SAARLAND UNIVERSITY

Faculty of Mathematics and Computer Science
Department of Computer Science
MASTER THESIS

Augmented Intelligence in Tutoring Systems: Real-Time Pose Tracking to Enhance the Self-learning of Fitness Exercises

Hitesh Kotte
Master's Program of Data Science and Artificial Intelligence
May 2024

Advisor:

Dr. Florian Daiber
German Research Center for Artificial Intelligence
Saarland Informatics Campus
Saarbrücken, Germany

Reviewers:

Prof. Dr. Antonio Krüger
German Research Center for Artificial Intelligence
Saarland Informatics Campus
Saarbrücken, Germany

Dr. Nghia Duong-Trung
Educational Technology Lab
German Research Center for Artificial Intelligence
Berlin, Germany

Submitted:
17th of May, 2024

Saarland University
Faculty MI – Mathematics and Computer Science
Department of Computer Science
Campus - Building E1.1
66123 Saarbrücken
Germany

Declarations

Eidesstattliche Erklärung

Ich erkläre hiermit an Eides Statt, dass ich die vorliegende Arbeit selbständig ver- fasst und keine anderen als die angegebenen Quellen und Hilfsmittel verwendet habe.

Statement in Lieu of an Oath

I hereby confirm that I have written this thesis on my own and that I have not used any other media or materials than the ones referred to in this thesis.

Saarbrücken, 17th of May, 2024

Einverständniserklärung

Ich bin damit einverstanden, dass meine (bestandene) Arbeit in beiden Versionen in die Bibliothek der Informatik aufgenommen und damit veröffentlicht wird.

Declaration of Consent

I agree to make both versions of my thesis (with a passing grade) accessible to the public by having them added to the library of the Computer Science Department.

Saarbrücken, 17th of May, 2024

Acknowledgements

I would like to express my sincere gratitude to all the people whose help and advice have been crucial in finishing my thesis.

First and foremost, I would want to express my sincere gratitude to Dr. Florian Daiber, my advisor, for his sound advice, continuous backing, and expert guidance on this subject.

Additionally, I would like to express my gratitude to Prof. Dr. Antonio Krüger, under whose guidance this thesis was conducted. His insights and opinions have greatly impacted my studies and given me an enjoyable educational journey.

I would specially like to thank Dr. Nghia Trung Duong and Dr. Milos Kravcik, who have both been an essential part of my journey. Throughout this journey, their positions as advisors and mentors have given me tremendous support and direction. Their knowledge and perspectives have really enhanced both my research experience and personal development.

I would want to express my gratitude to each and every study participant, as their participation was essential to the project's success. Their time and work have greatly influenced the study's conclusions.

I would especially want to thank M.Sc. Vinayak Reddy, whose help with the study was invaluable. In addition to helping with the practical components of this research, his function as a trainer has improved my educational experience.

Lastly, but definitely not least, I would want to sincerely thank my family and friends. Throughout this journey, their continuous support, understanding, and drive have been my pillars of strength. Their confidence in me has always been a source of inspiration and has been essential to my success.

Abstract

The pursuit of physical fitness often encounters the obstacle of maintaining proper exercise posture during gym workouts, which is crucial for the safety and effectiveness of fitness routines. This thesis introduces an advanced solution utilizing the YOLOv7 object detection model to offer real-time feedback on posture, aiming to enhance the quality of gym workouts. The primary objective of this investigation was to develop and assess a technology-supported system capable of improving exercise form in the absence of professional trainers. Employing computer vision and the YOLOv7 model, the system identifies human keypoints and tracks them with algorithms tailored for the human body structure. This approach provides immediate feedback on exercise technique, enabling self-correction and sustained motivation for individuals working out independently. The system also incorporates transfer learning to optimize its learning architecture, reducing the need for extensive model retraining. The study's results, derived from a user study involving 16 participants across different levels of fitness experience, revealed that the group receiving real-time feedback via our system exhibited marked improvements in exercise technique over the control group using traditional training methods. The intermediate group, in particular, showed significant advancements, highlighting the system's effectiveness. In conclusion, the investigation confirms the potential of AI-supported systems in enhancing traditional fitness training methods. The findings suggest that real-time feedback mechanisms can substantially benefit exercise form, especially for those without access to personal trainers. The research contributes to the fields of human-computer interaction and fitness technology, underlining the transformative impact of interactive models in replicating the advantages of personal training.

Contents

Chapter 1
Introduction

1.1 Motivation

Augmented intelligence, blending human cognitive capabilities with the power of artificial intelligence, is making significant inroads in various sectors, with the fields of sports and fitness being particularly transformative [19]. This synergy of human and machine intelligence is fueled by a harmonious integration of machine learning, state-of-the-art sensor technology, and sophisticated camera systems. The culmination of these technologies has led to the emergence of advanced fitness training platforms. These platforms have revolutionized traditional training methods by offering customized coaching and real-time feedback, which have been instrumental in enhancing performance in a multitude of physical activities [20].

In today's fitness landscape, the role of digital fitness applications has become increasingly prominent, especially when access to traditional personal training resources is hampered by logistical or economic constraints. The conventional model of hiring a personal trainer for each workout session is often fraught with financial and logistical challenges, making it an impractical solution for many. Conversely, fitness applications have emerged as a viable alternative, circumventing these constraints by providing expert-level guidance remotely. These applications leverage the expertise of personal trainers and make it accessible without the usual financial and scheduling burdens associated with traditional training methods [22].

The proliferation of smartphones and wearable technology has further amplified the reach and efficacy of these fitness applications. This technological ubiquity has democratized access to quality fitness training, enabling users from various demographics to receive tailored, data-informed training recommendations. The convenience and personalized nature of these applications have significantly expanded the scope and effectiveness of fitness training [22]. However, the shift towards app-based fitness regimes is not without its challenges. The absence of in-person expert supervision can lead to increased risks such as potential injuries and a decline in motivation, especially for novice and intermediate athletes or fitness enthusiasts [11].

To counteract these risks, recent innovations in safety equipment and augmented reality tools have been proposed. However, these solutions often come with their own set of limitations, such as restrictions on the freedom of movement, which is a critical component of effective workouts [9]. Our research aims to address these challenges by integrating advanced computer vision technologies into fitness training systems. This integration enables a detailed analysis of an individual's biomechanics during exercise, facilitating the identification and correction of improper forms or movements that could lead to injury [38].

maybe add research questions here

1.2 Method of the Investigation

In addressing the challenges of gym activity recognition and posture correction, this research adopts a computer vision approach, utilizing the YOLOv7 object detection model. This model is at the core of the FitSight application, a pioneering tool designed to enhance the safety and effectiveness of gym workouts by providing real-time feedback on exercise posture.

1.2.1 Introduction to YOLOv7

YOLOv7 represents the latest advancement in the YOLO (You Only Look Once) series, known for its exceptional speed and accuracy in object detection tasks. This model outperforms its predecessors by incorporating improved algorithms for detecting objects with higher precision, even in complex visual environments. Its architecture is optimized to process images rapidly, making it ideally suited for real-time applications. In the context of FitSight, YOLOv7's capabilities allow for the accurate detection and analysis of an athlete's posture during various gym exercises, such as pushups, shoulder presses, and bicep curls.

1.2.2 Real-Time Feedback System

FitSight leverages YOLOv7 to monitor the user's movements through a simple web camera setup. This setup does not require specialized hardware, making the application accessible to a wider audience. The camera captures the user's exercise in real-time, and the YOLOv7 model analyzes these movements to identify any deviations from the correct posture.

The application provides instant textual feedback to the user, suggesting adjustments to align with the optimal exercise form. This feature is crucial for preventing injuries that can occur when exercises are performed incorrectly, effectively substituting for the guidance of a personal trainer.

1.2.3 Webcam Setting and Augmented Mirrors Approach

The webcam setup in FitSight is inspired by the concept of augmented mirrors, which enhances the user's experience by allowing them to see themselves and their posture corrections in real-time. This approach not only aids in maintaining correct posture throughout the workout but also enhances the user's awareness of their own body movements. The visual feedback, coupled with textual suggestions from the application, empowers users to adjust their posture immediately, further minimizing the risk of injury.

This augmented mirrors approach, combined with the sophisticated analysis provided by YOLOv7, offers a comprehensive solution for at-home or gym-based workouts. It democratizes access to personalized fitness coaching, extending the benefits of expert guidance to users regardless of their location or access to personal trainers.

1.2.4 Rationale and Significance

The choice of YOLOv7 and the innovative webcam setup for FitSight are driven by the goal to create an accessible, effective, and user-friendly platform for exercise guidance. By integrating advanced computer vision technology with a simple camera interface, this research bridges the gap between high-tech fitness solutions and everyday users. The application's ability to provide real-time, accurate feedback on exercise form represents a significant step forward in preventing exercise-related injuries and enhancing the overall effectiveness of workouts.

In summary, the methodology behind this investigation showcases the potential of combining cutting-edge object detection models with user-friendly technology to create impactful fitness solutions. The FitSight application exemplifies how technology can be leveraged to improve health outcomes and exercise safety, paving the way for future innovations in digital fitness coaching.

1.3 Transition to Automated Feedback in Fitness Training

In the traditional fitness training model, as depicted in Figure 1.1, the trainee performs an exercise under the guidance of a trainer. The trainer observes the execution and provides feedback, which the trainee uses to adjust and improve their performance in subsequent attempts. This cycle of performance and feedback is fundamental to the learning and improvement of exercise techniques.

Transitioning from this conventional model, our research introduces a technology-supported fitness training paradigm, encapsulated in Figure 1.2. In this advanced model, the trainee performs an exercise while being recorded by a camera. The recorded video is then processed through computer vision algorithms, specifically leveraging the YOLOv7 object identification model. This model analyzes the exercise execution in terms of posture and joint angles, comparing it to ground-truth postures and angles that represent the correct execution of the exercise.

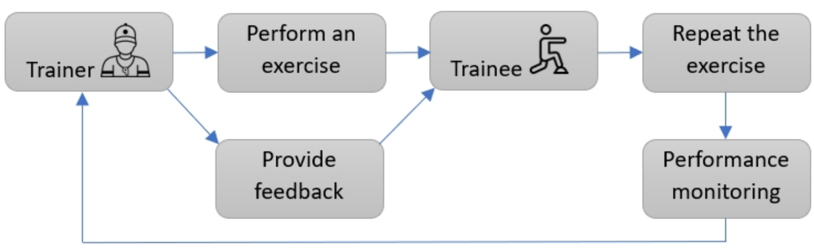

Figure 1.1: Feedback during the traditional fitness training.

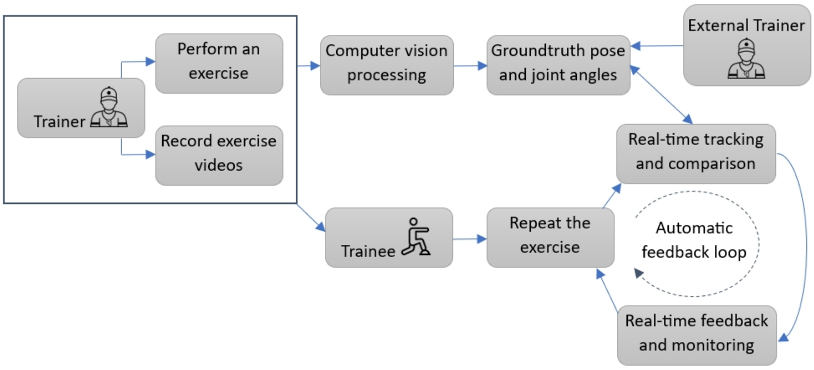

Figure 1.2: Feedback during the technology-supported fitness training.

The innovation lies in the automatic feedback loop. Once the system detects any deviation from the ideal exercise form, it immediately provides real-time feedback to the trainee. This feedback loop not only replicates but also enhances the traditional feedback mechanism by offering consistent, objective, and immediate corrections without the need for an external trainer's presence at all times. The technology-supported system monitors performance continuously, allowing the trainee to correct their posture in real-time, thereby fostering a safer and more effective workout environment.

This integration of an intelligent feedback system within the fitness training process marks a significant leap from the conventional training methods. By automating the feedback loop, the FitSight application empowers individuals to train effectively, with reduced risk of injury, and ensures that the benefits of expert guidance are not lost even when training in the absence of a personal trainer.

1.4 Principal Results of the Investigation

The core investigation of this study was conducted through a user study that divided participants into two distinct groups. One group was provided with real-time feedback on their exercise form using the FitSight application, while the control group was allowed to self-monitor their form using a traditional mirror setup, without any real-time feedback.

The study encompassed a diverse demographic to ensure a comprehensive analysis, with participants stratified across three fitness levels: beginner, intermediate, and expert. This stratification was crucial for assessing the efficacy of real-time feedback across varying degrees of exercise proficiency.

The results of the study were illuminating, revealing a clear distinction in performance improvements between the two groups. Notably, participants in the intermediate group who received real-time feedback from the FitSight application demonstrated a significantly higher improvement in exercise form and technique compared to their counterparts who relied on mirror feedback alone.

These results highlight the role of real-time, technology-enhanced feedback in improving gym workouts. Notably, intermediate-level exercisers—those beyond beginner skills but not yet at peak proficiency—benefited most from the instant adjustments offered by the FitSight system.

This significant differential suggests that the presence of real-time feedback can play a pivotal role in the progression of exercise capabilities, especially for those in the critical phase of transitioning from beginner to advanced levels. The implications of these results point to a potential paradigm shift in how feedback is utilized in fitness training, with technology-supported systems like FitSight offering a substantial advantage over traditional methods.

1.5 Outline

Existing literature has already established the capabilities of AI in sports, with a particular emphasis on activity recognition in gym environments [23]. However, these AI-driven methods often require extensive data collection and significant computational resources. While augmented reality solutions offer innovative approaches, they are not without drawbacks, including high costs and potential discomfort for the users [35].

The field of Multimodal Learning Analytics (MMLA) has further explored AI's ability to interpret complex data patterns in sports and fitness [13]. Despite the advancements, there remain challenges in ensuring the reliability and generalizability of these studies, indicating a need for more scalable and robust methodologies [26].

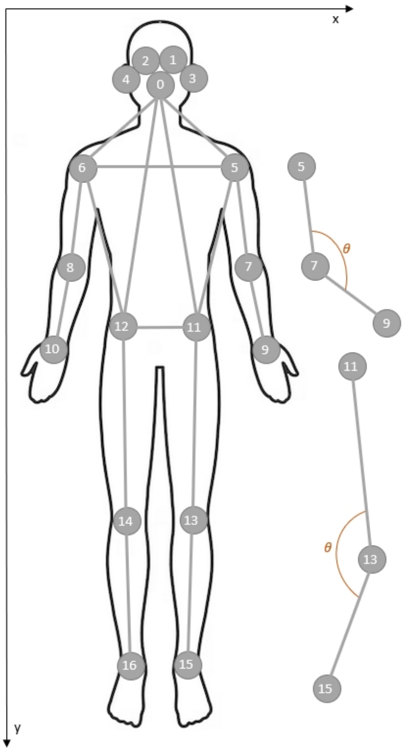

Figure 1.3: Human Keypoints provided by Yolov7

Our research contributes to this evolving landscape by employing the YOLOv7 model for human pose estimation, a model renowned for its accuracy and efficiency. This state-of-the-art model is engineered to detect body keypoints, as depicted in Figure 1.3, offering an in-depth understanding of an individual's posture and biomechanics during exercises. Utilizing these keypoints, we then calculate the angles between them in different movements to differentiate between correct and incorrect postures, providing real-time, personalized feedback.

The structure of the thesis is outlined in the following. The *Related work* chapter discusses sections of *Augmented Intelligence in Tutoring systems* and *Integration of Machine Learning in Human Activity Recognition*. The *Concept and Implementation* chapter gives an overview of development and integration of FitSight. The study structure, methods, results, and discussion can be found in chapters *User Study and Results Analysis*. Lastly, chapter *Conclusion* comprises the findings of the thesis and gives approaches and recommendations for future work.

Chapter 2
Related Work

Recent years have witnessed a surge in research focusing on activity recognition within gym environments, a subset of the broader field of human activity recognition. The gym setting offers a controlled yet dynamic ecosystem where multiple forms of exercises are performed, making it a fertile ground for investigating sophisticated object detection and machine learning algorithms. The task involves the automated identification of different gym activities—such as Pushups, Squats, Bench Press, Shoulder press through various computational models. Of particular interest in this realm is the application of object detection models like YOLO (You Only Look Once) to enhance the accuracy and efficiency of activity recognition [6]. This literature review aims to outline significant contributions to object detection models for activity recognition, particularly those applicable to or tested in gym settings, while excluding tangential topics like gamification, which are less relevant to the current study.

2.1 Augmented Intelligence

Augmented intelligence (AuI), an innovative technology that synergistically combines human cognitive function with artificial machine intelligence, is rapidly emerging as a transformative force in various sectors, including sports and fitness. The application of Augmented Intelligence, as explored in recent studies [15], offers substantial promise for enhancing the efficiency, safety, and enjoyment of physical activities. The implications of AuI are particularly profound within the domain of gym-based activity recognition, which forms a critical subset of the broader human activity recognition framework. Gym environments, characterized by their controlled yet dynamic nature, present a unique setting where diverse physical activities are executed, making them an ideal landscape for the application of sophisticated object detection and machine learning algorithms.

Leading the charge in this specialized field is the adoption of advanced object detection models, most notably the YOLO (You Only Look Once) framework. Celebrated for its high-speed processing and precision, YOLO has been pivotal in advancing the accuracy and effectiveness of activity recognition systems. Specifically, in the context of gym activities, YOLO has demonstrated exceptional proficiency in identifying and classifying various exercises, ranging from Pushups and Squats to Bench Press and Shoulder Press, thereby underscoring the model's versatility and reliability [6, 10].

The integration of AuI into gym activity recognition signifies a leap towards a more interactive and intelligent fitness experience. The real-time data processing and feedback provided by YOLO model implementations not only cater to the immediate corrective needs of individuals performing the exercises but also contribute to long-term health and performance tracking. This innovative approach aligns with the current trajectory of personalized fitness regimens and data-driven training programs. As such, the intersection of Augmented Intelligence and object detection models holds substantial potential for revolutionizing the way gym activities are monitored, analyzed, and enhanced, paving the way for a new era of intelligent fitness solutions.

2.1.1 Augmented Intelligence in Tutoring Systems

When designing intelligent tutoring and coaching systems, a key challenge is achieving a synergy of human intelligence and AI. It uses machine learning and deep learning to analyze data and recognize patterns, allowing humans to take over. This hybrid knowledge should improve the performance of humans not only in cognitive but also in psychomotor areas. A systematic review of intelligent tutoring systems based on gross body movement detected through computer vision [33] indicated that the state-of-the-art computer vision methods for detection and tracking along with the temporal aspects are not much explored, that there is a need for large-scale implementations and testing for generalizability as well as that mechanism used in providing feedback needs to be evaluated and optimized [33].

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 [36]. 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 high-impact activities like running or weightlifting. [20, 23].

Moreover, there has been a substantial amount of previous research in the domain of feedback methods in motor learning (see [30] for a thorough review). Recent research indicates that well-integrated visual cues significantly impact movement learning and execution [28], while auditory cues have been recognized as effective in enhancing athletic performance [32]. Techniques for motor learning have also been developed within the domain of Virtual Reality [34, 8] and Augmented Reality [1, 16].

2.1.2 Augmented Mirror Approaches

Augmented mirror approaches (e.g. [1, 16] allow movements of a user to be augmented with training content. FitSight also uses an augmented mirror approach, that augments the movements of the trainee with visual performance indicators in a mirror view, to guide the trainee through the exercises.

An impressive development in physical training is the incorporation of augmented reality mirrors, as demonstrated by the "YouMove" system. This novel method improves conventional movement training by employing a large-scale augmented reality mirror that offers immediate visual feedback. The technology facilitates the generation and dissemination of instructional material, enabling the capture of bodily motions that users can subsequently acquire. The training module of the system utilizes recorded data to lead users through interactive phases, gradually diminishing their dependence on supervision in order to foster independent learning.

The utilization of augmented reality (AR) mirrors in training systems such as "YouMove"[1] highlights the capacity of technology to enhance physical learning environments. AR mirrors offer instantaneous and accurate feedback, enabling learners to independently assess and modify their moves in real-time. The prompt visual input is essential for rectifying posture and technique, thereby expediting the learning process. Furthermore, research has demonstrated that these systems can greatly enhance learning results and long-term memory compared to conventional video demonstrations. The "YouMove" system, which utilizes Kinect technology for recording and interactive training, signifies a significant advancement in utilizing augmented reality (AR) for physical training. This has wide-ranging implications for sectors such as yoga, dancing, and physical workouts.

2.2 Integration of Machine Learning in Human Activity Recognition

The burgeoning field of human activity recognition has seen an exponential increase in studies leveraging the capabilities of machine learning algorithms, particularly within the health and fitness industry. This surge in research reflects a growing consensus about the pivotal role machine learning can play in monitoring and enhancing wellness and healthcare services. An array of machine learning architectures has been thoroughly investigated and implemented, including, but certainly not limited to, sophisticated neural networks, the robust Support Vector Machines (SVM), and the probabilistic Hidden Markov Models (HMM). These systems have been adeptly employed for the recognition, categorization, and analysis of a wide spectrum of physical activities.

For instance, one seminal study delved into the nuances of electrocardiograms (ECGs) by deploying neural network-based analytical frameworks to differentiate between a variety of intensive aerobic exercises [25]. The research provided compelling evidence showcasing the efficacy of machine learning algorithms in dissecting and understanding complex patterns within biometric data, specifically for the purpose of activity recognition. The success of this study underscored the potential of these computational techniques in potentially transforming how physical activities are monitored and assessed for health and performance outcomes.

In another innovative study, researchers embraced the use of wearable technology, employing wristband and belt-mounted accelerometers in conjunction with Artificial Neural Networks (ANNs). This approach aimed to meticulously classify different types of weightlifting activities. The results were impressive, indicating a high degree of accuracy in the classification processes [24]. Such endeavors not only reinforce the practical utility of machine learning strategies in capturing and interpreting human movement but also serve as a solid foundation of empirical evidence that enriches the academic dialogue surrounding physical activity recognition.

These pioneering research efforts collectively contribute to a robust intellectual framework, lending credence to the practicality and efficacy of machine learning methodologies in decoding the complex nature of human kinetics. Furthermore, they establish an academic bedrock upon which newer models and techniques can be tested and validated. This is particularly relevant to our current research on Gym Activity Recognition (GAR), where we are implementing the cutting-edge YOLO (You Only Look Once) model [19]. This model is renowned for its rapid and efficient object detection capabilities, which we are harnessing to identify and analyze gym activities with the goal of providing real-time, actionable feedback to users. By drawing on the strengths of these prior studies, our research aims to push the boundaries of what is possible in the realm of activity recognition, opening up new avenues for innovation and application in the ever-evolving intersection of machine learning and physical training.

2.2.1 Real-time Pose Detection and Tracking

YOLO (You Only Look Once) is a real-time object detection system that has gained popularity recently due to its high speed and accuracy. After more than a decade of continuous development, the YOLO model family has significant implications for computer vision research, with many applications poised to leverage this technology [29, 5, 37]. YOLO is a single-stage object detection model that uses a deep neural network to detect and classify objects in an image. It achieves high accuracy using a unified detection approach that considers the object's class and location in a single forward pass. YOLO also has a fast inference speed, making it suitable for real-time applications. While initially designed for object detection, the YOLO architecture has been adapted for human pose estimation, a critical task in computer vision for fitness applications. The YOLO pose estimation model is built initially on the YOLOv3 architecture and then YOLOv5, which uses a deep convolutional neural network (CNN) to detect and estimate human poses in real-time. Notably, YOLO-pose estimation has yielded impressive results and is the state-of-the-art for real-time human pose detection [19]. It does not use heat maps. Instead, it associates all key points of a person with anchors. YOLO models, in general, are easy to implement and train, making them accessible to researchers and developers with limited resources. YOLO's high accuracy, real-time performance, and simplicity make it an ideal model for pose detection and tracking applications. Consequently, we integrate the latest version of YOLO-pose [19] into our proposed fitness tutoring system as the backbone for real-time 17 key points detection. Moreover, thanks to the transfer learning technique [39, 7], we do not need to re-train the model on our experimented datasets because the YOLO has been trained on the Common Objects in Context (COCO) dataset over 200000 images with 250000 person instances.

2.3 Real-Time Feedback and Enhanced User Experience

The research landscape within physical exercise regimes is densely populated with studies that underscore the transformative potential of real-time feedback mechanisms. A particularly noteworthy investigation within this domain focused on the realm of weight training, implementing advanced clustering algorithms to provide users with immediate, data-driven feedback on their exercise performance [17]. The significance of this study is twofold: not only does it demonstrate the practical applications of machine learning in crafting personalized feedback loops, but it also cements the transition of real-time feedback from a supplementary feature to a critical element that tangibly enhances the quality and effectiveness of exercise routines.

In parallel, the field of Human-Computer Interaction (HCI) has been instrumental in redefining user experiences within digital environments. A suite of studies has emerged, advocating for the integration of key HCI principles—namely usability, learnability, and user satisfaction—into the evaluation process of activity recognition systems [3] [12]. These contributions are invaluable, as they delineate a multifaceted approach to user-centric design and evaluation, proposing a set of criteria that holistically measure system performance through the lens of user interaction.

Our forthcoming user study aims to intersect with these pivotal research trajectories. By adopting and integrating established HCI metrics, we intend to meticulously evaluate our YOLO-based gym activity recognition model. This evaluation will not only consider the technical accuracy of the model but will also critically assess how users interact with, perceive, and benefit from the system. In doing so, we will provide a nuanced understanding of the model's operational efficacy and its alignment with user experience paradigms. This dual-focused approach, rooted in both technical precision and user-centric design, aspires to yield a system that is not only high-performing but also intuitive and satisfying for the end-user, thereby embodying the essence of what HCI principles advocate for in the context of interactive systems.

2.4 Role of Multimedia and Gamification in User Engagement

The corpus of human activity recognition literature is continually being augmented with studies that emphasize the significant role of multimedia and gamification in engaging users in physical activities. The integration of multimedia features, particularly visual cues, has been closely scrutinized and evidenced to enhance the learning curve and execution of complex movements in exercise routines [28]. The precise and effective use of these visual stimuli has been recognized as a catalyst for improved cognitive and physical coordination, leading to more efficient and safer workouts.

In addition to visual aids, there has been an exploration into the realm of auditory feedback, with several research initiatives advocating for its use as an influential mechanism for boosting athletic performance. The rhythmic and timely delivery of auditory signals has been posited as a method to not only enhance focus and timing during exercises but also to potentially expedite the acquisition of new motor skills [32].

Simultaneously, the expanding domain of gamification in physical activities has garnered substantial scholarly attention. The persuasive power of game design elements, when implemented in non-gaming contexts like fitness applications, has been linked with increased motivation and adherence to exercise programs [14] [4]. The strategic incorporation of these elements can invoke a sense of achievement and progression, which are cornerstones of engaging gameplay. This has been evidenced to foster a more compelling and sticky user experience, leading to sustained user engagement.

As we consider the evolution of our YOLO-based Gym Activity Recognition (GAR) system, these studies offer insightful perspectives that can guide the integration of multimedia and gamification features into future iterations of our application. By weaving these elements into the fabric of the GAR system, we anticipate not only an elevation in the usability and enjoyment factors but also a significant enhancement in the overall effectiveness of the exercise routines. The potential to blend these innovative features into our system promises a richer, more interactive, and enjoyable user journey, potentially leading to higher levels of user satisfaction and superior exercise outcomes.

2.5 Upcoming User Study and Forward-Looking Insights

Our research endeavors to contribute to the existing corpus of literature by implementing an upcoming user study focused on the applicability and performance of our YOLO-based Gym Activity Recognition (GAR) system. This study is designed to yield empirical data, offering a comprehensive evaluation that extends beyond mere algorithmic accuracy to encompass metrics related to Human-Computer Interaction (HCI), specifically user engagement and satisfaction.

By integrating machine learning methodologies, real-time feedback mechanisms, and multimedia cues, the study aims for a multi-faceted assessment of the system's utility in enhancing gym exercise performance. This inclusive approach anticipates providing substantial insights that can inform the trajectory of future research efforts, specifically in the computational techniques, user interface design, and integrated systems within the GAR domain.

The systematic review of related works serves to contextualize our research within the broader scientific landscape, thereby solidifying its academic relevance. The forthcoming user study will add empirical evidence to this discourse, thereby enriching the academic conversations in the rapidly evolving fields of GAR and HCI.

Expanding upon these foundational aspirations, we also plan to explore additional applications of our architecture across varied contexts. Notably, the utilization of the YOLOv7 model to analyze and enhance golf postures presents an exciting avenue. Such adaptation could offer golfers precise, real-time feedback on their stances, potentially revolutionizing training methods in the sport. Similarly, the dance community could benefit from applying the YOLOv7 model to assess and improve dance postures. Integrating a virtual trainer system could provide dancers with immediate corrections and guidance, fostering an environment of continuous skill enhancement. These explorations into diverse use cases underscore our commitment to broadening the impact of our research, paving the way for novel integrations of technology in sports and performance arts.

Chapter 3
Concept and Implementation

This chapter describes the development and application of FitSight System that utilizes real-time feedback and object recognition to improve training experiences. FitSight aims to offer users actionable insights and real-time feedback to enhance their posture in gym activities through the application. The technology is specifically meant to encourage individual interaction, eliminating the requirement for traditional training method such as a trainer to be physically present. FitSight engages in real-time interaction with user movements, providing a variety of real-time feedback to help users improve their workout routines.

3.1 Self-practicing and Feedback in Fitness

Self-practicing of fitness exercises can be challenging for several reasons. First, it cannot be easy to maintain proper posture and technique without the guidance of a coach or trainer. Poor posture can lead to injuries and reduced effectiveness of the exercise. Second, staying motivated and tracking progress without external accountability and feedback can be challenging. Third, it can be challenging to personalize a workout program to meet individual needs and goals. Figure 1.1 presents a traditional scenario of practicing fitness. In the process, a coach or an instructor performs a particular exercise before a trainee. Then, the participant tries to repeat the movement by remembering what has been observed. Next, by monitoring the trainee's performance, the trainer can provide feedback or perform the exercise again. While having a trainer present during fitness practice can be helpful in terms of guidance, several potential issues can arise. Trainees may become overly reliant on the trainer and need their presence to motivate themselves or maintain a consistent fitness routine. Additionally, hiring a trainer can be expensive, and if the instructor is always present, the cost of their services can add up quickly. If the trainer is always leading the workout, the client may not have the opportunity to try different exercises or workout routines, which can lead to boredom and decreased motivation. It is also essential for clients to perform their workouts and make adjustments based on their individual needs and goals. Scheduling conflicts may also arise if the trainer is always present.

Therefore, a real-time feedback system powered by AI and computer vision can help address these challenges. We present the proposed technology-supported fitness training in Figure 1.2. Here, the trainers do not need to present the entirety, or even they can record fitness exercises once as instruction templates. Then, the proposed system calculates the groundtruth pose and joint angles for later reference. Indeed, no correct posture applies to every trainee in the context of fitness and exercise. Everyone has a unique body structure, meaning their ideal posture and alignment may differ. Thus, trainers need to assess each trainee individually and provide personalized guidance and adjustments to their posture. The trainers or external instructors can easily adjust the configuration because the ideal values of joint angles are the model's hyperparameters. Hence, the trainers do not need to re-perform the exercise again. Moreover, by providing immediate feedback on posture and technique, the system can help participants maintain proper form and reduce the risk of injury. Feedback, in our implementation, can be simply to count how many correct exercise execution has been performed and how likely the posture is compared to the groundtruth pose. The system can also track progress and provide personalized visualization for workout programs based on individual needs and goals. It can help users stay motivated and on track with their fitness goals. Overall, a real-time feedback system powered by AI and computer vision can help improve the effectiveness and safety of self-practicing of fitness exercises.

3.2 Proposed Fitness Tutoring System

We decompose our real-time fitness tutoring system into three phases: Keypoints Detection, Pose Tracking, and Output and Feedback. The concise pseudo-code in Algorithm 1 breaks down subroutines. In the Keypoints Detection Phase, human pose estimation boils down to a single class person detection problem, with each person having 17 associated key points. Each keypoint is again identified with a bounding-box location and the confidence score: $\{x_b, y_b, w_b, h_b, \text{cf}\}$. In practice, the model can detect multiple people in each frame. Each key point is represented by the anchor location and its confidence score.

Algorithm 1 Pipeline for Self-Learning of Fitness Exercises.

Keypoints Detection Phase:
Get input from recorded video or real-time streaming.
Detect $\{\text{batch\_id}, \text{class\_id}, x_b, y_b, w_b, h_b, \text{cf}, k_x^0, k_y^0, k_{cf}^0, k_x^1, k_y^1, k_{cf}^1, \ldots, k_x^{16}, k_y^{16}, k_{cf}^{16}\}$ in each frame.
Pose Tracking Phase:
Get predefined keypoint combination and range of joint angles.
Calculate θ and ϕ
Convert θ to 0-100% range using one-dimensional linear interpolation for monotonically increasing sample points.
Output and Feedback Phase:
Display the trainee's performance based on predefined difficulty level.

Exercises	Keypoints											
	5	7	9	11	13	15	6	8	10	12	14	16
Barbell Shrugs	[192°,320°]			-			[192°,320°]			-		
Bicep Curls	[010°,150°]			-			[010°,150°]			-		
Pushups	[210°,280°]			-			[210°,280°]			-		
Shoulder Lateral Raise	[171°,194°]			-			[171°,194°]			-		
Shoulder Front Raise	[145°,290°]			-			[145°,290°]			-		
Shoulder Press	[026°,180°]			-			[026°,180°]			-		
Chest Press Rope	[186°,199°]			-			[186°,199°]			-		
Upper Chest Press	[177°,285°]			-			[177°,285°]			-		
Lateral Pulls	[191°,327°]			-			[191°,327°]			-		
Pullups	[142°,321°]			-			[142°,321°]			-		
Lunges	-			[142°,321°]			-			[153°,066°]		
Squats	-			[220°,280°]			-			[220°,280°]		
Horizontal Leg Press	-			[157°,086°]			-			[157°,086°]		
Vertical Leg Press	-			[252°,170°]			-			[252°,170°]		
Chest Press	-			[151°,108°]			-			[151°,108°]		

Table 3.1: Keypoint combination for different sports exercises. The proposed range of joint angles can be adjusted easily by sports experts as the model's hyperparameters.

3.3 Human Key Points

Human pose estimation is a fundamental problem in computer vision with various applications, such as action recognition, human-computer interaction, and surveillance. [21, 31] For instance, it can be used to track the body movements of patients during rehabilitation exercises or analyze athletes' techniques during training. It can also control a robot or a virtual avatar in real time based on the user's body movements. There are various standards for human key points used in pose estimation. One of the most commonly used standards is the COCO dataset [18], which includes a set of keypoint annotations for human pose estimation. The COCO key points localize to the essential human body joins, vital for practical applications like fitness and dance, without consuming much computation power. Another widely used standard is the MPII (Max Planck Institute for Informatics) Human Pose dataset, including annotations for the 16 key points [2]. These standards provide a consistent set of key points that can be used across different datasets and algorithms for accurate and reliable pose estimation. We present our pose topology used in the experiments in Figure 1.3. Note that those 17 key points differ from what has been described in [19], which we adapt to our implementation.

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.

$$\theta(u, v, p) = \frac{180(\phi(y_p - y_v, x_p - x_v) - \phi(y_u - y_v, x_u - x_v))}{\pi} \tag{3.1}$$

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.

$$\phi(y,x) = \begin{cases} \arctan(\frac{y}{x}), & \text{if } x > 0, \\ \frac{\pi}{2} - \arctan(\frac{x}{y}), & \text{if } y > 0, \\ -\frac{\pi}{2} - \arctan(\frac{x}{y}), & \text{if } y < 0, \\ \arctan(\frac{y}{x}) \pm \pi, & \text{if } x < 0, \\ \text{undefined}, & \text{if } x = 0 \wedge y = 0 \end{cases} \tag{3.2}$$

3.4 System Components

A robust and effective hardware configuration is required for the successful deployment of FitSight which utilizes the Yolov7 model for object detection. This section provides a comprehensive overview of the main elements of the system, including a detailed description of the precise hardware utilized and their configuration inside the laboratory setting. Notably, the layout closely resembles the setup employed in the Slackliner - An Interactive Slackline Training Assistant.[16]

3.4.1 System resources

The central component of the system is a formidable Windows-based PC, outfitted with an advanced NVIDIA GeForce RTX 2080 GPU, boasting 8 GB of dedicated memory. The selection of this high-performance GPU was driven by the imperative need for substantial computational power, indispensable for the execution of sophisticated machine learning algorithms and the management of real-time video data processing. The GPU's formidable capabilities are integral to the efficient operation of the YOLO model, facilitating uninterrupted object detection and activity recognition in the gym environment.

For visual data acquisition, the system employs an uncomplicated yet effective external USB webcam. This camera was chosen for its straightforward operation and versatile connectivity options, ensuring ease of setup and consistent performance. Positioned intentionally at the gym participants' frontal aspect, the camera's vantage point is pivotal in capturing a comprehensive view of the exercise activities, thus enabling the system to meticulously observe and analyze participants' movements and postures.

At the heart of the system resides an Intel® Core™ i7-8700K CPU, operating at a base frequency of 3.70 GHz, coupled with a substantial 32.0 GB of installed RAM. This robust configuration underscores the system's capability to handle the intensive demands of processing and analysis with alacrity. The PC runs on Windows 10 Pro, edition 22H2, providing a stable and secure operating environment for the system's requirements.

Visual feedback is rendered through a projector, boasting a resolution of 1920x1080, to display real-time activity assessments. This high-definition projector is not merely a medium for output but a conduit for enhancing the interactivity of the exercise session. It achieves this by projecting various measurements and visual cues linked to the participants' performance, thereby enriching the training experience with an engaging and educational dimension.

The amalgamation of a high-definition projector and a strategically placed webcam creates a synergistic interface between the technological components and the physical exercises performed in the gym. This integration allows for the effective monitoring and evaluation of physical movements and augments the participants' experience with interactive and real-time feedback.

3.5 Software Architecture

This section delves into the involved software architecture of the gym activity detection system, which is divided into two main components: the front-end user interface and the back-end processing unit. The utilization of a dual-structure method is crucial in order to attain a seamless integration of user interaction with advanced computational processing.

3.5.1 Front-End Development

The user interface, created using HTML and JavaScript, serves as the graphical display of the application, offering users a user-friendly and dynamic experience. The HTML structure serves as the framework for the user interface, guaranteeing a well-defined, easily navigable, and visually appealing design. JavaScript enhances this framework by providing dynamic functions, enabling the interface to promptly adapt to user inputs and system outputs.

The primary characteristics of the front-end encompass:

1. **Live Video Stream Display:** Integrating the camera stream seamlessly into the user interface, this feature allows users to observe their activities in real-time. The live video stream is crucial for users to monitor their own movements and get immediate visual feedback.

2. **Activity Feedback and Visualization:** The system presents data collected from the backend, such as movement analysis and repetition counts, in a user-friendly manner. This visualization makes it easy for users to understand their performance and progress.

3. **Interactive Elements:** The front-end facilitates user interaction with the system. This includes features like starting or stopping activity recognition, and modifying settings based on individual preferences, enhancing the system's usability and adaptability.

3.5.2 Backend Processing

The system's functionality is mostly based on the Python-developed back-end. This decision capitalizes on Python's capabilities in managing data-intensive operations and its extensive collection of machine learning libraries. The YOLOv7 model has been incorporated into this configuration due to its exceptional performance in real-time object recognition and tracking as described in the previous sections. The source code is available at: `https://github.com/hiteshkotte/DFKI-fitsight`

The primary responsibilities of the back-end involve:

1. **Real-Time Video Processing:** This involves analyzing the live video stream captured by the webcam. The process focuses on detecting various gym activities

and tracking the movements associated with them, crucial for accurate activity recognition.

2. **Data Interpretation and Analysis:** The back-end is responsible for converting unprocessed video data into meaningful insights. This includes quantifying repetitions, assessing the correctness of movements, and other vital metrics related to gym activities.

3. **Data Preparation for Front-End Display:** The system generates feedback by ensuring the accuracy and timeliness of the data being sent to the front-end. This step is key to providing users with real-time insights and visual feedback on their performance.

The combination of the Python environment and YOLOv7 forms a robust and effective backend system that can effectively manage the intricate requirements of real-time activity recognition.

3.5.3 Integration and Workflow

The seamless integration of the front-end and back-end is crucial for the system's effectiveness. This integration guarantees a seamless process where each component enhances the other.

1. **Data Capture and Processing:** As users engage in gym exercises in front of the webcam, the system captures this visual data. The data is then fed into the YOLOv7 model on the back-end for real-time processing and analysis, a critical step for accurate activity recognition.

2. **Feedback Mechanism:** The system translates the analysis into understandable feedback for the users. This includes textual feedback on performance, such as advice on improving form or posture, and a repetition counter that tracks exercise progress. This feedback is instantly displayed on the screen, enabling users to quickly receive and act upon valuable workout insights.

3. **User Interaction and Responsiveness:** The interactive nature of the front-end facilitates direct user engagement with the system. Users can start or stop the activity recognition, adjust settings, and interact with the feedback provided. This interactivity transforms the system from a simple monitoring tool into an engaging workout companion.

This integration highlights the system's capability to not only recognize and analyze gym activities but also to enhance the overall workout experience through interactive and real-time feedback. The upcoming sections will delve into specific features of the system, such as the repetition counter and feedback algorithms, offering a comprehensive understanding of how the system supports and enriches gym workouts.

3.6 Interface Overview:

FitSight's user interface is designed for intuitive interaction, providing users with clear visual cues and real-time feedback to facilitate their gym activities.

- **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 keypoints 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.

3.6.1 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. The feedback is received to the front end from the IMPECT Platform which will be discussed in the later section.
- **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 served as a motivational tool, encouraging participants to persist and maintain proper form until all planned exercises were completed.

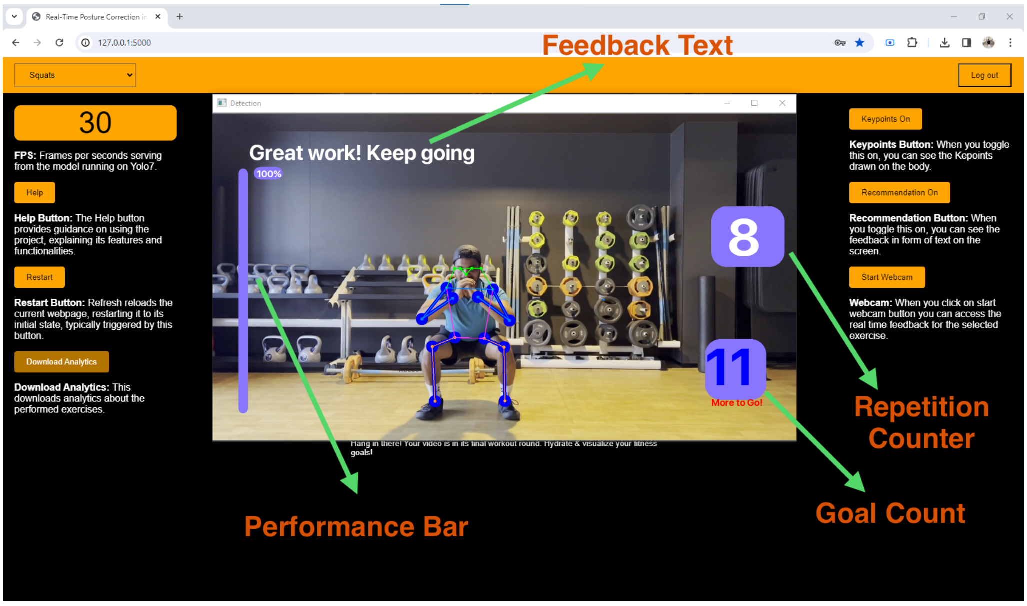

Figure 3.1: The real-time feedback system displaying the performance bar, feedback text, repetition counter, and goal tracker.

As depicted in Figure 3.1, these features collectively formed an integral part of the feedback loop, providing a comprehensive overview of performance 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 have the option to record their workout session and process the video later. This is ideal for those who prefer to review and reflect on their entire exercise routine post-workout, allowing them to assess their form and technique retrospectively and make improvements 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.

3.7 IMPECT System and Architecture

In this subsection, we describe the utilization of IMPECT (Immersive Multimodal Psychomotor Environments for Competence Training), a training toolkit that facilitates the development of immersive learning environments, leveraging sensors and immersive

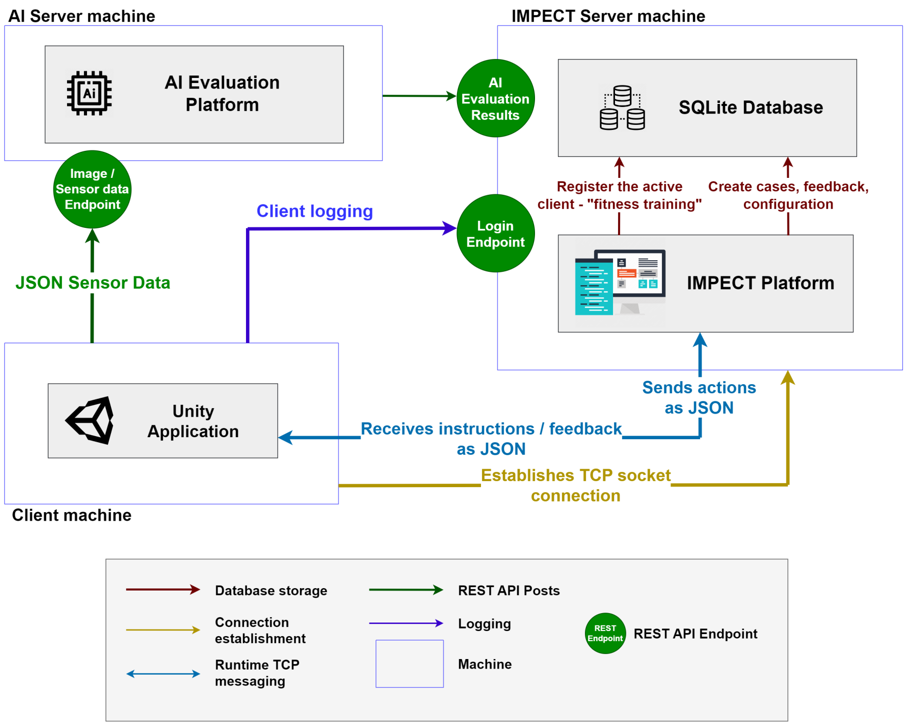

Figure 3.2: FitSight backend architecture.

technologies to cater to various psychomotor skills training needs [27]. The tool enables the transmission of image frames or sensor data from a client application equipped with sensor devices to the IMPECT server. The server is crucial in empowering the teacher with complete control over navigating the learning session. It also facilitates the creation of feedback cards based on identified mistakes. Consequently, the teacher can observe and assess the learner's performance by deciding which feedback card is the most suitable. In the context of our research, we extended this concept by integrating our YOLOV7 evaluation platform to evaluate and assess learner performance.

Figure 3.2 illustrates the system architecture and the data transmission flow. The IMPECT server operates as a Python Flask web application, incorporating a socket server and a REST API. An SQLite database is employed to gather and store data. The server-client connection is established through a TCP socket, transmitting information as JSON messages. Upon connection, the server registers the active client application in the database, assigns it a unique ID, and designates "fitness training" as its specific use case. The REST API endpoint facilitates the YOLOV7 evaluation platform's posting of evaluation results derived from the client's session recordings. The appropriate feedback card is dispatched to the client's application based on the identified mistake. Conversely, the Unity-based client application bridges the IMPECT server and the YOLOV7 evaluation platform. Initiating an active play session within the client application triggers the recording process from the camera and transmits the image frame data, identified with a session ID, to the evaluation platform. After completing the evaluation process, the client receives information from the IMPECT server and visualizes the feedback according to the selected feedback card.

Chapter 4
User Study

This chapter presents a detailed examination of the study. It begins with the *Introduction and Research Questions* section, where we outline the study's goals and hypotheses, clarifying its overall purpose. Following this, the *Participants* section details the profiles of the trainees involved, highlighting their relevant backgrounds and experiences.

In the *Method* section, we describe the study's framework, covering the conditions, apparatus used, design, procedures, and the variables involved. This lays the groundwork for understanding the study's approach.

4.1 Introduction and Research Questions

Understanding the importance of proper posture during exercise is key to both the safety and effectiveness of fitness routines. Typically, posture correction in gym settings relies on the guidance of fitness specialists. However, there's a growing interest in how advanced technology, especially machine learning models, can offer real-time assistance and feedback, potentially transforming this aspect of fitness training.

Our study embarks on a detailed exploration of these technological interventions. The main focus is to compare an interactive model, designed for posture correction, with the traditional training methods in a gym environment, especially in scenarios where continuous personal trainer support is not feasible. In many gyms, personal trainers, while beneficial, are often a luxury due to their cost. General trainers provide initial guidance but cannot offer sustained, individual attention. This common scenario leaves a gap in continuous posture supervision, which our study aims to address.

We are particularly interested in understanding the role of our real-time feedback system when a trainer is not present. The study hypothesizes that our system, acting as a virtual trainer, could effectively bridge the gap in supervision, offering consistent feedback akin to that of a personal trainer. This could be a crucial step in ensuring better posture and, consequently, more effective and safe workouts.

The primary goal is to scrutinize and contrast how individuals perform using our model compared to the traditional approach, with an emphasis on maintaining and improving posture quality. We're investigating whether our method can noticeably enhance posture during free weight exercises, compared to standard techniques that lack continuous technological support.

This research contributes to the broader conversation in Human-Computer Interaction (HCI) and fitness technology, focusing on the potential of interactive models to supplement, and in some cases, replicate the advantages of personal training in enhancing exercise form and posture.

Participant ID	Fitness Level	Gender	Age
P1	Intermediate	Male	24
P2	Intermediate	Female	28
P3	Beginner	Female	19
P4	Intermediate	Male	20
P5	Beginner	Male	19
P6	Beginner	Male	21
P7	Intermediate	Male	20
P8	Intermediate	Male	22
P9	Intermediate	Male	28
P10	Intermediate	Male	24
P11	Intermediate	Male	27
P12	Intermediate	Male	27
P13	Beginner	Male	24
P14	Beginner	Female	24
P15	Beginner	Male	25
P16	Intermediate	Female	26

Table 4.1: Demographic data of participants

Before the study began, we made sure each participant understood what the study involved and what was expected of them. Everyone agreed to be part of the study and to have their exercise sessions recorded on video. This was all done with the approval of the local ethics committee, ensuring that our study met all the required standards for ethical research.

Each *Participant ID* represents a unique individual in the study, ensuring their personal information remains private. A detailed breakdown of the participants' demographic information can be found in Table 4.1.

4.2 Method

Our study was supported by a diverse group of sixteen individuals from the *Informatics Campus* of Saarland University. This group included a mix of male and female participants, specifically twelve men and four women. The recruitment process, carried out via a campus-wide mailing list, attracted more male than female participants, likely reflecting the prevailing demographic trends in the campus's technical and scientific community. The research received ethical approval from the Ethics Review Board at Saarland University, and the approval document is included in the appendices.

The fitness experience of the participants spanned from novices to those with a moderate amount of gym experience. The participants were classified as either *Beginners*, who were new to structured exercise routines, or *Intermediates*, who had established a foundation of regular training and were familiar with exercise techniques.

The decision to exclude expert-level participants was intentional. 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. Including expert-level participants could skew the results, as their high baseline of posture correctness would not provide a significant scope for improvement through the model. Therefore, this categorization was chosen to capture a wide range of abilities and to ensure the technology being tested would have broad relevance.

4.2.1 Grouping and Balance

A balanced demographic profile is essential in experimental research to mitigate potential biases that may arise from uneven distribution of participant characteristics. Therefore, both groups were composed to have an equal representation of gender and fitness levels. Such equilibrium in group composition helps ensure that any differences observed between the groups can be confidently attributed to the intervention being tested rather than underlying demographic disparities.

In an effort to ensure a robust experimental design, participants in the study were stratified into two distinct groups with 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. Group A was designated as the Feedback Group, while Group B was the No Feedback Group.

ID	Level	Gender
P1	Intermediate	Male
P2	Intermediate	Female
P4	Intermediate	Male
P6	Beginner	Male
P9	Intermediate	Male
P11	Intermediate	Male
P14	Beginner	Female
P15	Beginner	Male

(a) Group A (Feedback)

ID	Level	Gender
P3	Beginner	Female
P5	Beginner	Male
P7	Intermediate	Male
P8	Intermediate	Male
P10	Intermediate	Male
P12	Intermediate	Male
P13	Beginner	Male
P16	Intermediate	Female

(b) Group B (No Feedback)

Table 4.2: Group Demographics

By mirroring the gender and fitness level proportions across both groups, the study aimed to achieve a level playing field. This balance allows for a clearer interpretation of the effect that real-time feedback has on exercise performance, isolating the variable of interest from confounding demographic factors. Table 4.2 outlines the specific demographic makeup of each group, illustrating the parity achieved in the study's design.

4.2.2 Instructional Setup

Both groups were introduced to the study's exercises through a standardized demonstration session, conducted to replicate a typical gym environment with a general trainer.

This session was designed to provide all participants with an equal foundation of understanding regarding the exercises, their form, and the expected posture.

Participants were encouraged to actively engage with the trainer, asking questions to clarify their understanding of the correct posture and to seek guidance on exercise execution. They were also permitted to perform one or two 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.

4.2.3 Experimental Conditions of FitSight Training

The design of this study was structured to meticulously evaluate the impact of FitSight on the improvement of exercise posture. Specifically, the research aimed to discern the effectiveness of posture correction strategies both with the inclusion of a technological aid and in its absence. To this end, the experimental framework was methodically crafted to facilitate a comparison under two distinct conditions.

Under the first condition, participants engaged in exercise routines with technological assistance, designed to offer immediate feedback on their posture. The second condition replicated the exercise sessions without such assistance, placing participants in a more conventional workout setting. This comparative approach allowed the study to provide insights into the potential benefits of incorporating FitSight into fitness training regimes.

The overarching goal was to determine whether FitSight could significantly enhance the accuracy of participant's posture during exercise, thereby contributing to safer and more effective workout practices. The following sections detail the specific nature of these two conditions, the setup of the exercises, and the manner in which participant performance was evaluated.

Condition 1 - Feedback Group

In the Feedback Group condition, we had eight participants perform 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. This system of feedback, which will be elaborated on in later sections, allowed participants to gauge their accuracy and make immediate corrections as needed.

Condition 2 - No Feedback Group

Group B, also consisting of eight participants, was instructed to perform the identical set of exercises as the previous group. However, in contrast to Group A, 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. This allowed participants to self-monitor to some extent while exercising. After completing the set of exercises, the participant's performance was evaluated by the fitness expert, who provided a score for each individual on a scale from 1 to 10, assessing their posture and technique without the aid of interactive feedback.

The direct comparison of these two conditions was pivotal in evaluating the FitSight's role in exercise execution. By ensuring both groups performed the same exercises in an identical order, the study aimed to isolate the effect of the interactive feedback from other variables, offering a clear perspective on the technology's impact on exercise posture and form.

4.3 Procedure

The procedure for the study was meticulously designed to ensure consistency and reliability of the data collected. Participants were welcomed warmly and were first provided with the "Participant Registration and Consent Form." It was imperative for the integrity of the study that participants were fully informed and had given their consent prior to participation.

Upon completion of the consent form, the study was explained in detail to the participants, and they were handed the "Participant Instructions Form." To gain insights into the participant's background, they were requested to fill out the "Preliminary Questionnaire." This initial data collection was crucial in assessing the participant's suitability for either the feedback or control group, with a particular focus on maintaining a balance of gender and fitness levels within the study.

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 key to preventing injury and ensuring the exercises were challenging yet achievable.

The study session commenced with the 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 had a clear understanding of the exercises they were about to perform.

The exercise session then began, with participants performing the exercises while their posture was evaluated by the trainer. 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 exercise they had just performed, and the researcher noted down their responses.

This cycle of exercise, evaluation, and survey was repeated for each of the six exercises in the session. Upon completion of 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.

The detailed procedure outlined above was designed to ensure that each participant's experience was standardized, thereby allowing for reliable comparisons to be made between the different groups within the study.

4.4 Apparatus

The apparatus for the study was designed to closely replicate a typical gym environment, providing a familiar setting for participants to perform exercises. The central feature of this setup was a projector screen, connected to a Windows system, which displayed real-time feedback and motivational cues. This feedback was projected in a large format in front of the participants, facilitating easy visibility during exercises. Notably, the layout closely resembles the setup employed in the Slackliner - An Interactive Slackline Training Assistant.[16]

Positioned in front of the exercising area, a webcam was mounted to record the participant's movements. Its placement was carefully chosen to ensure a clear and unobstructed view of the exercises, capturing details necessary for subsequent posture analysis. The designated exercise area was marked on the ground to standardize the location where participants stood, ensuring uniformity in the recording perspective. The floor was covered with interlocking gym mats, providing a stable and non-slip surface suitable for physical activity.

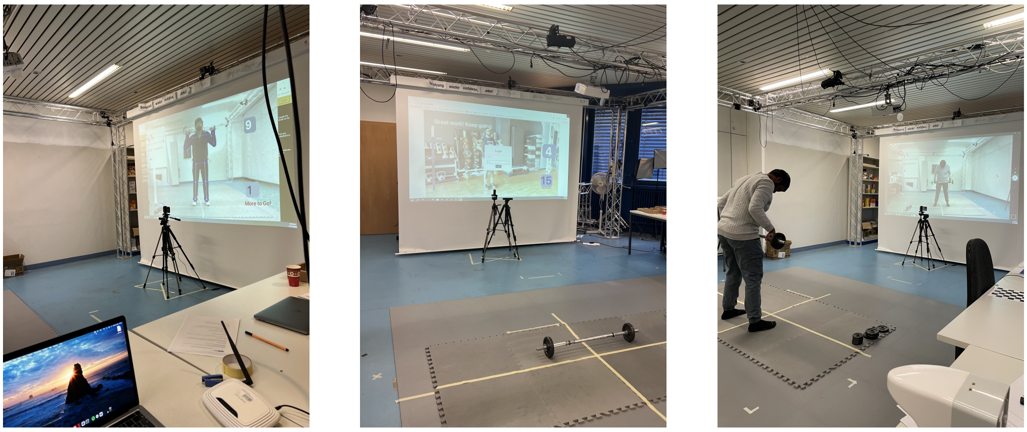

Figure 4.1: Experimental setup for the user study

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 that would allow participants to perform as they would in their regular gym sessions, thus providing authentic data for the study's evaluation criteria.

4.5 Variables

In any experimental study, it is crucial to define and distinguish between independent and dependent variables. The independent variables are those that are manipulated by the researcher to observe their effect on the dependent variables, which are the outcomes measured in the study.

By analyzing the dependent variables in relation to the independent variables, the study aims to draw conclusions about the impact of real-time feedback on exercise performance and participant perception. This analysis will provide insights into the potential benefits of integrating such feedback systems into regular fitness routines.

4.5.1 Independent Variables

In this study, the primary independent variable is the type of feedback provided to participants during their exercise sessions. This variable consists of two levels: feedback through the interactive model and no feedback, which corresponds to the traditional exercise approach. The allocation of participants to either the Feedback Group (Group A) or the No Feedback Group (Group B) represents the operationalization of this independent variable. Additionally, the weight of the dumbbells used during the exercise session was tailored to the participant's fitness level, representing a secondary independent variable.

4.5.2 Dependent Variables

The dependent variables in this study are the outcomes that reflect the efficacy of the feedback system. These include the accuracy of the participant's posture during exercises, as rated by a fitness expert on a scale of 1-10, and the participant's self-reported satisfaction and perceived effectiveness of the exercise session. Further, the improvement in the participant's performance over time, measured through the change in posture ratings across the exercises, is another critical dependent variable. These measures are intended to capture both the objective and subjective effects of the feedback system on the participant's exercise experience.

Hypothesis here and research questions in the intro

.

Chapter 5
Results and Analysis

Data from the user study was analyzed using various metrics to test our hypothesis that Group A, receiving feedback, outperforms Group B, which did not receive feedback. A simple mean comparison showed that the Feedback group scored an average of 42.1, while the No Feedback group scored 38.4, affirming our preliminary hypothesis. To further understand participant trends, we conducted detailed statistical analyses, including ANOVA and Tukey's HSD Post-Hoc Tests.

These methods helped identify significant performance differences among different participant groups—beginners, intermediates, and based on gender. ANOVA was crucial for detecting these differences, while Tukey's HSD Post-Hoc Tests provided deeper insights into the specific nature of these disparities. These analyses are key to understanding how training levels and gender influence performance outcomes.

This chapter presents details on the user study, including a section on **ANOVA** where we interpret the findings and compare groups using **Tukey's HSD**, detailed in the following section. The section on **Inter-Rater Reliability** addresses potential biases in trainer's ratings of posture.

5.1 Analysis of Variance (ANOVA)

APA style for reporting

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, individuals with intermediate skill levels, and grouped by gender.

ANOVA for Beginner vs Intermediate participants

The initial nature ANOVA analysis showed a group sum of squares (Sum Sq) of 75.0, which indicates the amount of variability caused by the treatment effect. The related degree of freedom (df) is 1, showing the number of treatment levels minus one. The computed F-statistic was 6.122449, representing the ratio of systematic variance to unsystematic variation. The p-value (PR(>F)) was 0.068625, showing a statistical trend that suggests potential importance, although it did not exceed the standard alpha level of 0.05. The p-value is the likelihood of detecting an F-statistic of this magnitude if the null hypothesis were true, indicating the potential presence of a treatment effect that requires additional investigation.

The ANOVA results were more prominent for the Intermediate group. The group sum of squares was much larger at 102.704167, showing that the therapy accounted for a greater amount of variability. Given a degree of freedom (df) of 1, the F-statistic significantly increased to 36.483256, indicating that the treatment impact is strong in comparison to the variability caused by errors. The observed p-value of 0.000309 is significantly lower than the predetermined alpha level of 0.05. This provides compelling evidence to reject the null hypothesis, showing a highly statistically significant difference between the group means as a result of the treatment.

ANOVA for Male vs Female participants

The male participants' group sum of squares was 14.116667, with a degree of freedom of 1. The F-statistic of 0.848868 falls below the significance level, indicating that the variation caused by the treatment is not significantly different from the variance caused by error. The p-value of 0.378564 above the alpha level of 0.05, suggesting that there is insufficient evidence to reject the null hypothesis for male participants. Therefore, no significant treatment effect is seen.

The total of squares for the female participant group was 93.520833, with a degree of freedom (df) of 1. The F-statistic yielded a value of 10.020089, signifying a significant treatment effect. The p-value obtained was 0.086977, indicating a slight inclination towards statistical significance. Although this does not satisfy the usual requirements for statistical significance, it does indicate that the treatment's impact is nearing significance and should be further explored, especially with a larger sample size.

Group	Variation Source	Sum Sq	df	F	P-value
Beginners	Between	75.0	1	6.12	0.069
	Within	49.0	4	-	-
Intermediates	Between	102.70	1	36.48	0.0003
	Within	22.52	8	-	-
Male	Between	14.12	1	0.85	0.379
	Within	166.30	10	-	-
Female	Between	93.52	1	10.02	0.087
	Within	18.67	2	-	-

Table 5.1: Detailed ANOVA Results for Different Participant Groups

The table 5.1 provides a comprehensive summary of 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 assumption that the null hypothesis is true. A p-value below the threshold of 0.05 typically indicates statistical significance.

5.1.1 Interpretation

The ANOVA results provide us with insights into the treatment's influence on various groups. Although the beginners and female participants exhibited indications of potential significance, implying that the treatment might have an impact that could be more evident with bigger 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 is not 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, in order to identify the precise differences between groups. These findings not only add to the current body of knowledge but also provide insights for future experimental designs and potential practical applications that consider gender and skill level disparities.

5.2 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 5.2: 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 5.3: Tukey's HSD Post-Hoc Tests: Intermediate Group Comparison

5.2.1 Overall Group Comparison

The comparison between groups A and B resulted in a mean difference of -0.3125, showing that group A has a marginally lower mean compared to group B. 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 that there is a strong likelihood that the observed difference in means may have been caused by random sampling variation, assuming that the true difference between the groups is actually 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 fails to yield adequate data to establish a notable impact of the conditions or interventions represented by groups A and B.

5.2.2 Intermediate Group Comparison

In contrast to the overall group comparison, the intermediate group comparison shows a substantial mean difference of -6.5417 between groups A and B, with a p-value of 0.0003. 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 substantial 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 true mean difference, confirming the observed substantial difference. The negative sign of the mean difference indicates that the mean of group A is significantly lower than that of group B, with statistical significance.

5.3 Observations of the User Study

In this particular section, we will look at the feedback obtained from users, which was gathered through a carefully structured procedure. The study participants were requested to provide their answers to a sequence of questions after each exercise session. The main objective of these inquiries was to determine the perceived level of difficulty of each exercise and assess the adequacy and helpfulness of the instructor's demonstrations.

Additionally, after completing all exercises, every participant was asked to fill out a thorough user survey. The purpose of this survey was to incorporate open-ended questions in order to ensure inclusivity. The inclusion of such questions aimed to acquire more profound and detailed insights into their experiences with FitSight. The inclusion of open-ended questions was vital in enabling participants to openly articulate their opinions and offer comprehensive feedback, which is essential for obtaining a comprehensive knowledge of the efficacy and user-friendliness of FitSight. This feedback is extremely essential as it provides an authentic viewpoint from the users, illuminating both the program's strengths and places where there is room for advancement.

5.3.1 Feedback and Rate of Perceived Exertion (RPE)

The trainer kept a close eye on each participant's posture during the exercise session. Upon the completion of an exercise, a 2-minute rest period was allotted to each participant. This rest period served not only as a physical recovery but also as an opportunity to gather immediate feedback through a short verbal survey. The survey comprised two 5-point scale questions designed to assess the participant's confidence and the effectiveness of the instructor's demonstration:

1. "I felt confident while performing the exercises."
2. "The demo by the instructor helped me maintain the correct posture during the exercises."

Participants responded to these statements on a 5-point scale, ranging from "Strongly Disagree" to "Strongly Agree." These questions were posed after each exercise to capture consistent and exercise-specific feedback.

The examination of the user survey provides significant observations regarding the degrees of confidence exhibited by participants in various exercises within Groups A and B. Both Group A, who received feedback, and Group B, which did not, showed similar confidence scores during the bicep curl exercise, suggesting that feedback had little effect on confidence for this specific task.

Exercise	**Group A Total**	**Group B Total**
Bicep Curl	37	36
Squats	33	30
Shoulder Lateral Raise	31	35
Bent Over Row	33	33
Lunges	32	33
Shoulder Press	37	33

Table 5.4: Total Confidence Scores by Exercise and Group

Group A exhibited a bit higher trust during squats, indicating that feedback may have had a minor positive impact. In contrast, Group B exhibited a somewhat higher level of confidence during the shoulder lateral lift. This phenomenon can be ascribed to the feedback mechanism's inclination to display a progress bar of low magnitude when the participants' arms were not precisely aligned, hence impacting the confidence of Group A.

Both groups had identical levels of confidence in their performance of the bent-over row, indicating that the execution of the exercise may not have been significantly influenced by the feedback given. Similarly, the lunges exhibited minimal variation in confidence levels among the groups, so corroborating the idea that feedback did not have a substantial impact on this exercise.

However, a significant disparity was noted in the shoulder press exercise, with Group A exhibiting greater confidence compared to Group B. This suggests that feedback may have a greater impact on activities that necessitate a significant level of precision in form, such as the shoulder press.

Presented in table 5.4 summarizes the overall confidence scores for each exercise in both groups:

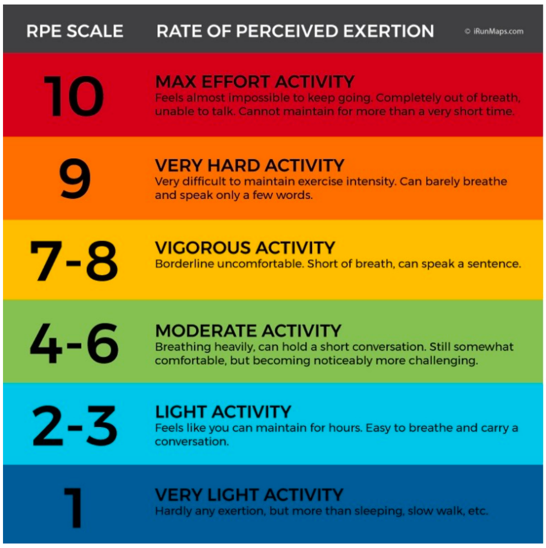

Figure 5.1: The Rating of Perceived Exertion (RPE) scale.

Additionally, to assess the exertion level of participants and maintain uniformity in physical effort across the exercises, an RPE (Rating of Perceived Exertion) score was also recorded after each exercise. The RPE scale is a standard tool used to estimate an individual's effort and exertion during physical activity. To ensure comparability of effort and to control for exertion bias, it was crucial that participants operated within a similar range on the RPE scale.

Group	**E1**	**E2**	**E3**	**E4**	**E5**	**E6**
Feedback Group	2	2.75	3.125	2.5	2.75	2.25
No Feedback Group	2	2.625	2.75	2	3.125	2.375

Table 5.5: Average RPE for Feedback and No Feedback Groups for exercises E1 to E6

Table 5.5 illustrates that the intensity of all exercises remained consistently scored between 2 and 3 on the Rate of Perceived Exertion (RPE) scale, indicating a level of light activity. The intentional maintenance of uniformity in the difficulty level was aimed at ensuring similarity among individuals' exercise experiences. The significance of this technique is two-fold: firstly, it takes into account the fact that participants were not wearing sportswear, as the study was conducted in a laboratory setting rather than a fitness center. Furthermore, the data unambiguously indicates that there was minimal difference in perceived exertion between the two separate groups, namely those who received feedback and those who did not. The absence of any notable variation suggests that the

RPE scale was consistently and without bias used within the group, hence strengthening the dependability of the gathered data in evaluating the intensity of the recommended workouts.

5.4 Inter-Rater Reliability

The user study featured a key trainer responsible for both demonstrating and setting the appropriate weights for each participant's exercises. Additionally, this trainer assessed the posture of each participant during the exercises.

However, this arrangement presented a potential issue of bias, as the trainer was aware of each participant's specific conditions. To ensure the evaluations were unbiased, it was crucial to first determine if the trainer exhibited any bias. To address this, two additional trainers were brought on board. These trainers, who were graduate students with a focus in Sports Science and expertise in weight training, independently rated the participants' exercises to ensure objectivity.

The inter-rater reliability among the three trainers was quantified using Spearman's rank correlation coefficient. This non-parametric statistic is ideal for data with ordinal properties, allowing for the evaluation of rank correlations without assuming a linear relationship. Spearman's rho ranges from -1 to 1, where 1 represents perfect positive correlation and -1 indicates perfect negative correlation. The correlation matrix detailed in Table 5.6 showcases these relationships, with the overall agreement score derived from the average of the pairwise correlations.

The analysis of the pairwise Spearman's rank correlations among the three trainers yielded an overall agreement score of $r = 0.6746$, indicative of a generally strong consistency in their evaluations. The specific pairwise correlations were as follows:

- Between Trainer 1 and Trainer 2, the correlation coefficient was $r = 0.6181$.
- Between Trainer 1 and Trainer 3, a higher correlation of $r = 0.7234$ was observed.
- The correlation between Trainer 2 and Trainer 3 was $r = 0.6823$.

These results demonstrate a moderate to strong positive agreement across all trainer evaluations, suggesting that the ratings were consistently aligned.

This assessment is crucial for validating the reliability and consistency of the evaluation process used in the study. By confirming that the trainers' ratings are harmonious and free from individual biases, we can assure stakeholders of the integrity and objectivity of the outcomes.

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

Chapter 6
Discussion

6.1 Real-Time Feedback in Exercise Performance

The investigation into real-time feedback mechanisms for exercise performance represents a promising avenue for advancement in gym activities. The utilization of advanced object detection algorithms, such as the YOLO model, for gym activity recognition offers a unique opportunity to assess the impact of real-time feedback on workout posture and effectiveness. This study has been dedicated to exploring these components with the goal of providing valuable insights into the potential advantages of integrating technological augmentation into training routines. The findings of this research have profound implications for the development of AI-supported training systems and their integration into regular exercise regimens.

An essential aspect of this study was to replicate a gym environment in a controlled laboratory setting, carefully designed to closely mimic the typical conditions found in a gym. One of the critical considerations was to maintain a consistent Rate of Perceived Exertion (RPE) among all participants. This approach was crucial to minimize potential bias in the perceived difficulty of exercises, which could have influenced the study's results significantly. By ensuring uniform RPE levels, the study aimed to isolate the impact of real-time feedback on exercise performance.

To further enhance the study's integrity and minimize bias, the trainer meticulously assigned weights to each participant, taking into account their varied Body Mass Index (BMI) values. This personalized approach to weight distribution was founded on logical and unbiased principles, aligning with common practices in sports science. The systematic allocation of weights was validated through the RPE results, confirming that all exercises were executed at a low intensity level, guaranteeing that participants did not experience fatigue. Additionally, short intervals of 1-2 minutes between workouts were introduced to reduce fatigue, preserving the accuracy and reliability of the study's outcomes.

6.1.1 Evaluation Findings and Inter-Rater Reliability

The ANOVA results from the intermediate group present remarkable insights, with a p-value of 0.0003, significantly lower than the recognized significance level of 0.05. These findings provide strong evidence of significant differences in exercise performance between Groups A and B, with Group A demonstrating superior outcomes. This outcome supports the hypothesis that real-time feedback positively influences exercise performance, particularly among athletes at an intermediate level.

Conversely, the p-value of 0.069 for the beginner group, while approaching statistical significance, does not conclusively establish a significant difference in posture correction between the groups. This result can be attributed to the limited knowledge of novice participants in gym workouts. Novices may require additional instructional support beyond real-time feedback at the initial stages of their fitness journey. Future research might explore more extensive assistance, potentially incorporating advanced visual aids or augmented reality avatars, to effectively improve posture and workout technique among beginners.

The analysis of exercise performance heavily relied on the assessment scores provided by Trainer 1. The analysis of the pairwise Spearman's rank correlations among the three trainers yielded an overall agreement score of $r = 0.6746$, indicative of a generally strong consistency in their evaluations. The reliance on Trainer 1's judgments is supported by the results of inter-rater reliability tests, which demonstrated that Trainer 1's assessments were unbiased.

6.1.2 Implications for Different User Groups

In examining the intermediate group, it becomes evident that real-time feedback complements their pre-existing knowledge and experience in gym exercises, resulting in a more precise execution and maintenance of correct posture during workouts. This observation is supported by the notable statistical disparities in performance results. The deliberate exclusion of the expert group from the user study is based on the assumption that their considerable expertise would yield only marginal improvements in posture through real-time feedback. This assumption is corroborated by the lack of significant results in the Tukey's HSD post-hoc analysis when comparing the entire group, as the adjusted p-value exceeded the predetermined significance level.

However, narrowing the analysis to the intermediate group unveils a significant difference, affirming the prediction that the real-time feedback system significantly enhances performance for individuals with intermediate exercise experience rather than extensive expertise. These insights underscore the importance of adapting fitness technology to cater to the diverse needs and abilities of various user groups. Real-time feedback technologies emerge as a valuable complement to the training programs of intermediate-level athletes.

6.2 Future Directions and Broader Implications

In conclusion, this research sheds light on the intricate relationship between training experience and the effectiveness of real-time feedback in exercise environments. This observation raises important questions regarding how to effectively assist individuals at all fitness levels using adaptive, intelligent technologies that enhance physical performance and adherence to proper technique. Future research avenues may explore the integration of the FitSight application with wearable technologies, its application in rehabilitation settings, and long-term studies to assess its impact on fitness progression. Additionally, addressing the identified limitations, such as diversifying the exercise repertoire and incorporating more objective measurement tools, will be vital for advancing this field.

This study contributes to the broader discourse on human-computer interaction and fitness technology, emphasizing the potential of interactive models to augment, and in some cases, replicate the advantages of personal training in improving exercise form and posture. As we continue to explore the synergies between computer science and user experience, it becomes evident that the journey toward more interactive, intelligent fitness solutions is just beginning.

.

Chapter 7
Conclusion and Future Work

This thesis has ventured into the intersection of computer vision and physical fitness, aiming to harness the potential of advanced object detection algorithms for enhancing gym workout experiences. The FitSight system, a cutting-edge AI solution integrating the YOLOv7 model for precise body posture estimation, marks a significant leap in fitness training. Our detailed evaluation with 16 participants underscores FitSight's effectiveness in boosting exercise performance, particularly through its real-time feedback feature. This feedback is pivotal in maintaining correct posture and minimizing accident risks, enhancing workout results. This concluding chapter reflects on the contributions, findings, implications, and future directions of this research.

7.1 Conclusion

The contribution of this thesis lies in its successful application of the YOLOv7 object detection model to the realm of fitness training, marking a significant advancement in the field of computer science and user experience research. The development of the FitSight application represents a pioneering approach to integrating real-time feedback mechanisms into exercise routines, offering a novel solution to the common problem of maintaining correct exercise posture without professional supervision.

Our investigation revealed that real-time feedback significantly enhances exercise performance, particularly among intermediate-level participants. By simulating a gym environment within a controlled laboratory setting, the study demonstrated that participants receiving real-time feedback exhibited notable improvements in maintaining correct exercise posture and technique. These findings underscore the effectiveness of the FitSight application in providing actionable, precise feedback, thereby contributing valuable insights into the use of technology-assisted training methods in physical fitness.

The study's emphasis on the intermediate group showed a notable improvement, indicating that real-time feedback significantly boosts performance in moderately experienced exercisers, not just those with advanced skills. This highlights the importance of fitness technology adapting to the diverse needs and abilities of various user groups, particularly benefiting intermediate-level athletes with real-time feedback.

FitSight's real-time feedback has proven instrumental in self-monitoring during workouts, particularly in the absence of professional supervision. The data gathered provided deep insights into the varied needs of fitness enthusiasts, showcasing the system's broad applicability and its capability to personalize AI-driven fitness solutions.

The practical implications of this research are manifold. For developers and researchers in computer science and fitness technology, the FitSight application serves as a prototype for future developments in AI-supported fitness training systems. Its potential for scalability and adaptation across various fitness levels and settings suggests a promising avenue for the broader implementation of similar technologies, aiming to make exercise safer and more effective for a wider audience.

7.2 Limitations and Future Work

Despite its successes, this study faced limitations, including the sample size and the diversity of exercises tested. Additionally, the reliance on manual assessments by trainers introduced subjectivity into the evaluation process, although measures such as inter-rater reliability tests were employed to mitigate bias. These limitations suggest caution in generalizing the findings without further research.

The current study relies primarily on a trainer's evaluations to assess human posture, which introduces potential limitations regarding subjectivity and bias. Future research could greatly benefit from developing a more objective and reliable metric for posture assessment. One promising direction is the automation of posture evaluation through the use of angular measurements, which could provide quantifiable and reproducible data. While the integration of such technology represents a potential advancement in eliminating human bias and streamlining the assessment process, it falls outside the scope of this thesis and warrants further investigation.

Looking forward, enhancing FitSight's precision remains a primary objective. This research aims to tailor the system to a wider customer base, boosting its adaptability. Future tests in real fitness centers and long-term studies will be crucial to understand the system's practical utility, user engagement, and its long-term impact on fitness motivation and outcomes.

Future research could also expand on this work by exploring the integration of FitSight with wearable technologies, investigating its application in rehabilitation, and conducting long-term studies to assess its impact on fitness progression. Moreover, addressing the limitations identified in this study, such as expanding the exercise repertoire and incorporating more objective measurement tools, will be crucial.

In conclusion, this thesis underscores the significant potential of applying computer vision and machine learning technologies to address real-world challenges in physical fitness. The FitSight application exemplifies how technological innovation can bridge the gap between the expertise of personal trainers and the autonomy of individual workouts, offering a new perspective on the future of fitness training. As we continue to explore the synergies between computer science and user experience, it is clear that the journey towards more interactive, intelligent fitness solutions is just beginning.

Bibliography

[1] Fraser Anderson, Tovi Grossman, Justin Matejka, and George Fitzmaurice. 2013. YouMove: enhancing movement training with an augmented reality mirror. In *Proceedings of the 26th Annual ACM Symposium on User Interface Software and Technology (UIST '13)*. Association for Computing Machinery, New York, NY, USA, 311–320. DOI:http://dx.doi.org/10.1145/2501988.2502045

[2] Mykhaylo Andriluka, Leonid Pishchulin, Peter Gehler, and Bernt Schiele. 2014. Human Pose Estimation: New Benchmark and State of the Art Analysis. In *Proceedings IEEE Conf. on Computer Vision and Pattern Recognition (CVPR)*. IEEE, 3686 – 3693. DOI:http://dx.doi.org/10.1109/CVPR.2014.471

[3] Fabio Buttussi, Luca Chittaro, and Daniele Nadalutti. 2006. Bringing Mobile Guides and Fitness Activities Together: A Solution Based on an Embodied Virtual Trainer. In *Proceedings of the 8th Conference on Human-Computer Interaction with Mobile Devices and Services (MobileHCI '06)*. Association for Computing Machinery, New York, NY, USA, 29–36. DOI:http://dx.doi.org/10.1145/1152215.1152222

[4] Yu Chen and Pearl Pu. 2014. HealthyTogether: Exploring Social Incentives for Mobile Fitness Applications. In *Proceedings of the Second International Symposium of Chinese CHI (Chinese CHI '14)*. Association for Computing Machinery, New York, NY, USA, 25–34. DOI:http://dx.doi.org/10.1145/2592235.2592240

[5] An Cong, Nghi Tran, and Nghia Duong-Trung. 2020. Recognition and Quantity Estimation of Pastry Images Using Pre-training Deep Convolutional Networks. 200–214. DOI:http://dx.doi.org/10.1007/978-981-33-4370-2_15

[6] Nghia Duong-Trung, Hitesh Kotte, and Milos Kravcik. 2023. Augmented Intelligence in Tutoring Systems: A Case Study in Real-Time Pose Tracking to Enhance the Self-learning of Fitness Exercises. 705–710. DOI:http://dx.doi.org/10.1007/978-3-031-42682-7_65

[7] Nghia Duong-Trung, Luyl Da Quach, and Chi-Ngon Nguyen. 2019. Learning Deep Transferability for Several Agricultural Classification Problems. *International Journal of Advanced Computer Science and Applications* 10 (02 2019). DOI:http://dx.doi.org/10.14569/IJACSA.2019.0100107

[8] Daniel Eaves, Gavin Breslin, Paul Schaik, Emma Robinson, and Iain Spears. 2011. The Short-Term Effects of Real-Time Virtual Reality Feedback on Motor Learning in Dance. *Presence* 20 (02 2011), 62–77. DOI:http://dx.doi.org/10.1162/pres_a_00035

[9] Çağatay Erdaş and Selda Güney. 2021. Human Activity Recognition by Using Different Deep Learning Approaches for Wearable Sensors. *Neural Processing Letters* 53 (02 2021). DOI:http://dx.doi.org/10.1007/s11063-021-10448-3

[10] Evianita Fajrianti, Sritrusta Sukaridhoto, Udin Al Rasyid, Bambang Suwito, Rizqi Putri Nourma Budiarti, Ilham Hafidz, Naufal Satrio, and Amma Haz. 2022. Application of Augmented Intelligence Technology with Human Body Tracking for Human Anatomy Education. *International Journal of Information and Education Technology* 12 (01 2022), 476–484. DOI:http://dx.doi.org/10.18178/ijiet.2022.12.6.1644

[11] 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 (01 2021), 107859. DOI:http://dx.doi.org/10.1016/j.comnet.2021.107859

[12] 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 (DIS '21)*. Association for Computing Machinery, New York, NY, USA, 1619–1637. DOI:http://dx.doi.org/10.1145/3461778.3462094

[13] Michail Giannakos, Daniel Spikol, Daniele Di Mitri, Kshitij Sharma, Xavier Ochoa, and Rawad Hammad. 2022. *Introduction to Multimodal Learning Analytics*. DOI:http://dx.doi.org/10.1007/978-3-031-08076-0_1

[14] Xinning Gui, Yu Chen, Clara Caldeira, Dan Xiao, and Yunan Chen. 2017. When Fitness Meets Social Networks: Investigating Fitness Tracking and Social Practices on WeRun. In *Proceedings of the 2017 CHI Conference on Human Factors in Computing Systems (CHI '17)*. Association for Computing Machinery, New York, NY, USA, 1647–1659. DOI:http://dx.doi.org/10.1145/3025453.3025654

[15] Justin Kim, Taylor Davis, and Lingzi Hong. 2022. Augmented Intelligence: Enhancing Human Decision Making. 151–170. DOI:http://dx.doi.org/10.1007/978-3-030-84729-6_10

[16] 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 (MM '18)*. Association for Computing Machinery, New York, NY, USA, 154–162. DOI:http://dx.doi.org/10.1145/3240508.3240537

[17] 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 (OzCHI '16)*. Association for Computing Machinery, New York, NY, USA, 517–526. DOI:http://dx.doi.org/10.1145/3010915.3010941

[18] Tsung-Yi Lin, Michael Maire, Serge Belongie, James Hays, Pietro Perona, Deva Ramanan, Piotr Dollár, and C. Zitnick. 2014. Microsoft COCO: Common Objects in Context.

[19] 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.

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

[21] Tewodros Legesse Munea, Zelalem Yalew, Halefom Tekle, Longbiao Chen, Chenxi Huang, and Chenhui Yang. 2020. The Progress of Human Pose Estimation: A Survey and Taxonomy of Models Applied in 2D Human Pose Estimation. *IEEE Access* PP (07 2020), 1–1. DOI:http://dx.doi.org/10.1109/ACCESS.2020.3010248

[22] Hristo Novatchkov and Arnold Baca. 2013a. Artificial Intelligence in Sports on the Example of Weight Training. *Journal of sports science medicine* 12 (03 2013), 27–37. DOI:http://dx.doi.org/10.1145/3010915.3010941

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

[24] 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. DOI:http://dx.doi.org/10.1109/ISPA-BDCloud-SocialCom-SustainCom51426.2020.00200

[25] 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). DOI:http://dx.doi.org/10.1109/JIOT.2018.2846359

[26] José A. Ruipérez-Valiente, Roberto Martínez-Maldonado, Daniele Di Mitri, and Jan Schneider. 2022. From Sensor Data to Educational Insights. (11 2022). DOI:http://dx.doi.org/10.3390/s22218556

[27] Khaleel Asyraaf Mat Sanusi, Daniel Majonica, Patrick Handwerk, Suhana Biswas, and Roland Klemke. 2023. Evaluating an immersive learning toolkit for training psychomotor skills in the fields of human-robot interaction and dance.. In *3rd International Workshop on Multimodal Immersive Learning Systems, MILeS 2023*. CEUR-WS, 70–78.

[28] 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 (CHI '22)*. Association for Computing Machinery, New York, NY, USA, Article 90, 16 pages. DOI:http://dx.doi.org/10.1145/3491102.3517735

[29] Mohammad Javad Shafiee, Brendan Chywl, Francis Li, and Alexander Wong. 2017. Fast YOLO: A Fast You Only Look Once System for Real-time Embedded Object Detection in Video. *Journal of Computational Vision and Imaging Systems* 3 (09 2017). DOI:http://dx.doi.org/10.15353/vsnl.v3i1.171

[30] 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.

[31] Liangchen Song, Gang Yu, Junsong Yuan, and Zicheng Liu. 2021. Human pose estimation and its application to action recognition: A survey. *Journal of Visual Communication and Image Representation* 76 (04 2021), 103055. DOI:http://dx.doi.org/10.1016/j.jvcir.2021.103055

[32] Fabrizio Sors, Mauro Murgia, Ilaria Santoro, and Tiziano Agostini. 2015. Audio-Based Interventions in Sport. *The Open Psychology Journal* 8 (12 2015), 212–219. DOI: http://dx.doi.org/10.2174/1874350101508010212

[33] Ashwin T S, Vijay Prakash, and Ramkumar Rajendran. 2023. A Systematic Review of Intelligent Tutoring Systems based on Gross Body Movement Detected using Computer Vision. *Computers and Education: Artificial Intelligence* 4 (01 2023), 100125. DOI:http://dx.doi.org/10.1016/j.caeai.2023.100125

[34] Emanuel Todorov, Reza Shadmehr, and Emilio Bizzi. 1997. Augmented Feedback Presented in a Virtual Environment Accelerates Learning of a Difficult Motor Task. *Journal of motor behavior* 29 (07 1997), 147–158. DOI:http://dx.doi.org/10.1080/00222899709600829

[35] 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.

[36] 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 (11 2022), 100139. DOI:http://dx.doi.org/10.1016/j.jjimei.2022.100139

[37] Yang Wang, Hongyuan Wang, and Zihao Xin. 2022. Efficient Detection Model of Steel Strip Surface Defects Based on YOLO-V7. *IEEE Access* PP (01 2022), 1–1. DOI: http://dx.doi.org/10.1109/ACCESS.2022.3230894

[38] Mingqiu Xie. 2020. Design of a physical education training system based on an intelligent vision. *Computer Applications in Engineering Education* 29 (06 2020). DOI: http://dx.doi.org/10.1002/cae.22259

[39] Qiang Yang, Yu Zhang, Wenyuan Dai, and Sinno Jialin Pan. 2020. *Transfer learning*. Cambridge University Press.

Appendices

Appendix A
Participant Handouts for User Study

This chapter outlines the various kinds of handouts provided to participants throughout the user study. The handouts given to Group B were mostly similar to those given to Group A, except for the sections related to AI feedback, which were omitted. Below is a comprehensive overview of the distributed materials:

1. **Participant Registration and Consent Form**: This form was given to each participant, providing them with information about video recording and data usage. This ensured that participants gave their informed consent.

2. **Participant Instructions**: This handout provided a detailed overview of the sequential processes to be followed during the user study, thereby familiarizing the participants with the study's workflow.

3. **Preliminary Questionnaire**: This form was created to collect data on the participants' exercise background. It played a crucial role in setting a starting point and helped in classifying the participants into appropriate groups.

4. **User Survey (*Likert Scale*)**: This survey was created to obtain precise feedback on several parts of the user study using a Likert scale, allowing for measurable insights.

5. **User Survey (*Open-Ended*)**: This survey consisted of questions that were not limited in their response options, giving participants the opportunity to freely express their feedback without any restrictions. The purpose of included this survey was to encompass a wider range of participant experiences and perceptions.

Each of these handouts had a key part in the careful collection of data and feedback, which was essential for a thorough analysis of the user study.

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Participant Registration and Consent

Project Title: Enhanced Self-Learning of Fitness Routines through Real-Time Pose Monitoring in Augmented Intelligence Tutoring Systems

Researcher: Hitesh Kotte
Contact Information: hiteshkotte@gmail.com

Introduction

Welcome to our study! We appreciate your interest in participating. Before you proceed, please read the following information carefully. By completing and signing this form, you are giving your informed consent to participate in the study.

Study Overview

- You will be part of a research study that assesses the posture quality of gym exercises using AI.
- The study involves performing 6-7 exercises.

Participant Information

- Full Name: ____________________________

Confidentiality and Data Usage

- Your participation is completely voluntary.
- Your personal information provided on this form will be kept confidential and will only be used for study-related purposes.
- Any data collected, including video recordings, will be used solely for research and analysis purposes and will not be shared with any unauthorized individuals or organizations.
- Data will be anonymized, and any identifying information will be removed.

Video Recording

- The study may involve video recording of your participation.

- The video recordings will be used for research and analysis.
- You may be asked to repeat exercises for data collection purposes.

Rights of the Participant

- You have the right to withdraw from the study at any time without any penalty or consequence.
- You may refuse to be recorded on camera or decline to repeat exercises.
- You can ask questions or seek clarification at any point during the study.
- You can ask for deletion of personal data or recordings.

Consent

I have read and understood the information provided in this form. I agree to participate in the study and consent to the use of my data and potential video recordings as described.

Participant's Signature: ______________________ Date: ______________________

Researcher's Statement

I have explained the study to the participant and have answered any questions they may have had. I confirm that the participant has voluntarily consented to participate.

Researcher's Signature: ______________________ Date: ______________________

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Participant Instructions

Project Title: Enhanced Self-Learning of Fitness Routines through Real-Time Pose Monitoring in Augmented Intelligence Tutoring Systems

Researcher: Hitesh Kotte
Contact Information: hiteshkotte@gmail.com

Greetings!

Welcome to our study! We appreciate your participation. Please follow these steps as you enter the room:

Note: If you have any questions or need assistance, feel free to ask the researcher. Your comfort and understanding are important to us, and we are here to help throughout the study.

Step 1: Participant Registration and Consent

As a participant, you will be provided with information regarding the study's objectives, procedures, and potential risks. Your informed consent will be sought, and it will be made clear that you have the right to withdraw from the study at any point, with an emphasis on this right.

Step 2: Preliminary Questionnaire

You will fill out a preliminary questionnaire to provide information about your demographics, prior experience with gym activities, and any pre-existing medical conditions that could affect your performance.

Step 3: Task Set

You will perform a set of exercises with a brief demo by the expert. While performing the exercises, You will receive real-time feedback and guidance from the Interactive model to help you improve your posture and form during the exercise. The exercise you will be performing are:

1. Bicep Curl
2. Squats
3. Lateral Shoulder Raise
4. Bent Over Row
5. Lunges
6. Shoulder Press

Note: Please refer to the researcher for the specific exercise.

Step 4: Transition and Rest Period

A brief transition and rest period will be provided between Task Set 1 and Task Set 2 to minimize potential fatigue effects.

Step 5: User Survey

A user survey will be conducted to gather your qualitative insights into experiences, challenges faced, and perceptions of the task set.

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

Preliminary Questionnaire

Project Title: Enhanced Self-Learning of Fitness Routines through Real-Time Pose Monitoring in Augmented Intelligence Tutoring Systems

Researcher: Hitesh Kotte
Contact Information: hiteshkotte@gmail.com

Participant Information

Please provide the following information:

1. **Name:** ______________________________
2. **Age:** __________
3. **Occupation:** ______________________________
4. **Gender:**

☐ Male

☐ Female

☐ Other: _______________

5. **Have you been to a gym before?**

☐ Yes

☐ No

6. **If yes, how often do you visit the gym?**

☐ 0-2 days per week

☐ 3-5 days per week

☐ Daily

7. **Do you have any pre-existing medical conditions that may affect your physical activity or performance?**

☐ Yes (please specify): ______________________________

☐ No

8. **What is your fitness level?**

☐ Sedentary (little to no physical activity)

☐ Moderately active (moderate exercise or sports 3-5 days a week)

☐ Extremely active (very hard exercise, physical job, or training twice a day)

9. **Please list fitness routines or exercises you are familiar with assigned by the researcher:**

10. **Are you currently taking any medication or supplements that may affect your physical activity?**

☐ Yes (please specify): ________________________________

☐ No

11. **Any other information you would like to share about your health, fitness, or exercise habits:**

Deutsches Forschungszentrum für Künstliche Intelligenz GmbH

User Survey

Project Title: Enhanced Self-Learning of Fitness Routines through Real-Time Pose Monitoring in Augmented Intelligence Tutoring Systems

Researcher: Hitesh Kotte
Contact Information: hiteshkotte@gmail.com

How much do you agree with these criteria after performing the exercises?
Please rank between (*Strongly Disagree to Strongly Agree*).

	Question	Strongly Disagree	Disagree	Neutral	Agree	Strongly Agree
1	The weights provided for the exercises were appropriate for my body weight.	○	○	○	○	○
2	The AI system will help me in maintaining better posture and form during exercises.	○	○	○	○	○
3	The AI system's real-time feedback made the exercises more engaging.	○	○	○	○	○
4	I am likely to recommend this AI-based exercise system to others.	○	○	○	○	○
5	I had no difficulty following the AI system's instructions or feedback.	○	○	○	○	○
6	Overall, how satisfied are you with the system (On scale 1-5)?	○	○	○	○	○

Deutsches
Forschungszentrum
für Künstliche
Intelligenz GmbH

User Survey

Project Title: Enhanced Self-Learning of Fitness Routines through Real-Time Pose Monitoring in Augmented Intelligence Tutoring Systems

Researcher: Hitesh Kotte
Contact Information: hiteshkotte@gmail.com

1. How would you describe your overall experience using the AI in the gym setting, considering the system you just saw? What are your initial thoughts or expectations?

2. Can you share any specific instances where you think the system you just saw might impact your gym routine? Which elements of the system do you believe you would use and/or focus on during your performance?

3. Have you noticed any potential changes in your workout efficiency or effectiveness that could occur by using the system you just saw? What improvements or challenges do you anticipate?

4. Do you have additional feedback, comments, or remarks based on the system you just witnessed? What are your initial impressions or concerns?

.

Appendix B
Statement of the Ethical Review Board

Prof. Antonio Krüger · Univ. des Saarlandes · Campus D3 2 · D-66123 Saarbrücken

UNIVERSITÄT DES SAARLANDES

Prof. Dr. Antonio Krüger
Chair
Ethical Review Board
Department of Computer Sciences

Universität des Saarlandes
Saarland Informatics Campus D3 2
D-66123 Saarbrücken
Deutschland

Fon: +49 (0) 681 85775 5006
Fax: +49 (0) 681 85775 5007
krueger@cs.uni-sb.de

Saarbrücken, 18.12.2023

Statement of the Ethical Review Board (ERB)
in response to your application (No. 23-12-08)

Dear Hitesh Kotte,

the ERB has reviewed your research project "Fitsight". According to its regulations, approved by the Department of Computer Science of Saarland University on October 26, 2016, the ERB has come to the following conclusion:

"There are no ethical concerns against the implementation of the research project"

If you have any questions, please let me know. We wish you all the best in your future research endeavours.

Best regards, on behalf of the ERB

Prof. Dr. Antonio Krüger
Chair of the Ethical Review Board
of the Department of Computer Sciences at Saarland University

Figure B.1: Approvement of the Ethical Review Board in response to the reasearch project