AI GOLF COACH: A POSE-BASED GOLF COACHING SYSTEM USING ARTIFICIAL INTELLIGENCE AND COMPUTER VISION

Andy Yu[1], Yu Sun[2]

[1]Los Osos High School, 6001 Milliken Ave, Rancho Cucamonga, CA 91737
[2]California State Polytechnic University, Pomona, CA, 91768,
Irvine, CA 92620

ABSTRACT

In the sport of golf, several main mechanics in each successful golf swing bring consistency [1]. In addition, there are specific nuances in the golf swing that may be difficult for recreational players to analyze and detect. The main question is how an app can potentially detect these main techniques and compare them when inputting two different golf swings. Guided by Mediapipe and cv2 technologies, we have built software to analyze two golf swings simultaneously, and to compare them with one another, all into a smartphone application [2][3]. This paper will discuss key technologies that were utilized to create the program, the various challenges that were faced in the process of creating the software, along with our methodologies [4]. The end result was an application that would detect various angles created based on certain golf positions, and then utilize them to compare with another video, whether that be a professional player or a friend's golf swing. Using a Python backend and a Flutter-based frontend, the program can be utilized through a mobile device, and thus makes the program accessible to the public [5]. With the publishing of this application, users can start to deconstruct their own golf mechanics without the complications and the high fee that is generally charged for one-on-one coaching.

KEYWORDS

Golf, AI, Golf coach, Visual feedback

1. INTRODUCTION

Golf is an extremely precise sport with many factors that need to be evaluated in order to find success in the sport. One of such is the golf swing and specifically the technique and mechanics that come with the swing. One major obstacle to entry and success in the game of golf is sub-par coaching, as the cost and time barrier is profound. However, our app aids in this issue as the user, after implementing their own swing and a professional golfer's swing of their choice into our app can see the exact frames in which both swings are different, and thus start to fix their swing without needing a professional coach. The popularity of the app can easily be realized as the software is a free app, making it accessible to anyone with a mobile device. In addition, the app acts as a portable coach, as you can take a video of your swing and always compare it with another golfer, whenever the user wants to. This benefits both casual and competitive golfers as it can be used as a fun way to gauge different swings with each other, or as a way to check up on the specific mechanics of the golf swing. There are no negative consequences to the existence of this app, as it is specialized to address a niche issue. With the greater promotion of this app, it will not only bring more people to start understanding their technique, but it will likely bring

David C. Wyld et al. (Eds): ICAITA, CDKP, CBW, SOFT, CMC, Signal, MLT-2023
pp. 83-98, 2023. CS & IT - CSCP 2023 DOI: 10.5121/csit.2023.130607

more people into the sport of golf, as self-coaching is now very accessible with the help of artificial technologies.

Some related methods and tools on this topic are similar to artificial intelligence software which measures various angles in the golf swing [6]. Renowned world-class instructors have started to use artificial intelligence to measure those measurements to the optimal positions. In addition, institutions such as GolfTec have also implemented artificial intelligence into their technological lessons, in order to attempt to make coaching easier to visualize and implement. Areas such as GolfTec need specific coaches to interpret the information that is brought upon through the AI software and thus create more barriers to the idea of AI coaching. One popular AI technology company specifically designed for golfers is SportsBox AI, which they created using 3D motion analysis [7]. The main flaw that comes with AI technologies such as SportsBox AI is that instructors must be taught the optimal positions in such technologies, and thus the barrier in terms of experience and cost for everyday and competitive golfers will skyrocket. Another flaw that comes with these AI technologies that are now on the golfing market is the idea that they only analyze the player's angles, and cannot compare from one golfer to another. As a result, not only that these types of technology is not available to the everyday or competitive golfer and are especially rare to come by, non-specialized golfers cannot interpret the information that is portrayed with these types of apps. Combined with the extremely high cost, our application breaks through these two obstacles as our free application can now be utilized by everyday and competitive golfers in order to consciously improve their game with no cost. In addition, everyday golfers, using our 3-frame comparison system, can quickly see the differences between their swings versus the professionals, and thus can start to gain an understanding of the golf swing and continue to self-correct their swings.

Similar to some of the other AI technologies geared toward golf coaching on the market right now, our app also utilizes AI, specifically Mediapipe, to interpret videos and create angles based on various body parts [8]. Some features that are in our app are that it allows two videos to be compared side by side after our analysis of the videos, and it also shows the two videos implemented and replayed, with the Mediapipe angles on the video as it plays, simultaneously changing throughout the video to fit the frames. In addition, we also include a history tab in order to rediscover older analyses for storage purposes and to also reflect on the various changes in mechanics that have progressed over the time of using this app. In addition, by utilizing Firebase for storage in the backend, images that are deduced after the algorithm is always saved in that storage. Dissimilar to other AI technologies, the application is completely free to download, once it reaches app stores, as we believe that the lack of solid coaching should not be a barrier to entering the sport. In addition, our ability to analyze two videos simultaneously, and then put them together, side by side, within the same frames give the users a unique ability to check their swings at specific checkpoints, something that is vital in the golf swing.

We will prove the results of our application by evaluating the results of the application, and seeing if the side-by-side comparisons are accurate and thus helpful. By comparing the two videos, it is clear to see whether or not the algorithm has picked up the correct two pairs of photos to contrast, and as long as those two images are in the same position in that frame, it will be accurate. Since the positions are generally accurate, the user from there can analyze the two points themselves, start interpreting the visual illustrations of their swing mechanics and start to understand where their flaws are in the golf swing. In this topic specifically, I simply used the application on a variety of golfers in order to test the program's effectiveness in detecting similarities between two different swings, one usually being an amateur and the other being a professional golf swing. In addition, the history tab on our app can enable users to look back on their previous runs with the application, and notice the differences in their swing. The general accuracy of the program can be beneficial, and the results thus can be implemented in either an

everyday golfer or a serious competitive golfer's game and routine. In conclusion, the results stemming from our app are generally accurate, and as a result can add to the sophistication and depth of the program, as it uniquely compares two different videos with each other to find similarities and differences in technique.

The rest of the research paper is organized into various sections: Section 2 will expand on the various challenges faced during the development of the software; Section 3 will explain the methodology and process behind our solution and the application; Section 4 will convey the details on experiments, and Section 5 will be a short but vital literature review on research papers previously published in this area of topic and discipline. Section 6 will be the conclusion of the paper, and also illustrate limitations along with areas to improve, with ideas on how the future of the work will look on the application.

2. CHALLENGES

In order to build the project, a few challenges have been identified as follows.

2.1. Realizing Mediapipe with cv2 and Implemented Videos

The first major challenge we faced when building this application was hiccups in the Mediapipe process, specifically how to implement the software into our videos to be displayed. This was especially important, as cv2 and Mediapipe are the two aspects of the backbone of the software, as it is what allows the videos to be interpreted in the first place. Without the two functioning correctly, there was no way to read the videos to even give feedback to the user. As a result, after some research on the Internet and Mediapipe's own website, we were eventually able to diagnose the issue and reset our environment to sufficiently provide an area that could allow the cv2 and Mediapipe to work together seamlessly. Finally, we solved the first major challenge, as the videos that were originally implemented from our computers could be read, along with Mediapipe angles to portray alongside.

2.2. Building the K-Means Clustering

The next major obstacle was challenges regarding the building of k-means clustering, the main The next major obstacle regards the building of k-means clustering, the main way our software detects the similarities in positions between the two videos [9]. Initially, the whole program would fail if we tried to run the application with k-means clustering, and after many different bug fixes, the clustering was still off and did not bond with the correct pairs of pictures. Oftentimes, the pictures would get mixed up, not match at all, or even pair with frames from the same video. Thus, there had to be adjustments in the code to specifically work well within my system, and eventually was able to fix the challenge. This was especially important, as it enabled us to use both cv2 and Mediapipe, in conjunction with k-means in order to truly compare the two videos. As a result, this was a vital component of the program that was now fixed.

2.3. Flutter

Finally, the last major challenge was the usage of Flutter and implementing it as a frontend to connect to our backend python server. It was extremely important for Flutter to function correctly, as it was the parts that the user would interact with, and thus would resonate with the most. At the beginning of the Flutter implementation, many areas did not connect properly, and some graphics overrode the pixel-by-pixel dimensions of the various emulators we were testing out. In addition, some buttons were not able to be seen as a result of these boundaries being

crossed in terms of dimensions. As a result, creative solutions had to be made up, such as minimizing the size of the button in order to fit on a smaller screen or adding png images to portray the logo of the application. The small details were especially important, as it gives the user a sense of connectedness in our application and in order to make the app as easy to use as possible, and utilize abstraction to simplify things on the user interface.

3. SOLUTION

AI Golf Coach runs off of a variety of different components in order to create the most accurate outputs for the application. First off, users will input two videos through their connected Google Drive to upload their corresponding "student" and "coach" videos to be analyzed. The analyze videos button will then be clicked, in which the python-based backend will utilize OpenCV cv2, MediaPipe, and k-means clustering within an Anaconda environment in order to interpret the videos, allow the AI technologies to read the video and essentially document the positions at each given frame, and then compare and contrast the various angles formed by the body imagery provided by Mediapipe to eventually cluster angles together to form accurate outputs of images [10]. AI Golf Coach is built off of a Flask and AWS ec2 instance server, which allows for the potential scalability of the application moving in the future and sets up the foundation for machine learning to occur. The output images are then connected to Google Firebase and then output onto the user interface of the application, which is developed through Flutter. If a user ever wants to check back on a specific analysis, they can also utilize the History tab in the application in order to retrace their analysis, which is all backed up through Google Firebase. The components of the application all work in coordination with each other to analyze the given videos by the user and portray the most accurate analyses of such videos and offer accessibility to the user. The major processes can be seen in Figure 1, which outlines the methodology of AI Golf Coach and how it operates.

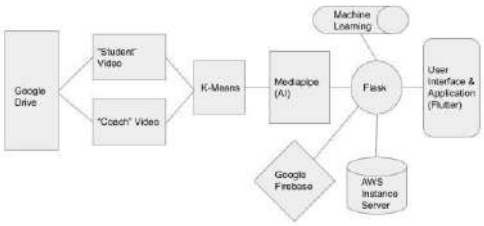

Figure 1. Overview of AI Golf Coach's Components

First, the user starts off their journey by implementing videos from their linked Google Drive to be inputted into the application (Figure 2). Through Google Drive, the user will then implement two videos, a coach and a student, into the app for them to be eventually analyzed (Figure 3). Next, the two videos go through the k-means clustering portion, which in conjunction with Mediapipe's AI technologies that provide various angles on the person, can compare the two videos in order to find similarities (Figure 4). In addition, the Kmeans portion of the program utilizes other functions such as get_nearest_neighbor to find such clusters. By interpreting these similarities, the output images can be connected from the server and returned to the user interface, controlled by Flutter. Speaking of the server, the server runs off of an AWS ec2 instance server in order to operate the app (Figure 5). This server is then connected to the Flutter portion of the app, which leads to the outputted images after the "Analyze Videos" button is clicked, leading to the loading screen (Figure 6). The images generated through the k-means clustering are then uploaded to the user interface as shown (Figures 7, 8, 9, 10). The images

generated are then uploaded to Google Firebase, a service for hosting databases. Through Firebase, a "History" tab in our application can be utilized, as Firebase generates public URLs for each generated image in order to potentially be recalled (Figure 11). These components were connected through the Python-based backend, in order to create a fully functioning application for mobile devices. In the process of building the application, Anaconda for Python was utilized to simplify data management, especially because of the utilization of machine learning. The machine learning aspect came in conjunction with the k-means clustering, as the algorithm developed in the application's backend could be used in order to find patterns in videos. These videos implemented by the user are interpreted by using OpenCV, a Python open-source library that offers various methods to read videos, which offered functionality that played a crucial role in the building of AI Golf Coach (Figure 12). In order to test out the functionality, Android Studio was utilized to simulate a mobile device screen on my computer, as the emulator ended up being connected to the whole project itself. Thus, by utilizing Android Studio's emulators, I could test out various Flutter widgets or changes to my Python backend to see the immediate feedback and changes such as the splash screen (Figure 13). In summary, the application utilizes a Python backend and a Flutter frontend utilizing services such as Google Firebase, Google Drive, AWS Instance Servers, libraries such as OpenCV and Mediapipe, algorithms such as k-means, and other additions such as Anaconda.

Figure 2. Google Drive Access

Figure 3. Analyze Videos Page

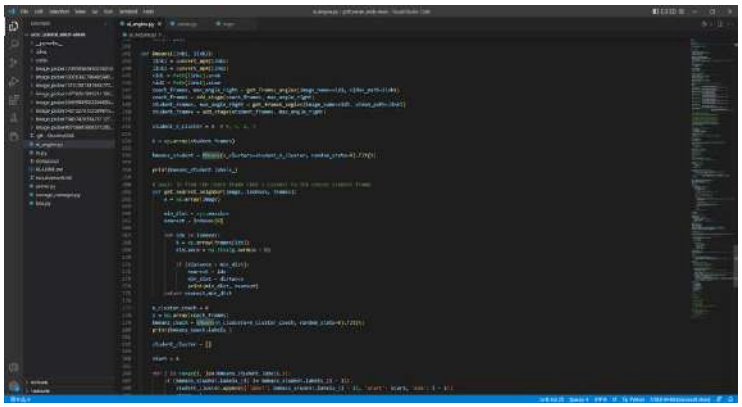

Figure 4. Screenshot of Code #1

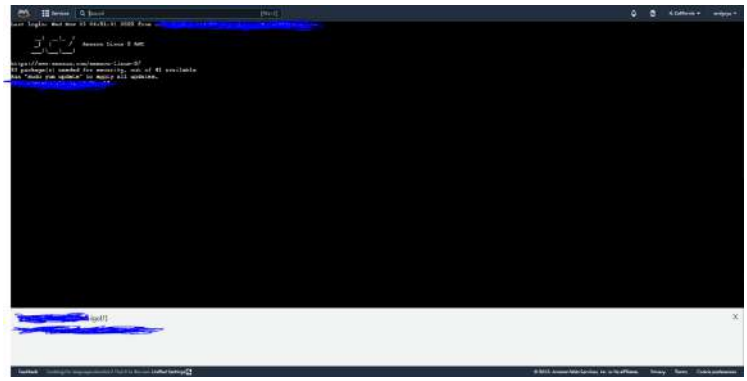

Figure 5. AWS Instance Server

Figure 6. Analyzing Videos Processing

Figure 7. Analyzed Posture #1

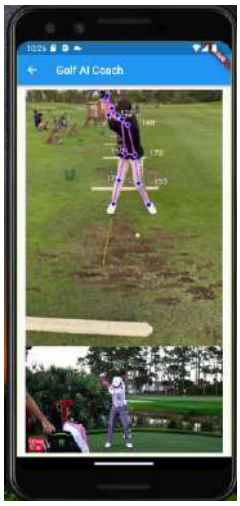

Figure 8. Analyzed Posture #2

Figure 9. Analyzed Posture #3

Figure 10. Analyzed Posture #4

Figure 11. History Page

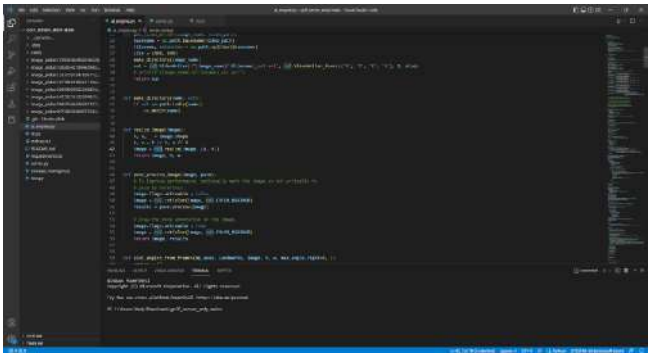

Figure 12. Screenshot of Code #2

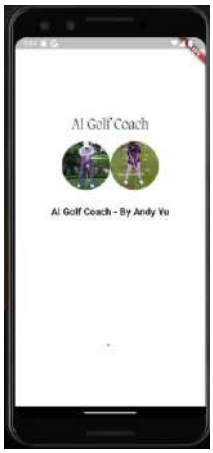

Figure 13. Splash Screen

4. EXPERIMENT

4.1. Experiment 1

Using the AI Golf Coach app through an Android emulator, we can imitate how the app will operate by inputting two videos from the front camera: Tiger Woods and myself, Andy. We will then see the outputs of the two videos, and analyze the frames. The success of the experiment will be based on the accuracy of the compared photos, which will output multiple images side-by-side. The angles created by the various body parts should all be similar, and thus give the user a good understanding of the differences between themselves and the "coach" video.

My app AI Golf Coach solves the issues of pricey, high-level coaching, the large time commitment of traditional coaching, and the huge discipline needed to continue working on specific techniques. With my application, the various issues aforementioned can be bypassed, as my application gives users the ability to capture various frames in order to dissect their swing easily and improve their technique.

I have 2 videos, my own video, and Tiger Woods' swing video as the sample. And 4 pairs of images, one from each video, at a certain position in the golf swing in order to compare the various positions in the golf swing as feedback.

The user first implements two videos, a "student" and a "coach" video through Google Drive to be compared against each other. After running the initial analysis, four pairs of images are portrayed as a result of the algorithm. Each pair of photos indicate a similar position in the golf swing, as a result of the cv2, Mediapipe, and k-means clustering working simultaneously. The user then can immediately detect the differences in their swing versus a professional golfer's swing, and thus change their mechanics as needed. In this specific trial, the positions were at setup, the top of the backswing, right after setup/right before impact, and the finish. Thus, the experiment was a success, as the outputs were similar and gave the user feedback compared to their "coach" video. Therefore, the implementation of AI technologies to assist with golf coaching is very feasible and can continue to be explored in the future.

Figure 14. AI analysis result 1

Figure 15. AI analysis result 2

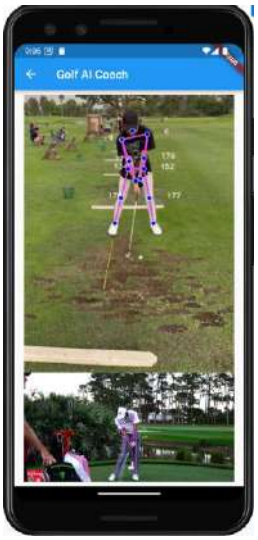

Figure 16. AI analysis result 3

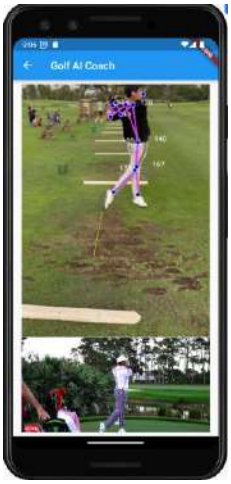

Figure 17. AI analysis result 4

4.2. Experiment 2

Using the AI Golf Coach app through an Android emulator, we can imitate how the app will operate by inputting two videos from the front camera: my friend Samson and myself, Andy. We will then see the outputs of the two videos, and analyze the frames. The success of the experiment will be based on the accuracy of the compared photos, which will output multiple images side-by-side. The angles created by the various body parts should all be similar, and thus give the user a good understanding of the differences between themselves and the "coach" video.

My app AI Golf Coach solves the issues of pricey, high-level coaching, the large time commitment of traditional coaching, and the huge discipline needed to continue working on specific techniques. With my application, the various issues aforementioned can be bypassed, as my application gives users the ability to capture various frames in order to dissect their swing easily and improve their technique.

I have 2 videos, my own video and Samson's swing video as the sample. And 4 pairs of images, one from each video, at a certain position in the golf swing in order to compare the various positions in the golf swing as feedback.

The user first implements two videos, a "student" and a "coach" video through Google Drive in order to be compared against each other. After running the initial analysis, four pairs of images are portrayed as a result of the algorithm. Each pair of photos indicate a similar position in the golf swing, as a result of the cv2, MediaPipe, and k-means clustering working simultaneously. The user then can immediately detect the differences in their swing versus a professional golfer's swing, and thus change their mechanics as needed. In this specific trial, the positions were at setup, the middle of the backswing, the top of the backswing, and the finish. Thus, the experiment was a success, as the outputs were similar and gave the user feedback compared to their "coach" video. Therefore, the implementation of AI technologies to assist with golf coaching is very feasible and can continue to be explored in the future [14].

Figure 18. AI analysis result 5

Figure 19. AI analysis result 6

Figure 20. AI analysis result 7

Figure 21. AI analysis result 8

Figure 22. AI analysis result 9

5. RELATED WORK

This research article written by Peng Ding in the Journal of Physics discussed the implementation of AI in sports and the specific obstacles that would need to be overcome for proper application [11]. The author uses value analysis in order to deconstruct AI and its current position in the world of sports, and provides 5 strategies in order to improve AI for its efficient use in sports. My work and this research article's work are similar as we both utilize AI and sports. I start the conversation that the author is wary about, which is the application of AI in sports. The strengths of their research paper are that it is backed by other research articles (as shown in the references page) and articulates an understanding of the various nuances that are needed for proper application in sports, while my strength in my research paper is that my app is a clear indication of the potential of AI in sports. Their research paper may be weak in specific analysis, as most ideas are portrayed without experiments, however, my experiment may be weak as I have not developed a large sample size for my software.

The second research paper was very recently published and discusses an extremely similar topic of mine, which is AI implementation in golf techniques [12]. The researcher's main points were that beginners often have issues in pinpointing the specific positions in which to fix, and how that can be compared to professionals using data graphs and 3D data models. Compared to my work, their research is much more extensive and largely implemented among amateurs in order to be compared to professionals. In addition, they utilize mathematical concepts, diagrams, and image comparisons in order to thoroughly address their process. The strengths of their research are that it is extremely in-depth, and combines a variety of factors such as the aforementioned 3D data models, video analysis, and data tables from specific information gathered, and thus is much more scientific, however, my research is strong in which my software can capture most of the essence in a single application, and thus is much more applicable to most people. The main weakness of "AI Golf: Golf Swing Analysis Tool for Self-Training" is that it is extremely nuanced and hard to interpret for the general audience attempting to understand their program, while my main weakness is the lack of scientific documentation and the lack of sample size.

The third research paper I analyzed was titled "Golf Swing Segmentation from a Single IMU Using Machine Learning" which utilized principles of machine learning in order to divide the golf swing into specific phases to eventually assist in the development of other swing analysis software [13]. Their main contribution was applying and proposing two different methods with inertial measurement units (IMUs) to divide such phases. Compared to my work, their research was similar to the second research paper, as both were scientific and utilized many different visual aspects in order to attempt to simplify complex algorithms. In addition, the two proposed methods of bidirectional long short-term memory-based and convolutional neural network-based methods were tested with multiple professional golfers for feedback. The strengths of their paper were mainly the level of scientific accuracy and complexity to the discipline, as their discovery was relatively new, while the strengths of my research are the ability for much complexity of comparing golf positions abstracted into a mobile app. The weakness of the research may be the complexity of their methods, as they might not be practical to be implemented widely and with a relatively small sample size, while the weakness of my research is the relative lack of scientific discussion and small sample size with various videos with various camera angles.

6. CONCLUSIONS

The future of AI in the implementations of golf lies in utilizing it for various coaching functions, such as the one we are currently proposing with this application, in which two videos are compared against each other to find similarities and differences. By using the ability of artificial

intelligence, we can provide flexible feedback to all users of all skill levels, and can be implemented with data clustering with other sports, not just golf. The results see the potential, effectiveness, and accuracy that these pictures and comparisons may present. Through the utilization of Python as a backend, Flutter-based Dart as a frontend, Mediapipe for AI technologies, cv2 for reading videos, K-means for clustering and machine learning, Firebase for storage, and AWS for running the server, AI Golf Coach serves as a large step in the future of coaching and sport improvement for anyone at any skill level [15]. The potential of the application to further give explicit feedback and scale the app even larger proves the practicality and demand for this new market of sorts. By running the application with various videos from the front-on-camera view, the flexibility of the app is demonstrated with the continued accurate portrayals of various positions in the golf swing. In conclusion, the future of coaching has a bright potential in the world of artificial intelligence and the implementation of artificial intelligence as an efficient way to compare multiple videos for instructional purposes.

Some current limitations in my application are the accuracy of the comparison of frames, as my application consistently only gets 1-2 out of the 3 displayed pairs of photos matched up. In addition, I would like to make the application's visuals more appealing and straightforward, and also eventually get our algorithm to line up in the P classification system, in which there are ten. In addition, the algorithm may be better optimized in order to help the app run slightly faster, and adding a slider to indicate when the process is finished may assist the user in their usage. Looking into the long term, it is possible that based on these AI deductions, we can begin to explicitly explain to the golfer what to focus on.

I plan to solve these limitations in the future by evaluating the algorithm used in order to deduce the comparisons between the two swings, and then attempting to improve such to garner better results. In addition, adding more visuals for the user to understand at which point the app is at will better appeal to the user's interaction with the app.

REFERENCES

[1] Langdown, Ben L., Matt Bridge, and Francois-Xavier Li. "Movement variability in the golf swing." Sports Biomechanics 11.2 (2012): 273-287.

[2] Lugaresi, Camillo, et al. "Mediapipe: A framework for perceiving and processing reality." Third Workshop on Computer Vision for AR/VR at IEEE Computer Vision and Pattern Recognition (CVPR). Vol. 2019. 2019.

[3] Nesbit, Steven M., and Monika Serrano. "Work and power analysis of the golf swing." Journal of sports science & medicine 4.4 (2005): 520.

[4] Avison, David E., and Guy Fitzgerald. "Where now for development methodologies?." Communications of the ACM 46.1 (2003): 78-82.

[5] Punja, Shafik G., and Richard P. Mislan. "Mobile device analysis." Small scale digital device forensics journal 2.1 (2008): 1-16.

[6] Harman, Mark. "The role of artificial intelligence in software engineering." 2012 First International Workshop on Realizing AI Synergies in Software Engineering (RAISE). IEEE, 2012.

[7] Ford, Kevin R., Gregory D. Myer, and Timothy E. Hewett. "Reliability of landing 3D motion analysis: implications for longitudinal analyses." Medicine and science in sports and exercise 39.11 (2007): 2021-2028.

[8] Zhang, Fan, et al. "Mediapipe hands: On-device real-time hand tracking." arXiv preprint arXiv:2006.10214 (2020).

[9] Hamerly, Greg, and Charles Elkan. "Learning the k in k-means." Advances in neural information processing systems 16 (2003).

[10] Krishna, K., and M. Narasimha Murty. "Genetic K-means algorithm." IEEE Transactions on Systems, Man, and Cybernetics, Part B (Cybernetics) 29.3 (1999): 433-439.

[11] Ding, Peng. "Analysis of artificial intelligence (AI) application in sports." Journal of Physics: Conference Series. Vol. 1302. No. 3. IOP Publishing, 2019.

[12] Liao, Chen-Chieh, Dong-Hyun Hwang, and Hideki Koike. "AI Golf: Golf Swing Analysis Tool for Self-Training." IEEE Access 10 (2022): 106286-106295.
[13] Kim, Myeongsub, and Sukyung Park. "Golf swing segmentation from a single IMU using machine learning." Sensors 20.16 (2020): 4466.
[14] Smith, Aimée, et al. "Professional golf coaches' perceptions of the key technical parameters in the golf swing." Procedia Engineering 34 (2012): 224-229.
[15] Jordan, Michael I., and Tom M. Mitchell. "Machine learning: Trends, perspectives, and prospects." Science 349.6245 (2015): 255-260.