trung
German Research Center for Artificial Intelligence (DFKI)
Berlin, Germany
nghia_trung.duong@dfki.de

ABSTRACT

Physical fitness presents a significant challenge in ensuring proper exercise posture. Individuals who work out need help maintaining correct exercise posture during their workouts. Maintaining correct form is critical for ensuring the safety and effectiveness of fitness routines. Yet, it is often challenging for individuals to keep proper form without professional guidance, which usually comes at expensive costs. The paper presents a novel method that utilizes the capabilities of YOLOv7 and a primary web camera to offer immediate feedback and correction on body posture during gym activities. Such a method empowers individuals to correct themselves and promotes motivation even without the presence of a professional trainer. This system has been developed to provide immediate, personalized feedback for various fitness exercises. It efficiently counts repetitions and provides textual guidance for improvement, tailored to the specific requirements of fitness enthusiasts. To determine the efficacy of our technology, we carried out a user study in a controlled laboratory setting simulating a gym environment. The study compares our interactive system with the traditional training method, involving participants of varied fitness levels. It showed significant improvements in exercise technique with real-time feedback. These findings are crucial for AI-supported training systems in strength training, underscoring the need for adaptive technologies for different user experiences. The research contributes to human-computer interaction and fitness technology discussions, highlighting interactive models' potential to augment and sometimes replicate personal training benefits in exercise form and posture improvement.

CCS CONCEPTS

• **Human-centered computing** → **User studies**; **User models**; **Heuristic evaluations**; *Visualization design and evaluation methods*; *Empirical studies in interaction design.*

UMAP '24, July 01–04, 2024, Cagliari, Italy

ACM ISBN 979-8-4007-0433-8/24/07
https://doi.org/10.1145/3627043.3659547

KEYWORDS

Pose Tracking, Feedback, YOLO, Fitness, Personalized Training.

ACM Reference Format:
Hitesh Kotte, Florian Daiber, Miloš Kravčík, and Nghia Duong-Trung. 2024. FitSight: Tracking and Feedback Engine for Personalized Fitness Training. In *Proceedings of the 32nd ACM Conference on User Modeling, Adaptation and Personalization (UMAP '24), July 01–04, 2024, Cagliari, Italy*. ACM, New York, NY, USA, 9 pages. https://doi.org/10.1145/3627043.3659547

1 INTRODUCTION

Augmented intelligence (AuI), a blend of human cognitive skills and artificial intelligence, is transforming various sectors, notably sports and fitness [16]. This transformation is driven by the amalgamation of machine learning, advanced sensor technology, and sophisticated camera systems, leading to the creation of state-of-the-art fitness training platforms. These platforms offer personalized coaching and real-time feedback, proving to be highly effective in improving performance in a wide range of physical activities [14, 17]. The role of fitness applications has become vital significantly when logistical or financial constraints limit access to personal training. Hiring a personal trainer for each workout session is often unaffordable, and coordinating schedules can be challenging. In contrast, fitness applications provide expert guidance without the logistical and financial challenges associated with personal trainers [15].

The ubiquity of smartphones and wearable technology further enhances the attractiveness of these applications, democratizing access to quality fitness advice. These technologies enable users to receive personalized, data-driven training advice, expanding the reach and effectiveness of fitness training [15]. However, the absence of expert supervision in these routines can lead to risks such as injuries and reduced motivation, especially for novice athletes or fitness enthusiasts [7]. Innovations in safety equipment and augmented reality tools have been suggested to mitigate these risks, but they often restrict movement during workouts [5]. Our research addresses these challenges by integrating computer vision technologies into existing fitness systems. By analyzing an individual's biomechanics through computer vision, we can identify incorrect forms or potentially hazardous movements [29].

Previous studies have validated AI's effectiveness in sports, particularly in recognizing activities in gym settings [18]. However, these methods usually require significant data collection and computational resources. Augmented reality solutions show promise but are limited by high costs and potential discomfort for users [26]. Multimodal Learning Analytics (MMLA) has demonstrated AI's ability to interpret complex learning data in sports [9]. However, issues related to these studies' reliable identification and replicability for generalizable outcomes remain unresolved, highlighting the need for more scalable and robust methodologies [21].

Our research addresses these issues by utilizing the YOLOv7 model for human pose estimation. This advanced model detects body key points, providing detailed insights into an individual's posture and biomechanics during exercises. We can discern correct from incorrect postures by analyzing these key points and offering real-time, personalized feedback.

2 RELATED WORK

AuI, an innovative technology that merges human and machine intelligence, has been gaining traction for enhancing performance and productivity. The use of AuI in tutoring systems [11] holds much potential to improve the effectiveness, safety, and enjoyment of sports and fitness activities. This concept is highly relevant in the specialized field of gym-based activity recognition, a significant part of human activity recognition. This area, characterized by a controlled yet dynamic environment where various exercises are performed, is ideal for deploying advanced object detection and machine learning algorithms. At the forefront of this field is the implementation of object detection models, particularly the YOLO (You Only Look Once) framework, which stands out for its improved accuracy and efficiency in recognizing a range of gym activities such as Pushups, Squats, Bench Press, and Shoulder Press [3, 6].

In tutoring systems for sport and fitness, augmented intelligence refers to a collaboration between humans and artificial intelligence (AI), where AI assists humans in decision-making and problem-solving [27]. By utilizing various sensors, cameras, and machine learning algorithms, tutoring systems can provide personalized coaching and feedback to individuals engaged in sports or fitness activities. The technology can track an individual's movements, analyze their performance data, and provide suggestions for improvement, which is especially important for activities that require the learning of complex motor movements like, climbing, running or weightlifting [10, 17, 18].

Pioneering research, for instance, has utilized neural networks to interpret electrocardiograms (ECGs) for distinguishing different aerobic activities [20]. Another notable study employed wristband and belt-based accelerometers alongside Artificial Neural Networks (ANNs) to accurately classify weightlifting activities [19]. These studies underscore the effectiveness of machine learning in decoding human actions and lay a solid foundation for applying the YOLO model in our Gym Activity Recognition (GAR) research. An integral component of our research intersects with studies on real-time feedback in exercise routines. For example, a ground-breaking study in weight training used clustering algorithms to provide immediate performance feedback [15]. This research highlights the critical role of real-time feedback in enhancing exercise outcomes. Additionally, the influence of Human-Computer Interaction (HCI) in shaping user experiences is paramount. Various studies have emphasized usability, learnability, and user satisfaction as critical metrics for evaluating activity recognition systems [2, 8]. These insights are instrumental in our upcoming user study, which will incorporate HCI principles to evaluate the effectiveness of our YOLO-based model.

Moreover, there has been a substantial amount of previous research in the domain of feedback methods in motor learning (see [23] for a thorough review). Recent research indicates that well-integrated visual cues significantly impact movement learning and execution [22], while auditory cues have been recognized as effective in enhancing athletic performance [24]. Techniques for motor learning have also been developed within the domain of Virtual Reality [4, 25] and Augmented Reality [1, 13, 28]. Augmented mirror approaches (e.g. [1, 12, 13] allow a user's movements to be supplemented with training content. FitSight also uses an augmented mirror approach that augments the trainee's movements with visual performance indicators in a mirror view to guide the trainee through the exercises.

Our research, through the upcoming user study of the FitSight system, seeks to extend the current literature by exploring the practicality and effectiveness of our YOLO-based gym activity recognition and feedback system. This study is designed to generate empirical data, moving beyond algorithmic accuracy to include HCI-related metrics such as user engagement and satisfaction. We aim to evaluate the system's impact on enhancing gym exercise performance, integrating machine learning methodologies, real-time feedback, and multimedia cues. This holistic approach is expected to provide valuable insights that will guide future research in computational techniques, user interface (UI) design, and integrated systems in the sports and fitness arena.

3 SELF-PRACTICING IN FITNESS TRAINING

The practice of fitness exercises independently presents several challenges. Maintaining correct posture and technique without a coach or trainer's guidance can be problematic. Incorrect posture increases the risk of injuries and diminishes the exercise's effectiveness. Moreover, self-motivation and progress tracking without external accountability and feedback are often challenging hurdles. Additionally, tailoring a workout program to individual needs and goals is crucial. Typically, a coach demonstrates an exercise, followed by the trainee attempting to replicate the observed movement. The trainer then monitors and provides feedback or re-demonstrates the exercise as needed. While the presence of a trainer during fitness sessions offers guidance, it can also lead to several issues. Trainees might become dependent on the trainer for motivation or to maintain a consistent workout regimen. Furthermore, employing a trainer can be costly, especially if they are required to be present at all times. This reliance may also limit the variety of exercises or routines a client tries, potentially leading to boredom and a decline in motivation. Clients must manage and adjust their workouts according to their needs and objectives.

We introduce a real-time feedback system utilizing AI and computer vision in response to these challenges. Here, trainers are not required to be continuously present; they can record exercise demonstrations as templates. The system then establishes a reference for the ideal posture and joint angles. Importantly, fitness has no universally correct posture; each individual's perfect posture varies due to unique body structures. Therefore, trainers must evaluate and provide customized posture guidance for each trainee. The proposed system allows trainers to easily modify the model's hyperparameters, which are the ideal joint angles, eliminating the need for repeated demonstrations. Moreover, the system's immediate feedback on posture and technique helps participants maintain proper form, reducing injury risks and aiding them in staying motivated and aligned with their fitness goals. Overall, this AI and computer vision-powered real-time feedback system enhances the efficacy and safety of independent fitness exercise practice.

4 IMPLEMENTATION

4.1 Calculating Joint Angles from Human Pose Key Points

Our approach utilizes a 17-keypoint pose topology to track and analyze body movements during fitness exercises [3]. Given three key points $u(x_u, y_u), v(x_v, y_v), p(x_p, y_p)$, the joint angle $\theta(u, v, p)$ (in degree) between two rays formed by three mentioned points is calculated as follows: $\theta(u, v, p) = \frac{180(\phi(y_p - y_v, x_p - x_v) - \phi(y_u - y_v, x_u - x_v))}{\pi}$. where the angle $\phi(y, x)$ (in radian) between the ray from the origin to the point (x, y) and the positive x-axis in the Cartesian plane is calculated as follows: $\phi(y, x) = arctan(\frac{y}{x})$, if $x > 0$; $\phi(y, x) = \frac{\pi}{2} - \arctan(\frac{x}{y})$, if $y > 0$; $\phi(y, x) = \arctan(\frac{y}{x}) \pm \pi$, if $x < 0$; $\phi(y, x) = -\frac{\pi}{2} - \arctan(\frac{x}{y})$, if $y < 0$. Table 1 showcases the angles between two rays formed by three joint points on different exercises. The correct angles were defined by fitness experts. We decompose our real-time fitness tutoring system into three phases: key points detection, Pose tracking, and output and feedback, see Algo 1.

Algorithm 1 Pipeline for Self-Learning of Fitness Exercises.

Keypoints Detection Phase:
Get input from recorded video or real-time streaming.
Detect $\{x_b, y_b, w_b, h_b, \text{cf}, k_x^0, k_y^0, k_{cf}^0, \ldots, k_x^{16}, k_y^{16}, k_{cf}^{16}\}$ in each frame.
Pose Tracking Phase:
Get predefined keypoints combination and range of joint angles.
Calculate θ and ϕ.
Convert θ to 0-100% range using one-dimensional linear interpolation.
Output and Feedback Phase:
Display the trainee's performance based on a predefined difficulty level.

4.2 FitSight System

4.2.1 Hardware Components: The system is centered around a Windows PC, powered by an Intel® Core™ i7-8700K CPU at 3.70 GHz and equipped with 32 GB of RAM, a high-performance NVIDIA GeForce RTX 2080 GPU with 8 GB of dedicated memory. Visual input is captured through an accessible external USB webcam to gain an unobstructed view of the participants' movements. The system utilizes a high-definition projector with a 1920x1080 resolution to deliver visual output.

Exercises	angleLH	angleRH	angleLL	angleRL
Bicep Curl	(30-178)°	(182, 327)°	(90, 180)°	(90, 180)°
Squats	(190, 234)°		(127, 172)°	
Lateral Shoulder Raise	(170, 192)°	(170, 192)°	(90, 180)°	(90, 180)°
Bent Over Row	(190, 238)°	(165, 106)°		
Lunges	(95, 165)°		(95, 165)°	
Shoulder Press	(50, 175)°	(305, 182)°	(90, 180)°	(90, 180)°

Table 1: Range of joint angles for different exercises. The angle data represents the measure of the angle created by the intersection of two rays, each line being defined by a pair of joint points. Angles are specified for the left hand (LH), right hand (RH), left leg (LL), and right leg (RL) where applicable. Measurements are given in degrees.

4.2.2 User Interface Overview: FitSight's UI is designed for intuitive interaction, providing users with clear visual cues and real-time feedback to facilitate their gym activities (see Figure 1). The UI contains the following components. **FPS Display**: This element shows the frames per second (FPS) delivered by the model running on YOLOv7, indicating the system's responsiveness. **Help Button**: Clicking this button offers users a guide on navigating and utilizing the system's features and functionalities. **Restart Button**: This button refreshes the current webpage, effectively restarting the application and resetting the session for a new exercise. **Download Analytics**: By pressing this button, users can download a report detailing the analytics of their performed exercises, providing insights into their workout session. **Feedback Text**: Displayed prominently at the top of the screen, this feature provides motivational messages to the user, such as "Great work! Keep going," enhancing the interactive workout experience. **Keypoints Button**: This toggle button, when activated, overlays key points on the user's body, aiding in the correct alignment and posture during exercises. **Recommendation Button**: This feature, when enabled, presents text-based feedback on the screen, offering suggestions and corrections for the user's form and technique. **Webcam Activation**: The 'Start Webcam' button initiates the real-time video feed, allowing the system to provide immediate feedback for the selected exercise.

4.2.3 Real Time Feedback: The real-time feedback system was designed to provide participants with immediate, actionable feedback to maintain correct posture during exercise routines. The system incorporated several key elements, each contributing to the overall effectiveness of the feedback provided. **Performance Bar**: The performance bar was a prominent feature of the system, visually representing the accuracy of the participant's posture in real time. It scaled from 0% to 100%, where 0% indicated a posture with substantial room for improvement, and 100% represented an ideal execution of the exercise form. This instantaneous feedback allowed participants to adjust their posture continuously throughout their workout. **Feedback Text**: Complementing the performance bar, feedback text was displayed on the top left corner of the screen. This running commentary provided targeted advice based on the participant's current posture, analogous to the corrections a personal trainer might offer during a training session. **Repetition Counter**: The feedback system also included a repetition counter which autonomously tracked and displayed the number of repetitions performed. This feature enabled participants to concentrate on their form rather than counting reps, promoting better overall exercise quality. **Goal Tracker**: To keep participants informed of their progress, the system featured a goal tracker indicating the remaining repetitions needed to complete the workout session. This goal-oriented metric is a motivational tool that encourages participants to persist and maintain proper form until all planned exercises are completed.

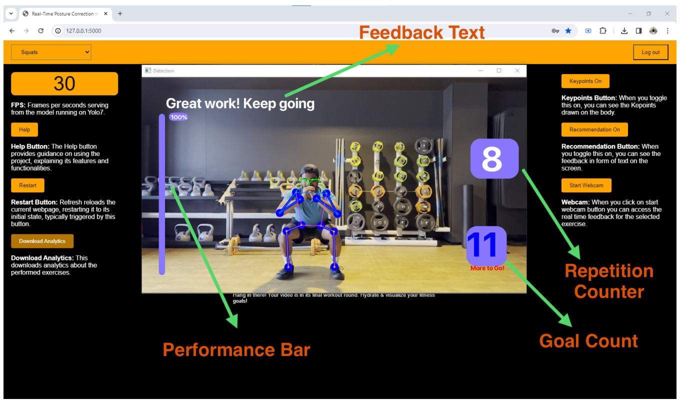

Figure 1: The real-time feedback system displays the performance bar, feedback text, repetition counter, and goal tracker.

As depicted in Figure 1, these features collectively formed an integral part of the feedback loop, providing a comprehensive performance overview in a single glance. In addition to these features, the system operates in two distinct modes to accommodate different user preferences and scenarios:

Webcam Mode: This mode provides live feedback to participants by analyzing their posture in real time as they perform exercises. It is particularly beneficial for immediate correction and ensuring that each movement is executed with proper form.

Recorded Processing Mode: In this mode, users can record their workout session and process the video later. It is ideal for those who prefer to review and reflect on their entire post-workout exercise routine, allowing them to assess their form and technique retrospectively and improve for future sessions.

These dual modes ensure that the system is versatile and adaptable, catering to the varied needs of users whether they seek immediate feedback or post-workout analysis.

5 USER STUDY

The purpose of this user study is to understand the effectiveness of the system, specifically comparing the performance of a real-time feedback system using the YOLO model and augmented feedback interface against the traditional method of human training. The focus of the study is on evaluating how well the real-time feedback system works compared to traditional training, rather than concentrating on the UI elements of the system.

5.1 Participants

A total number of 16 participants were recruited (twelve male, four female). The ages ranged from 19 to 28 years ($M = 23.63, SD = 3.04$), and their fitness levels were categorized between Beginner, Intermediate, and Expert levels. Only beginner and intermediate-level athletes were selected for the study because experts, by definition, have honed their exercise form over years of consistent practice and are likely to have developed an advanced understanding of correct posture and technique. Consequently, the posture correction model employed in this study was anticipated to have minimal impact on their already well-established routines. More details on demographics are summarized in Table 2.

5.2 Design

The performance of the proposed approach was compared to a traditional feedback method using a between-subjects design. Participants were assigned to either the Feedback Group or the No Feedback Group. Both groups were provided with the same exercises and amount of training. Both groups were introduced to the study's exercises through a standardized demonstration session conducted to replicate a typical gym environment with a human trainer. This session was designed to provide all participants with an equal foundation of understanding regarding the exercises, their form, and the expected posture.

In the introduction phase, the participants were encouraged to engage with the trainer actively, asking questions to clarify their understanding of the correct posture and to seek guidance on exercise execution. They were also permitted to perform a few trial repetitions, during which the trainer confirmed the accuracy of their form. This interactive approach ensured that all participants started the exercise sessions with a clear and precise understanding of the movements, fostering a controlled experimental setting. Participants performed six selected exercises for the study, namely: Bicep Curl, Squats, Shoulder Lateral Raise, Bent Over Row, Lunges, and Shoulder Press. The difference in each condition lies in the training method itself, how instructions are provided to the participant, and how feedback about the execution is given (i.e., Feedback Group: FitSight vs. No Feedback Group: human trainer).

Feedback Group: In the Feedback Group condition, eight participants performed exercises demonstrated by a fitness expert. During their workout, participants received real-time feedback displayed as text on a screen, along with a performance indicator that ranged from 0% to 100%, providing them with ongoing updates on their performance.

No Feedback Group: The No Feedabck Group, consisting of eight participants, was instructed to perform the same exercises as the previous group. However, compared to Feedback Group, these participants did not receive real-time feedback from FitSight. Instead, they carried out the exercises independently, with only a webcam turned on to simulate a mirror-like setup similar to that found in a typical gym environment. It allowed participants to self-monitor to some extent while exercising.

To ensure a robust experimental design, participants in the study were stratified into two distinct groups with careful consideration for demographic balance. This stratification was crucial to maintain the integrity of the study's outcomes, allowing for a fair comparison between the two sets of participants namely Feedback Group and the No Feedback Group.

ID	Level	Gender	Age
A1	Intermediate	Male	24
A2	Intermediate	Female	28
A3	Intermediate	Male	20
A4	Beginner	Male	21
A5	Intermediate	Male	28
A6	Intermediate	Male	27
A7	Beginner	Female	24
A8	Beginner	Male	25

Feedback Group

ID	Level	Gender	Age
B1	Beginner	Female	19
B2	Beginner	Male	19
B3	Intermediate	Male	20
B4	Intermediate	Male	22
B5	Intermediate	Male	24
B6	Intermediate	Male	27
B7	Beginner	Male	24
B8	Intermediate	Female	26

No Feedback Group

Table 2: Participant grouping strategy for the user study.

Gender and fitness level were balanced across both groups, aiming to achieve a fair comparison of the exercise performance of the real-time feedback effect vs. a traditional approach. Table 2 outlines the specific demographic distribution in each group.

5.3 Procedure

Participants were welcomed and fully introduced to the procedure study. Participants were informed that they would be recorded using a camera and that this data would be used only for research purposes. They were informed that they are allowed to take breaks and/or stop the study at any time. Their informed consent was obtained. The study was approved by the ethical review board of the faculty of Mathematics and Computer Science at Saarland University.

The study was explained in detail to the participants upon completing the consent form, and they were handed the "Participant Instructions Form." To gain insights into the participants' backgrounds, they were asked to complete the "Preliminary Questionnaire". Based on their background (i.e. gender and fitness levels) they were distributed into the feedback or no feedback group to maintain the balance of gender and fitness levels between groups.

Next, the appropriate weight of the dumbbells was selected for each participant to ensure the exercises were tailored to their strength and fitness level. This personalization was crucial for preventing injury and ensuring the exercises were challenging yet achievable. The study session commenced with a video recording to capture the participant's form and technique during the exercises. The trainer provided a comprehensive demonstration of all six exercises, during which participants were encouraged to ask any clarifying questions. This step was critical to ensure participants clearly understood the exercises they were about to perform.

The exercise session began with participants performing the exercises while the trainer evaluated their posture. After completing each exercise, the participant was given a 2-minute rest period. During this interlude, a verbal survey was conducted to capture the participant's immediate feedback on the workout they had just performed, and the researcher noted down their responses. This exercise, evaluation, and survey cycle was repeated for each of the six exercises in the session. Upon completing all exercises, participants were presented with a comprehensive user survey to collect their overall feedback on the session.

For those in the Control Group or Non-Feedback Group, the procedure concluded with a demonstration of the Feedback System, regardless of their previous engagement with it. This step was followed by a series of questions to gauge their perception of how such a system might have impacted their exercise session.

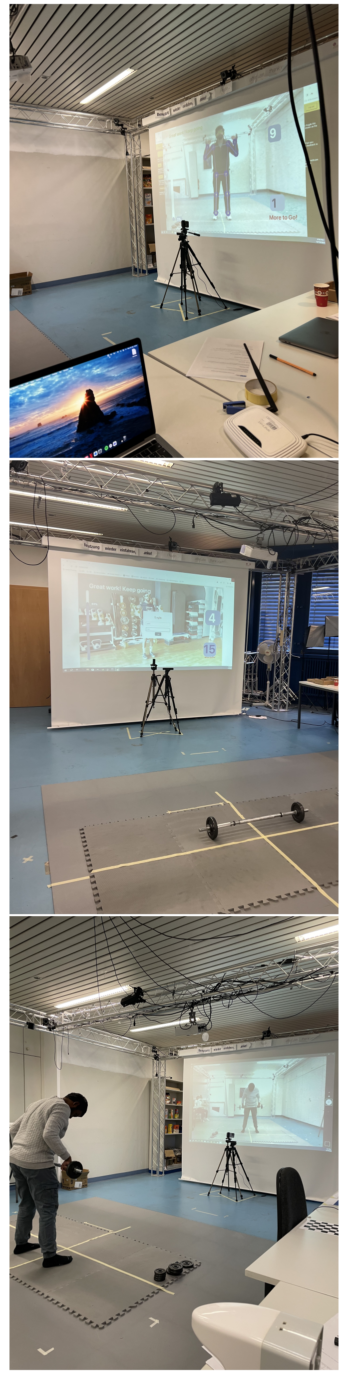

Figure 2: Our experimental site for the user study.

Elements such as ambient lighting and space arrangement were considered to enhance the realism of the gym setting. The goal was to create a conducive environment, allowing participants to perform as they would in their regular gym sessions, thus providing authentic data for the study's evaluation criteria.

6 RESULTS

The data collected from the User study was further examined using various metrics. The primary aim was to support our hypothesis that the group receiving feedback (Feedback Group) outperforms the group not receiving feedback (No Feedback Group). A simple mean of Trainer One score was obtained. The mean score for the Feedback Group was 42.1, while the average score for the No Feedback Group was 38.4. The preliminary findings were already affirmative. However, conducting more analysis was to comprehend the participant trends. Therefore, we give the outcomes of our statistical analyses focusing on implementing ANOVA and Tukey's HSD Post-Hoc Tests. We used these methods to analyze the differences in performance among various groups in our study, including beginners, intermediates, and male and female participants. The ANOVA tests played a crucial role in detecting any significant differences among groups, taking into account their levels of experience and gender. Subsequently, Tukey's HSD Post-Hoc Tests were performed to delve into the specifics of these differences. These studies are essential for comprehending the influence of varying training levels and gender on performance outcomes.

6.1 Inter-Rater Reliability

The user study involved a trainer who conducted a demonstration for all participants. It was essential to have this trainer present throughout the study to deliver the demonstration and determine the weights to be assigned to the participants. This trainer was also responsible for evaluating each participant's posture for each exercise.

It gives rise to another issue where the trainer may have exhibited bias due to being aware of the conditions in which the participants found themselves. Therefore, assessing if the trainer exhibited bias before relying on his scores for the remainder of the user study is imperative. In order to avoid bias in these ratings two more trainers were selected based on their suitability for the role. These trainers were graduate students in Sports Science specializing in weight training. All trainers rated the exercises independently and based on these ratings the inter-rater reliability was assessed.

The inter-rater reliability was measured using Spearman's rank correlation coefficient between the three trainers. Spearman's rho is a non-parametric measure that evaluates the rank correlation between two variables by examining how effectively their relationship can be expressed using a monotonic function. It is employed when the data possess ordinal characteristics but may not necessarily exhibit linearity. The values go from -1 to 1, with 1 indicating a flawless positive correlation and -1 representing a flawless negative correlation. Table 3 provides the correlation matrix, and the overall agreement score is based on the average of the pairwise correlations.

The overall agreement score among the three trainers, based on pairwise Spearman's correlations, is $r = 0.6746$. The pairwise correlations between the trainers were as follows: - Trainer 1 and Trainer 2: $r = 0.6181$ - Trainer 1 and Trainer 3: $r = 0.7234$ - Trainer 2 and Trainer 3: $r = 0.6823$

All correlations suggest a moderate to strong positive agreement between the trainers.

The assessment is essential in establishing the consistency and reliability of results from the study, guaranteeing that the evaluations are unbiased and unaffected by personal biases.

	Trainer1	**Trainer2**	**Trainer3**
Trainer1	1.0000	0.6181	0.7234
Trainer2	0.6181	1.0000	0.6823
Trainer3	0.7234	0.6823	1.0000
Overall Agreement Score		**0.6746**	

Table 3: Pairwise Spearman's Rank Correlations Between Trainers with Overall Agreement Score

6.2 Analysis of ANOVA Results

ANOVA is a statistical technique used to assess the degree of variation or difference between two or more groups in an experiment. The study utilized ANOVA to determine whether there are statistically significant differences among the means of the groups. The subsequent subsections provide a comprehensive account of the ANOVA analysis carried out for different participant categorizations, encompassing beginners and individuals with intermediate skill levels grouped by gender.

Beginners: A one-way ANOVA was conducted to compare the effect of treatment on beginners. The analysis showed no significant difference between groups, $F(1,4) = 6.12, p = .069$, although a trend towards significance was observed.

Intermediates: A one-way ANOVA revealed a significant effect of treatment among intermediates, $F(1,8) = 36.48, p < .001$, indicating a strong treatment effect.

Male Participants: For male participants, a one-way ANOVA showed no significant treatment effect, $F(1,10) = 0.85, p = .379$.

Female Participants: A one-way ANOVA for female participants did not reveal a significant effect of treatment, $F(1,2) = 10.02, p = .087$, although a trend towards significance was observed.

Table 4 comprehensively summarizes the ANOVA results. The 'Variation Source' column distinguishes between the variance attributable to the experimental groups (Between Groups) and the variance within each group (Within Groups, also known as Residual). The Sum of Squares (Sum Sq)' column quantifies the variability, and the' Degrees of Freedom (df)' column relates to the number of independent levels within the data. The' F-Statistic (F)' column represents the ratio of the variance between the groups to the variance within the groups, and the P-value (PR(>F))' column provides the probability of observing the calculated F-statistic, or one more extreme, under the 28 assumption that the null hypothesis is true. A p-value below the threshold of 0.05 typically indicates statistical significance.

The ANOVA results provide insights into the treatment's influence on various groups. Although the beginners and female participants exhibited indications of potential significance, implying that

Group	Variation Source	Sum Sq	df	F	P-value
Beginners	Between Groups	75.0	1	6.12	.069
	Within Groups	49.0	4	-	-
Intermediates	Between Groups	102.70	1	36.48	<.001
	Within Groups	22.52	8	-	-
Male	Between Groups	14.12	1	0.85	.379
	Within Groups	166.30	10	-	-
Female	Between Groups	93.52	1	10.02	.087
	Within Groups	18.67	2	-	-

Table 4: Detailed ANOVA results for different participant groups.

the treatment might have an impact that could be more evident with more significant sample numbers or more investigation, the intermediates and male participants had unequivocal outcomes. Further inquiry is required to examine the effectiveness of the treatment in the intermediate group, as it has shown a significant outcome. However, the lack of significance in the male participants indicates that the treatment's effect needs to be distinguishable from random variation.

It is crucial to acknowledge that whereas ANOVA detects disparities in averages, it does not indicate the specific locations of those disparities. Therefore, when the analysis of variance (ANOVA) yields statistically significant results, it is common practice to do post-hoc tests, such as Tukey's honestly significant difference (HSD) test, to identify the precise differences between groups. These findings add to the current body of knowledge and provide insights for future experimental designs and potential practical applications that consider gender and skill level disparities.

6.3 Tukey's HSD Post-Hoc Tests

Tukey's Honest Significant Difference (HSD) Post-Hoc test is a reliable approach in inferential statistics for comparing the means of various groups. It is specifically designed to control the family-wise error rate. This section explores the comparative study of two different group settings: the comparison of the entire group and the comparison of an intermediate group.

Group1	Group2	MeanDiff	P-adj	Lower	Upper	Reject
A	B	-0.3125	0.8983	-5.4642	4.8392	False

Table 5: Tukey's HSD Post-Hoc Tests: Overall Group Comparison

Group1	Group2	MeanDiff	P-adj	Lower	Upper	Reject
A	B	-6.5417	0.0003	-9.0391	-4.0442	True

Table 6: Tukey's HSD Post-Hoc Tests: Intermediate Group Comparison

6.3.1 Overall Group Comparison. A Tukey's HSD Post-Hoc test was conducted to compare the mean difference between the Feedback Group and the No Feedback Group. The results revealed a mean difference of -0.3125, $p = 0.8983$, 95% CI $[-5.4642, 4.8392]$, suggesting no significant difference between the groups. Nevertheless, the p-value related to this disparity in means is 0.8983, significantly beyond the standard alpha limit of 0.05. The high p-value indicates a strong likelihood that the observed difference in means may have been caused by random sampling variation, assuming that the actual difference between the groups is zero. Hence, the null hypothesis, which states that there is no substantial disparity between the means of the groups, cannot be disproven.

Furthermore, the confidence interval for the mean difference ranges from -5.4642 to 4.8392, including zero inside its boundaries. The 95% confidence interval suggests that the actual difference in means might range from -5.4642 to 4.8392. Given that the interval encompasses the value of zero, it provides more evidence to support the conclusion that the observed difference in means is statistically insignificant. Hence, the comparison needs to yield adequate data to establish a notable impact of the conditions or interventions represented by both the groups.

6.3.2 Intermediate Group Comparison. The Tukey's HSD Post-Hoc test was also used to compare the intermediate groups. The results showed a significant mean difference of -6.5417, $p = 0.0003$, 95% CI $[-9.0391, -4.0442]$, confirming that the Feedback Group's mean is significantly lower than the No Feedback Group's mean. The p-value is significantly smaller than the alpha threshold of 0.05, suggesting a highly improbable occurrence of such a substantial difference in means if the null hypothesis were valid. Thus, the null hypothesis is refuted, confirming a significant disparity between the two groups. The confidence interval for the mean difference is asymmetric, with lower and upper bounds of -9.0391 and -4.0442, respectively, and does not encompass zero. This interval establishes a 95% confidence level for the actual mean difference, confirming the observed substantial difference. The negative sign of the mean difference indicates that the mean of Feedback Group is significantly lower than No Feedback Group's, with statistical significance.

7 DISCUSSION

Investigating real-time feedback mechanisms in exercise performance offers a promising opportunity for advancement in gym activities. Utilizing advanced object detection algorithms, such as the YOLO model, for gym activity recognition allows an additional chance to evaluate the impact of real-time feedback on workout posture and effectiveness. This study has focused on these components to offer insights into the possible advantages of incorporating technological augmentation into training routines. The results have more significant implications for developing AI-supported training systems and their incorporation into regular exercise routines.

7.1 Beginners vs. Intermediate Athletes

The ANOVA findings for the intermediate group are highly significant, indicating a strong likelihood that the observed differences are due to the factors being studied rather than by random chance. These findings demonstrate substantial differences in performance between groups, with Feedback Group exhibiting better outcomes. The conclusive outcome indicates that the real-time feedback given to Feedback Group has positively influenced their exercise performance, therefore supporting the hypothesis that athletes at an intermediate level can significantly gain from real-time performance-related information.

However, no significant differences in posture correction between conditions could be shown in the beginner group. This result can be explained by considering the relative lack of knowledge of the novice participants with gym workouts. At the beginning of their physical training experience, they may want extra instructional assistance beyond real-time feedback. Beginners may benefit from more extensive assistance, possibly through advanced visual aids or augmented reality avatars, to increase their posture and workout technique.

In the Tukey's HSD post-hoc analysis, the comparison between the entire group and the intermediate groups for Feedback Group and No Feedback Group highlights a significant performance difference in the intermediate group. This suggests that the intermediate group benefited more from the intervention or conditions being tested, as compared to the overall group. The findings underscore the effectiveness of the feedback provided specifically to the intermediate group.

7.2 Gym vs. Lab

The study was conducted in a very close setting to the gym environment, i.e., the study tried to mimic the gym environment but was conducted in the lab environment. The RPE was maintained for every participant, ensuring there was no bias in the difficulty of performing the exercises, which might directly influence our results. Also, the trainer's choice of proving the weights for participants varied between participants, making it not biased as each participant's BMI is different, and proving higher weights for higher BMI is explainable and not biased. The trainer was successful in doing so with sport science calculations, and this can be verified with the RPE results; it was painted with light activity, ensuring the participants were not experiencing fatigue between exercises, and hence, additionally, they were given 1-2 minutes break between exercises to reduce fatigue. Overall, this study indicates a connection between the amount of training experience and the effectiveness of real-time feedback in exercise environments. It raises the question of how individuals can be effectively assisted at all fitness levels using adaptive, intelligent technologies that improve physical performance and adherence to proper technique.

7.3 Limitations

The study encounters several limitations. While based on their presumed proficiency, the decision to exclude the expert group potentially overlooks insights into the upper echelons of exercise mastery. This choice must be revised to understand how highly skilled individuals might benefit from or perceive the feedback system. Moreover, the reliance on technology for analyzing gym activities, though innovative, may need to capture the full complexity of physical movements and individual biomechanical differences. The controlled environment of the study and its duration may also not fully represent the varied and dynamic conditions of typical gym settings or capture the long-term effects and adaptability of the feedback system in real-world scenarios. These limitations underscore the need for further research, particularly in developing more comprehensive and objective evaluation methods in gym activity recognition.

8 CONCLUSION

This research represents a notable advancement in fitness training by introducing the FitSight system, an innovative AI system incorporating the YOLOv7 model to estimate body postures accurately. The results of our comprehensive evaluation, which included 16 participants, highlight the system's effectiveness in improving exercise performance. The real-time feedback functionality of FitSight has been crucial in ensuring proper posture, reducing the risk of accidents, and maximizing the effectiveness of workout outcomes.

Yet, when limiting the intermediate group for analysis, the findings shift, uncovering a significant difference and confirming that the real-time feedback system notably improves performance for those with intermediate exercise experience rather than extensive expertise. These insights highlight the significance of adapting fitness technology to suit various user group's different needs and abilities. Real-time feedback technologies are a valuable complement to the training program of intermediate athletes. The results of our user study indicate that FitSight's real-time feedback greatly assists in self-monitoring workout practices. It is crucial in situations where professional guidance is absent. The collected data also yielded profound insights into the diverse needs of fitness enthusiasts, emphasizing the system's wide-ranging effectiveness and ability to personalize AI-powered fitness equipment.

As we look ahead, our primary goal is to enhance the precision of FitSight. We aim to customize the system to fit a broader range of customer requirements, improving its versatility. Conducting tests in actual fitness centers will provide valuable insights into the practical usefulness and user engagement in realistic circumstances. Furthermore, conducting studies to assess the system's long-term influence on fitness motivation and outcomes would be crucial for comprehending and improving user engagement and continued utilization. Ultimately, the FitSight technology signifies a significant breakthrough at the convergence of artificial intelligence and fitness training. It provides a scalable and efficient alternative to existing training methods and creates future research and development opportunities in AI-assisted individualized fitness training.

ACKNOWLEDGMENTS

The authors would like to thank the German Federal Ministry of Education and Research (BMBF) for their kind support within the project *Multimodal Immersive Learning with Artificial Intelligence for Psychomotor Skills* (MILKI-PSY) under the project ID 16DHB4014 and within the project Life-Long Federated User and Mental Modeling for Health and Well-being (FedWell) under the project ID 01IW23004.

REFERENCES

[1] F. Anderson, T. Grossman, J. Matejka, and G. Fitzmaurice. 2013. YouMove: Enhancing Movement Training with an Augmented Reality Mirror. *Proceedings of the 26th annual ACM symposium on User interface software and technology - UIST '13* (2013), 311–320.

[2] Fabio Buttussi, Luca Chittaro, and Daniele Nadalutti. 2006. Bringing Mobile Guides and Fitness Activities Together: A Solution Based on an Embodied Virtual Trainer *(MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 29–36.

[3] Nghia Duong-Trung, Hitesh Kotte, and Miloš Kravčík. 2023. Augmented Intelligence in Tutoring Systems: A Case Study in Real-Time Pose Tracking to Enhance the Self-learning of Fitness Exercises. In *Responsive and Sustainable Educational Futures*, Olga Viberg, Ioana Jivet, Pedro J. Muñoz-Merino, Maria Perifanou, and Tina Papathoma (Eds.). Springer Nature Switzerland, Cham, 705–710.

[4] Daniel L. Eaves, Gavin Breslin, and Paul van Schaik. 2011. The Short-Term Effects of Real-Time Virtual Reality Feedback on Motor Learning in Dance. *Presence: Teleoperators and Virtual Environments* 20, 1 (2011), 62–77.

[5] Çağatay Berke Erdaş and Selda Güney. 2021. Human activity recognition by using different deep learning approaches for wearable sensors. *Neural Processing Letters* 53 (2021), 1795–1809.

[6] Evianita Dewi Fajrianti, Sritrusta Sukaridhoto, M Rasyid, Bambang Edi Suwito, Rizqi Putri Nourma Budiarti, Ilham Achmad Al Hafidz, Naufal Adi Satrio, and Amma Liesvarastranta Haz. 2022. Application of augmented intelligence technology with human body tracking for human anatomy education. *IJIET: International Journal of Information and Education Technology* 12, 6 (2022), 476–484.

[7] Alireza Farrokhi, Reza Farahbakhsh, Javad Rezazadeh, and Roberto Minerva. 2021. Application of Internet of Things and artificial intelligence for smart fitness: A survey. *Computer Networks* 189 (2021), 107859.

[8] Andrew Garbett, Ziedune Degutyte, James Hodge, and Arlene Astell. 2021. Towards Understanding People's Experiences of AI Computer Vision Fitness Instructor Apps. In *Designing Interactive Systems Conference 2021* (Virtual Event, USA) *(DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1619–1637.

[9] Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad. 2022. *The Multimodal Learning Analytics Handbook*. Springer Nature.

[10] Mahmoud Hassan, Florian Daiber, Frederik Wiehr, Felix Kosmalla, and Antonio Krüger. 2017. FootStriker: An EMS-based Foot Strike Assistant for Running. *Proc. ACM Interact. Mob. Wearable Ubiquitous Technol.* 1, 1, Article 2 (mar 2017), 18 pages. https://doi.org/10.1145/3053332

[11] Justin Kim, Taylor Davis, and Lingzi Hong. 2022. Augmented Intelligence: Enhancing Human Decision Making. In *Bridging Human Intelligence and Artificial Intelligence*. Springer, 151–170.

[12] Felix Kosmalla, Florian Daiber, Frederik Wiehr, and Antonio Krüger. 2017. ClimbVis: Investigating In-situ Visualizations for Understanding Climbing Movements by Demonstration. In *Proceedings of the 2017 ACM International Conference on Interactive Surfaces and Spaces* (Brighton, United Kingdom) *(ISS '17)*. Association for Computing Machinery, New York, NY, USA, 270–279. https://doi.org/10.1145/3132272.3134119

[13] Felix Kosmalla, Christian Murlowski, Florian Daiber, and Antonio Krüger. 2018. Slackliner - An Interactive Slackline Training Assistant. In *Proceedings of the 26th ACM International Conference on Multimedia* (Seoul, Republic of Korea) *(MM '18)*. Association for Computing Machinery, New York, NY, USA, 154–162.

[14] Hitesh Kotte, Milos Kravcik, and Nghia Duong-Trung. 2023. Real-Time Posture Correction in Gym Exercises: A Computer Vision-Based Approach for Performance Analysis, Error Classification and Feedback. In *Proceedings of the Third International Workshop on Multimodal Immersive Learning Systems (MILeS 2023). International Workshop on Multimodal Immersive Learning Systems (MILeS-2023), located at Eighteenth European Conference on Technology Enhanced Learning (EC-TEL 2023), September 4-8, Aveiro, Portugal*, Khaleel Asyraaf Mat Sanusi, Bibeg Limbu, Jan Schneider, Milos Kravcik, and Roland Klemke (Eds.), Vol. 3499. CEUR Workshop Proceedings, 64–70. CEUR-WS.org ISSN 1613-0073.

[15] Yousef Kowsar, Masud Moshtaghi, Eduardo Velloso, Lars Kulik, and Christopher Leckie. 2016. Detecting Unseen Anomalies in Weight Training Exercises. In *Proceedings of the 28th Australian Conference on Computer-Human Interaction* (Launceston, Tasmania, Australia) *(OzCHI '16)*. Association for Computing Machinery, New York, NY, USA, 517–526.

[16] Debapriya Maji, Soyeb Nagori, Manu Mathew, and Deepak Poddar. 2022. YOLO-Pose: Enhancing YOLO for Multi Person Pose Estimation Using Object Keypoint Similarity Loss. In *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition*. 2637–2646.

[17] Katya N Mileva and Lisa Zaidell. 2022. Sport and Exercise Science and Health. In *Health studies: An introduction*. Springer, 85–124.

[18] Hristo Novatchkov and Arnold Baca. 2013. Artificial intelligence in sports on the example of weight training. *Journal of sports science & medicine* 12, 1 (2013), 27.

[19] Xiyang Peng, Xiang Wang, Jun Qi, Yun Yang, Jie Li, and Po Yang. 2020. Experimental Analysis of Artificial Neural Networks Performance for Accessing Physical Activity Recognition in Daily Life. In *2020 IEEE Intl Conf on Parallel & Distributed Processing with Applications, Big Data & Cloud Computing, Sustainable Computing & Communications, Social Computing & Networking (ISPA/BDCloud/SocialCom/SustainCom)*. 1348–1353.

[20] Jun Qi, Po Yang, Martin Hanneghan, Stephen Tang, and Bo Zhou. 2018. A Hybrid Hierarchical Framework for Gym Physical Activity Recognition and Measurement Using Wearable Sensors. *IEEE Internet of Things Journal* PP (05 2018).

[21] José A. Ruipérez-Valiente, Roberto Martínez-Maldonado, Daniele Di Mitri, and Jan Schneider. 2022. From Sensor Data to Educational Insights. *Sensors* 22, 21 (2022). https://doi.org/10.3390/s22218556

[22] Alessandra Semeraro and Laia Turmo Vidal. 2022. Visualizing Instructions for Physical Training: Exploring Visual Cues to Support Movement Learning from Instructional Videos. In *Proceedings of the 2022 CHI Conference on Human Factors in Computing Systems* (New Orleans, LA, USA) *(CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 90, 16 pages.

[23] Roland Sigrist, Georg Rauter, Robert Riener, and Peter Wolf. 2013. Augmented visual, auditory, haptic, and multimodal feedback in motor learning: A review. *Psychonomic Bulletin & Review* 20, 1 (2013), 21–53.

[24] Fabrizio Sors, Mauro Murgia, Ilaria Santoro, and Tiziano Agostini. 2015. Audio-Based Interventions in Sport. *The Open Psychology Journal* 8 (12 2015), 212–219.

[25] E. Todorov, R. Shadmehr, and E. Bizzi. 1997. Augmented feedback presented in a virtual environment accelerates learning of a difficult motor task. *Journal of Motor Behavior* 29, 2 (1997), 147–158.

[26] Mamoru Ueta. 2022. Improving Mirror Fitness through augmented reality technology. In *2022 3rd International Conference on Big Data, Artificial Intelligence and Internet of Things Engineering (ICBAIE)*. IEEE, 186–194.

[27] Parvathy Venkatachalam and Sanjog Ray. 2022. How do context-aware artificial intelligence algorithms used in fitness recommender systems? A literature review and research agenda. *International Journal of Information Management Data Insights* 2, 2 (2022), 100139.

[28] Yihong Wu, Lingyun Yu, Jie Xu, Dazhen Deng, Jiachen Wang, Xiao Xie, Hui Zhang, and Yingcai Wu. 2023. AR-Enhanced Workouts: Exploring Visual Cues for At-Home Workout Videos in AR Environment. In *Proceedings of the 36th Annual ACM Symposium on User Interface Software and Technology (UIST '23)*. Association for Computing Machinery, New York, NY, USA, Article 121, 15 pages. https://doi.org/10.1145/3586183.3606796

[29] Mingqiu Xie. 2020. Design of a physical education training system based on an intelligent vision. *Computer Applications in Engineering Education* 29 (2020), 590 – 602. https://api.semanticscholar.org/CorpusID:225787194