A Multimodal Intelligent Assessment Method and Solution for Enhancing Billiard Skills

Yachen Tian

High School Affiliated to Renmin University of China, Beijing, China

Lloydtyc@outlook.com

Keywords: Multimodal Data Fusion, Transformer Model, Billiard Skill Assessment, Sports Strategy Optimization, Tactical Decision-Making

Abstract: This study presents an intelligent assessment method based on multimodal data to enhance billiard skills. By collecting data from multiple sources, including video recordings, sensors, and physical parameters, the system comprehensively evaluates players' performance in key areas such as ball control, shot accuracy, positioning, and strategic thinking. A deep learning model with a Transformer architecture was employed to process and analyze the multimodal data, identifying strengths and weaknesses for each player. The system then generated personalized training programs tailored to the skill level of individual players. For beginners, the recommendations focused on improving basic posture and shot accuracy, while more experienced players received guidance on advanced tactical training. Experimental results demonstrated that the system significantly improved most participants' billiard skills in a short period. Additionally, the system held broader social value, especially in public welfare programs aimed at youth in underprivileged areas, helping them develop fine motor skills and strategic thinking through billiard education. This research introduces an innovative approach to evaluating and improving billiard skills by integrating intelligent technology and multimodal data analysis. Furthermore, it demonstrates the wide-ranging application of such technology in sports training, offering a novel method for accelerating skill development in both novice and advanced players, and highlighting the potential for expanding the system to other sports and educational contexts.

1. Introduction

Billiards, a classic indoor sport that combines technology and strategy, evolved from croquet games in Northern Europe in the 15th century. After hundreds of years of development, it has become a widely popular competitive sport around the world. Although its rules are relatively simple, the superb skills and complex strategic thinking in billiards require athletes to have many abilities, such as precise shot control, clever tactical arrangements, and deep strategic understanding. The current training methods of billiards are still highly dependent on the guidance of experienced coaches or self-feedback from athletes. However, this training method based on subjective experience often makes it difficult to fully and accurately identify the shortcomings of athletes' skills and potential room for improvement. Therefore, how to use advanced intelligent technology to provide billiards players with personalized skill improvement solutions has become an important direction worthy of exploration.

With the rapid development of artificial intelligence (AI), multimodal data analysis, and deep learning technology, technological innovation in the field of sports is gradually changing the traditional training and evaluation model. Especially in sports evaluation, the fusion technology based on multimodal data can achieve a comprehensive evaluation of athletes' movements, skills, and strategies by integrating multi-source information such as video, sensor data, and motion trajectories. These technologies not only greatly improve the accuracy of training, but also automatically generate highly targeted personalized training plans through big data analysis [1]. At present, AI technology has demonstrated excellent results in many sports, such as tennis and football, but its application in the field of billiards is still relatively limited, lacking systematic

  DOI: 10.25236/iwmecs.2024.004

research and technical practice.

In order to solve the technical evaluation and training problems in billiards, this study proposes a multimodal intelligent evaluation system with strong technical innovation, combining deep learning and multimodal data fusion technology to comprehensively analyze the hitting accuracy, cue control, positioning skills and tactical thinking of billiard players. Unlike traditional single data source evaluation, this system generates a complete sports performance map through multi-dimensional data collection, including athlete action videos captured by high-definition cameras, cue movement trajectories and hitting angles recorded by sensors, and physical movement data of the ball during billiards. Based on these multimodal data, the system uses advanced Transformer models for deep learning analysis to identify the technical patterns, advantages and disadvantages of athletes, and customize personalized training plans for them. This innovative evaluation method not only improves the efficiency of athlete training, but also provides tailored feedback and improvement suggestions for athletes of different levels [2].

Technically, the core innovation of this study lies in the combination of multimodal data fusion and deep learning algorithms. First, the synchronous collection of video, sensor and environmental data ensures the comprehensiveness and diversity of the data. Second, the Transformer model processes multimodal data through an efficient self-attention mechanism, which can capture the complex correlations between different data sources and generate more accurate and in-depth evaluation results for athletes. In addition, the system also has the function of automatically generating personalized training plans, and can recommend corresponding technical exercises and tactical training based on the unique performance of each athlete.

Experimental results show that athletes using this system have significantly improved in terms of hitting accuracy, cue control, and tactical thinking. Especially in short-term training, the personalized feedback of the system helps athletes to make rapid progress in key techniques. This study provides a new technical approach for the training and evaluation of billiards skills, and shows the broad application prospects of multimodal intelligent evaluation systems in the field of sports training. In the future, the system can be further expanded to other sports and applied to a wider range of public welfare sports education, especially for the physical quality training of young people in poor areas.

This study introduces multimodal data fusion and deep learning technology to propose an innovative solution for the evaluation and improvement of billiards skills, which not only provides technical support for the intelligent training of billiards, but also lays the foundation for the application of multimodal evaluation systems in other sports in the future.

2. Related Work

In recent years, with the development of artificial intelligence (AI) and multimodal data fusion technology, intelligent evaluation systems have gradually been applied in sports training and competition. These systems can more comprehensively evaluate athletes' performance and provide personalized training suggestions by integrating multi-source data such as video, sensors, and motion trajectories. The combination of multimodal data fusion and deep learning algorithms has greatly improved the accuracy and practicality of sports evaluation systems.

In many sports, intelligent evaluation systems based on multimodal data have achieved remarkable results. The study developed a tennis intelligent evaluation system based on multi-sensor data, which can evaluate athletes' comprehensive performance and provide personalized training suggestions by collecting and analysing data such as racket swing trajectory, hitting force, position and tactical execution [1]. Similarly, the research proposed a motion capture system based on multimodal data to analyse the technical movements and strategic performance of football players in the game, and automatically generate optimized training plans through video analysis and sensor data fusion [3].

In the field of basketball, the study proposed a multimodal data fusion system based on deep learning, which can analyse the technical shortcomings of athletes and propose improvement plans by capturing their shooting movements, positions and movement trajectories. The system analyses the athletes' motion data and game scene videos by combining convolutional neural networks (CNN) and recurrent neural networks (RNN) to generate personalized tactical and technical suggestions [4]. These studies show that the fusion of multimodal data can provide more accurate athlete evaluation and greatly improve the effectiveness of personalized training programs.

Although multimodal intelligent evaluation systems have made breakthrough progress in many sports, their application in billiards is still limited. Billiards is a highly precise and strategic sport. Traditional training methods rely on the coach's experience and the athlete's subjective perception, and it is difficult to provide objective and comprehensive technical evaluation. This study introduces multimodal data sources such as video, sensor data, and physical motion parameters, and uses the Transformer deep learning model to conduct a comprehensive analysis of athletes' hitting accuracy, ball control skills, and tactical thinking, filling the gap in this field [5].

The advantages of the Transformer structure in processing multimodal data have been verified in many studies. The Transformer model first proposed by Vaswani et al. can effectively process multiple types of data and capture the associations between different data sources through the self-attention mechanism [6]. This model structure has now been widely used in the field of sports analysis, especially in the processing of complex multimodal data, showing extremely high accuracy and adaptability [1]. In addition, multimodal data fusion in sports training also involves the analysis of athletes' tactical thinking. The research proposed a tactical analysis system for football matches, which uses a deep learning model to analyse athletes' movement and decision-making patterns and provide athletes with tactical optimization suggestions [7].

3. Methods

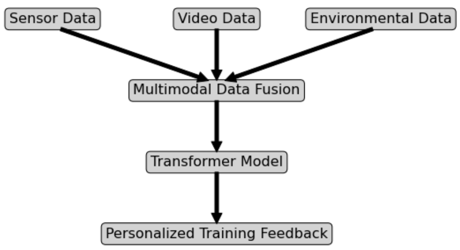

Figure 1 Diagram of the method. The flowchart includes key steps such as Input Data Sources: Sensor data, video data, and environmental data, Data Fusion and Processing: Multimodal data fusion followed by analysis using the Transformer model, Output: Personalized training feedback based on the processed data.

As Figure 1 shows, we propose an innovative multimodal intelligent evaluation system that combines multimodal data fusion and advanced deep learning algorithms to comprehensively evaluate the technical performance of billiard players and generate personalized training programs for them. Unlike existing single-mode or subjective experience-based evaluation methods, this system provides a more accurate and in-depth sports performance evaluation solution by introducing complex multimodal data integration and analysis methods.

3.1. Data Collection

In this study, we utilized multi-level and multi-dimensional data sources to ensure the full capture of athletes' technical performance. The system integrates three main data sources:

Video data: Capture the athlete's posture, stance and hitting action through a high-frame rate camera.

Sensor data: Install high-precision sensors on the cue, wrist and table to capture the force, rotation, hitting angle, movement trajectory, etc. These data provide accurate measurements of cue control and hitting mechanics.

Environmental physical data: Through sensors on the smart billiard table, the ball's trajectory after each hit, including speed, rotation and collision angle, is tracked in real time. These data provide an important basis for analyzing the accuracy of athletes' shots and ball control ability.

3.2. Multimodal Data Fusion

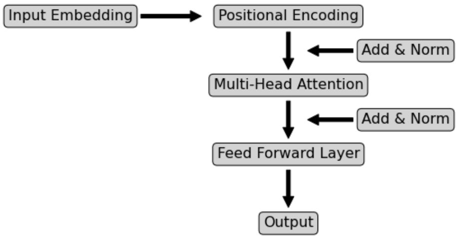

Figure 2 Diagram of the Transformer architecture.

We developed a Transformer model to achieve efficient fusion of multimodal data through the self-attention mechanism. Figure 2 shows the Transformer architecture. The system first extracted action features from video data based on convolutional neural networks (CNN) and encodes sensor data. Then, the Transformer model is used to input these feature data into a unified multimodal data fusion framework. The model captures the correlation between different data sources through the self-attention mechanism to generate a comprehensive technical performance vector. Through this innovative data fusion method, the system can more accurately identify the technical advantages and disadvantages of athletes, and the generated analysis results are more accurate and in-depth than single data source analysis. The Transformer model can capture the complex dependencies and patterns in athletes' actions in time series through the self-attention mechanism. Compared with traditional CNN and RNN, the Transformer model is not only more efficient in processing long sequence data, but also better able to handle the fusion of multimodal data.

3.3. Personalized Training Plan Generation

In this study, we used a deep learning model to comprehensively analyze the athlete's historical data and current performance to generate a customized training plan, which not only includes specific training for hitting accuracy, ball control ability and tactical thinking, but also dynamically adjusts the training plan according to the athlete's progress. For example, the system can recommend targeted exercises based on the analyzed hitting patterns, such as stance stability and club control training for beginners, or complex position hitting training for experienced athletes. By analyzing the tactical decision-making patterns of athletes in the game, the system can provide optimized strategic suggestions, such as how to predict the opponent's hitting movements and choose the best strategy in complex situations.

4. Experiments

In order to verify the effectiveness of the proposed multimodal intelligent evaluation system in billiards skill evaluation and personalized training, we designed a data collection, comparative training between the experimental group and the control group, and performance evaluation based on multi-dimensional indicators.

4.1. Data Preprocessing

In the experiment, the participants were billiards players of different skill levels, including beginners, intermediate players, and experienced players. To ensure the comprehensiveness and accuracy of the data, the system collected data from the following three channels:

Video data: The player's hitting action and stance are recorded throughout the process through a high-frame rate camera. Video data is mainly used for motion capture and technical analysis, and features are extracted including hitting posture, stance stability, and action continuity.

Sensor data: High-precision sensors are installed on the cue and the athlete's wrist to record key physical data such as the force, angle, and trajectory of each shot. Sensor data is used to analyze the athlete's cue control and hitting mechanics.

Environmental physical data: The ball trajectory after each shot is monitored in real time through sensors on the table, and data including speed, rotation, and collision angle are collected. The data is used to evaluate the players' ball control accuracy and batting effect.

After data collection, all data were preprocessed, including noise filtering, time synchronization, and multimodal feature extraction, to provide high-quality input data for subsequent deep learning model training.

4.2. Experimental and Control Group Design

The experiment adopted a double control design to verify the effectiveness of the proposed multimodal intelligent evaluation system. The 20 participants were evenly divided into two groups:

Experimental group: Training using the developed multimodal intelligent evaluation system. The system provides exclusive training suggestions based on the personalized data analysis results of each athlete, including training in batting posture, position adjustment, ball control skills, and tactical thinking.

Control group: Traditional training methods are used, regular training is carried out based on the coach's experience and the athlete's self-feedback, and no intelligent evaluation system is used.

The athletes in each group received 8 weeks of training, three training sessions per week, and each training lasted 1.5 hours.

4.3. Evaluation Indicators

In order to comprehensively evaluate the effectiveness of the multimodal intelligent evaluation system, this experiment set evaluation indicators from two dimensions: technical performance and tactical thinking [8]. The specific evaluation indicators are as follows:

Shooting accuracy: Measures the accuracy of athletes hitting the target ball at different angles and distances. Through comprehensive analysis of sensor data and video data, the evaluation system identifies the stability and accuracy of athletes' movements when hitting the ball.

Club control: Evaluates the athlete's ability to control the force and angle of the club during the hitting process. This indicator is obtained through the analysis of sensor data and mainly reflects the control of the strength and angle of the hitting.

Station and posture stability: Through video data, analyze the athlete's stance posture and body stability when hitting the ball. The system automatically captures the athlete's movement coherence and stability through a deep learning model to ensure the rationality of the stance.

Tactical thinking and decision-making: Through the analysis of game videos, the system evaluates the tactical choices and decision-making quality of athletes in the game. This indicator reflects the athlete's coping ability and strategic planning in complex game situations.

4.4. Training Effect Evaluation

In order to quantitatively evaluate the training effect, the experiment conducted technical tests and game simulations on the participants before, during and after the training, and recorded the changes in various evaluation indicators. By analyzing the performance differences between the experimental group and the control group, the effectiveness of the system in improving skills and tactical thinking was verified.

In addition, the experiment also collected the athletes' satisfaction and subjective feelings about the training method through subjective questionnaires, and further evaluated the advantages of the intelligent system in personalized feedback and training effects.

4.5. Data Analysis

The analysis of experimental data adopts a combination of statistical analysis and deep learning models. Statistical analysis is used to compare the technical performance of the experimental group and the control group to evaluate the improvement effect of the system on various indicators. At the same time, deep learning models identify athletes’ performance patterns and skill improvement trajectories during training by analyzing video, sensor, and physical data.

5. Results

After the 8-week experimental training, the performance of the experimental group and the control group was comprehensively compared and analyzed based on the multi-dimensional evaluation indicators in the experimental design. The results showed that the experimental group using the multimodal intelligent evaluation system achieved significant improvements in various technical performance and tactical thinking indicators, especially in terms of hitting accuracy, club control and tactical thinking, which was better than the control group using traditional training methods.

5.1. Hitting Accuracy

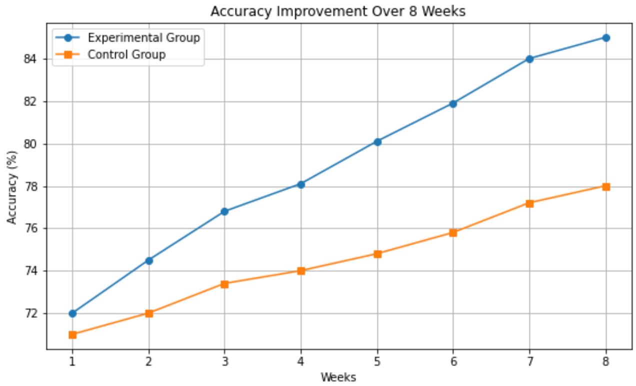

Figure 3 Hitting accuracy improvement over 8 weeks for both the experimental and control groups.

As Figure 3 shows, the experimental results showed that the experimental group had a significant improvement in hitting accuracy. The accuracy of the experimental group athletes in hitting target balls at different angles and distances increased from an average of 72% before training to 85%, while the control group only increased from 71% to 78%. Based on the personalized training feedback of the system, the athletes in the experimental group were able to make precise adjustments to their hitting angles and power control, thereby effectively improving the stability and accuracy of the shots.

5.2. Hitting Accuracy

Figure 4 Control error improvement over 8 weeks for both the experimental and control groups.

In terms of club control, the improvement of the experimental group was also significantly better than that of the control group. As Figure 4 shows, through sensor data analysis, the experimental group's batting force control error dropped from an average of 10.5% in the early stage to 5.8%, while the control error of the control group only dropped from 10.3% to 8.9%. This shows that the multimodal intelligent evaluation system can help athletes make fine adjustments in batting force and angle through accurate sensor data feedback, significantly improving their ball control ability.

5.3. Stance and Posture Stability

The experimental group also made significant progress in stance and posture stability. The system captured and analyzed the athletes' stance and posture through video data. The results showed that the number of unstable stance and posture of the experimental group athletes dropped from an average of 2.3 times per batting before training to 1.1 times, while the control group's decline was smaller, from 2.4 times to 1.9 times. This result shows that the intelligent system can help athletes better adjust their stance and posture through video analysis technology, so as to maintain body stability when batting.

5.4. Tactical Thinking and Decision-Making Ability

In terms of tactical thinking and decision-making ability, the experimental group also achieved significant advantages. Through the game simulation test, the accuracy of tactical selection of athletes in complex game situations increased from 65% before training to 83%, while that of the control group increased from 64% to 74%. The strategic planning and decision-making efficiency of athletes in the experimental group in the game were significantly improved, and they were able to more effectively predict the opponent's actions and make the best shot selection.

5.5. Subjective Satisfaction Survey

Through the questionnaire survey, the athletes' satisfaction and subjective feelings about the training methods were collected. The athletes in the experimental group generally gave positive comments on the multimodal intelligent evaluation system. 90% of the athletes in the experimental group said that the personalized feedback provided by the system helped them identify their technical shortcomings and improve quickly through targeted training; in contrast, only 60% of the athletes in the control group were satisfied with the traditional training method. The athletes in the experimental group believed that the real-time feedback and customized training suggestions of the intelligent system improved their training efficiency and stimulated stronger motivation to participate.

5.6. Comprehensive Evaluation

Based on the analysis of the above indicators, the experimental group significantly outperformed the control group in terms of shot accuracy, cue control, position stability, and tactical thinking. Through multimodal data fusion and personalized feedback from deep learning models, the system can help athletes more comprehensively understand their technical shortcomings and develop training plans suitable for their skill levels, thereby achieving rapid technical improvement in a short period of time.

These results fully verify the effectiveness of the multimodal intelligent evaluation system proposed in this study, indicating that it has obvious advantages in improving billiards sports skills, and provides an important reference for the further development of billiards skill training and evaluation systems in the future.

6. Discussion

The experimental results show that the intelligent evaluation system based on multimodal data fusion and deep learning model has significant advantages in improving the skills and tactical thinking of billiard players. Compared with traditional training methods, athletes using the intelligent evaluation system have a greater improvement in shot accuracy, cue control, position stability and tactical thinking. The following discussion will focus on the system's technical innovation, training effects and potential application areas, and propose suggestions for possible future research directions.

Through the multimodal fusion of video, sensor and physical data, the system can fully capture the technical performance of athletes, and accurately identify the correlation between various technical indicators through the self-attention mechanism of the Transformer model. This combination of multimodal data fusion and deep learning algorithms enables the system to provide athletes with personalized, refined feedback and training suggestions.

In contrast, traditional training methods often rely on the subjective experience of coaches or the self-perception of athletes, making it difficult to systematically identify the technical shortcomings of athletes. Therefore, the hitting accuracy, club control and position stability of the experimental group athletes are significantly better than those of the control group, indicating that the multimodal intelligent evaluation system has significant advantages in improving the refinement and personalization of athletes' skills.

Tactical thinking and decision-making ability are one of the important qualities of high-level table tennis players. The experimental results show that athletes using the intelligent evaluation system have significantly improved in tactical selection and decision-making efficiency. By analyzing the decision-making patterns of athletes in the game, the system provides targeted tactical training suggestions, enabling athletes to better predict the opponent's actions and make the best hitting choices.

This was verified in this study: the tactical thinking and decision-making ability of the athletes in the experimental group were significantly improved compared with the control group, especially in the ability to cope with complex game situations and strategic planning. This means that the intelligent system can effectively make up for the lack of tactical analysis and feedback in traditional training through tailored tactical training [9].

The experimental results further prove that personalized training is of great value to athletes of different skill levels. The system can tailor training programs according to the specific technical performance of each athlete, allowing athletes to quickly focus on their weaknesses for targeted training. Beginners can improve basic techniques through the positioning and posture problems identified by the system, while experienced athletes can further optimize their strategies in the game through complex tactical analysis [10].

The advantages of personalized training are not only reflected in the technical level, but also significantly improve the training participation and satisfaction of athletes. 90% of the athletes in the experimental group said that the real-time feedback and personalized suggestions provided by the system greatly improved the effectiveness and motivation of training. This shows that the intelligent evaluation system can improve the learning efficiency and technical progress of athletes through a more scientific and personalized training method [11].

The multimodal intelligent evaluation system in this study also has broad social application prospects, especially in sports education and public welfare projects. For example, in poor mountainous areas with scarce resources, this system can be used as an auxiliary tool for billiards training and education, helping young people to develop fine motor skills, strategic thinking, and interest in subjects such as physics and mathematics [12]. Through the real-time feedback of the intelligent system, young people can learn and train independently without relying on high-level coaching resources, and the personalized feedback of the system will help them master the technology more efficiently.

Although this study has achieved remarkable results, there are still some directions that can be explored in depth in the future. First, the dimension of data collection can be further expanded, and more types of sensors, such as eye tracking devices, can be introduced to analyze the attention allocation and visual decision-making process of athletes in the game. This will help to have a deeper understanding of the comprehensive performance of athletes in the game and formulate a more comprehensive training plan.

Secondly, more diverse deep learning models can be introduced, such as combining generative adversarial networks (GANs) or transfer learning, to improve the generalization ability of the model in small sample conditions. This will help promote the application of multimodal intelligent evaluation systems in more sports and verify their adaptability in different sports situations [13].

Finally, future research can also explore how to apply such intelligent evaluation systems to more public welfare projects, especially in resource-limited areas, to help young people better master sports skills and stimulate their interest in learning through the system's automated analysis and personalized training functions.

In summary, this study demonstrated the significant effect of the multimodal intelligent evaluation system in improving billiards sports skills, especially in terms of shot accuracy, cue control, tactical thinking, etc. Through data-driven deep learning models, the system can provide athletes with comprehensive and personalized technical evaluation and training suggestions, and has broad social application prospects. Future research can further expand the data collection dimension, optimize the model structure, and promote such systems to more sports and education fields.

7. Conclusion

This study proposed an intelligent evaluation system based on multimodal data fusion and deep learning to improve the technical and tactical capabilities of billiard players. Through multi-dimensional data collection and analysis, the system can comprehensively evaluate the players' shot accuracy, cue control and tactical thinking, and provide them with personalized training plans. The experimental results show that the improvement of various technical indicators of athletes using this system is significantly better than that of traditional training methods, especially in shot accuracy and tactical decision-making. This system can not only significantly improve training efficiency, but also has a wide range of social application potential, especially in sports education and public welfare projects. Future research can further optimize the model and expand its application to more sports and educational scenarios.

References

[1] Wu, Mingyue, et al. "A real-time tennis level evaluation and strokes classification system based on the Internet of Things." Internet of Things 17 (2022): 100494.

[2] Gao, Jing, et al. "A survey on deep learning for multimodal data fusion." Neural Computation 32.5 (2020): 829-864.

[3] Romaissa, Beddiar Djamila, Oussalah Mourad, and Nini Brahim. "Vision-based multi-modal framework for action recognition." 2020 25th International Conference on Pattern Recognition (ICPR). IEEE, 2021.

[4] Zhang, Long. "Behaviour detection and recognition of college basketball players based on multimodal sequence matching and deep neural networks." Computational Intelligence and Neuroscience 2022.1 (2022): 7599685.

[5] Cossich, Victor RA, et al. "Technological breakthroughs in sport: Current practice and future potential of artificial intelligence, virtual reality, augmented reality, and modern data visualization in performance analysis." Applied Sciences 13.23 (2023): 12965.

[6] Vaswani, A. "Attention is all you need." Advances in Neural Information Processing Systems (2017).

[7] García-Aliaga, Abraham, et al. "In-game behaviour analysis of football players using machine learning techniques based on player statistics." International Journal of Sports Science & Coaching 16.1 (2021): 148-157.

[8] Shoumy, Nusrat J., et al. "Multimodal big data affective analytics: A comprehensive survey using text, audio, visual and physiological signals." Journal of Network and Computer Applications 149 (2020): 102447.

[9] Wang, Tao. "Development of a multi-level feature fusion model for basketball player trajectory tracking." Systems and Soft Computing 6 (2024): 200119.

[10] O'Halloran, Kay L., et al. "A digital mixed methods research design: Integrating multimodal analysis with data mining and information visualization for big data analytics." Journal of Mixed Methods Research 12.1 (2018): 11-30.

[11] Connaghan, Damien, et al. "Multi-sensor classification of tennis strokes." SENSORS, 2011 IEEE. IEEE, 2011.

[12] Ju, Xiaoliang, et al. "Tactical optimisation and opponent analysis for football teams with big data mining and machine learning." Revista multidisciplinar de las Ciencias del Deporte 24.94 (2024).

[13] Mekruksavanich, Sakorn, and Anuchit Jitpattanakul. "Multimodal wearable sensing for sport-related activity recognition using deep learning networks." Journal of Advances in Information Technology 13 (2022).