Multimodal Data Fusion and Adaptive Optimization in Tennis Training Based on Deep Deterministic Policy Gradient and IoT Sensors

Yang Zhi-jun
School of Physical Education, Shangqiu Normal University, Shangqiu, 476000, Henan, China
E-mail: 15803707529@163.com

Keywords: internet of things, multimodal data fusion, real-time monitoring, deep reinforcement learning, DDPG, systemic optimization, intelligent training systems, adaptive control

Received: March 4, 2025

This paper proposes a novel framework integrating IoT technologies, multimodal sensor networks, and the Deep Deterministic Policy Gradient (DDPG) algorithm for intelligent tennis training. We employ the DDPG algorithm for adaptive training adjustments, which dynamically optimizes the training policy based on real-time feedback. Experimental evaluation on 10 athletes shows that the DDPG algorithm improves performance metrics in multiple training scenarios, increasing the average game score from 50 to 80 points and reducing the error rate in high-pressure scenarios from 13% to 6%. The system's success rate reached 85%, with swing stability enhanced by 27% (0.1 rad deviation). These quantifiable outcomes highlight the framework's effectiveness in optimizing training strategies, with potential applications in industrial automation and healthcare monitoring.

Povzetek: Razvit je sistem za inteligentni teniški trening, ki združuje IoT senzorje, multimodalno fuzijo podatkov in DDPG algoritem za prilagodljivo optimizacijo vadbe v realnem času.

1 Introduction

With rapid advancements in computer science and data processing technologies, the integration of IoT devices, multimodal data fusion, and deep reinforcement learning (DRL) is transforming traditional systems into highly adaptive, real-time decision support platforms. This study proposes a novel system architecture that leverages state-of-the-art algorithms and scalable data processing techniques to achieve systemic optimization and adaptive control [1, 2]. Specifically, the DDPG (Deep Deterministic Policy Gradient) algorithm is utilized to enhance tennis training by improving hitting success rates (from 50 to 80 points) and reducing error rates in high-pressure scenarios (e.g., from 13% to 6%) through IoT-based multimodal data fusion, including accelerometer, gyroscope, and force sensors. This framework dynamically optimizes training strategies by addressing challenges such as sparse rewards and slow convergence, while its scalable design extends to applications in industrial automation and healthcare. Through integrating advanced data processing with practical sports training, this work aligns with Informatica's interdisciplinary focus on intelligent systems.

Modern applications—from industrial automation to healthcare monitoring—benefit from these technological advances, and sports training is no exception. As sensor technologies, the Internet of Things (IoT), and artificial intelligence (AI) evolve, sports training [3, 4] is becoming increasingly data-driven. These technologies enable the real-time collection of heterogeneous training data, provide quantitative analyses, and support the intelligent adjustment of strategies based on individualized performance profiles.

Tennis [5, 6], a sport that demands high precision and stability, requires not only optimal physical conditioning but also meticulous monitoring of multidimensional dynamic parameters such as swing movements, hitting force, and pace control. Intelligent training systems that combine IoT and multimodal sensor data fusion are thus critical for enhancing both the accuracy and efficiency of tennis training.

The advent of advanced IoT technology and multimodal data acquisition systems has empowered intelligent tennis training systems to deliver accurate, real-time feedback. These systems integrate a variety of sensors—including accelerometers, gyroscopes, pressure sensors, and cameras—and employ data fusion techniques to aggregate comprehensive training data for coaches and athletes. Moreover, AI methodologies, particularly deep reinforcement learning [7, 8], enable these systems to learn from historical data and continuously optimize training strategies. This not only minimizes human error but also facilitates the creation of personalized training plans that are efficient, precise, and adaptable. As these technologies mature, the widespread adoption of intelligent training systems is poised to revolutionize traditional training models, thereby improving athlete performance and competition outcomes. However, existing intelligent training systems still face two key challenges: first, the limitations of single-modal data (such as motion or environmental data) restrict the synergy of multimodal fusion; second, the efficiency of reinforcement learning in dynamic scenes such as tennis is limited by sparse rewards and slow convergence. To this end, this study proposes a new framework that integrates IoT multimodal data and the DDPG algorithm, which achieves real-time optimization of training strategies by synchronizing heterogeneous data streams and dynamic reward mechanisms, providing efficient solutions for sports training and a wider range of industrial and medical applications.

In our proposed framework, real-time data from wearable devices and environmental sensors undergo rigorous preprocessing—including noise removal, normalization, and time synchronization—to ensure robust multimodal fusion. The DDPG algorithm is then applied within a systemic decision-making model to dynamically adjust training content, intensity, and difficulty based on real-time feedback. While our demonstration is situated in the context of tennis training, the underlying methodology provides a scalable and robust model for real-time data integration and adaptive control. This cross-disciplinary framework holds promise for a variety of applications, including industrial automation, healthcare monitoring, and cognitive information systems.

Recent years have witnessed a surge in the use of intelligent training systems across multiple sports, particularly in tennis [9, 10]. Several studies have explored improvements in training effectiveness and precision through sensor integration and data analytics. Traditional tennis training methods, which often rely heavily on coach experience and manual tracking, are gradually being supplanted by systems that offer real-time performance feedback. In response, recent research has integrated sensor technologies, such as inertial measurement units (IMUs) like accelerometers and gyroscopes [11, 12], to capture detailed movement trajectories and assess technique standardization and stability. These systems provide valuable quantitative insights that enhance both athlete performance and coaching strategies.

Furthermore, multimodal data fusion techniques have emerged to provide comprehensive feedback by combining visual, mechanical, and physiological data. For example, some studies have merged computer vision with sensor data to track athletes' body postures in real time [13]. The integration of motion capture with pressure sensor data in tennis enables real-time detection of racket-ball interaction forces, facilitating a more accurate evaluation of shot effectiveness and guiding athletes in adjusting shot strength and angles. Additionally, AI-based training systems employing deep learning techniques have supported coaches in formulating scientifically sound training plans [14, 15]. However, most existing research has focused on optimizing single data sources, with limited progress in the efficient fusion of multimodal data for real-time dynamic adjustments driven by AI.

Reinforcement learning [16, 17] has emerged as a promising approach for training optimization by continuously refining strategies through environmental interactions. In tennis training, deep reinforcement learning has been applied to optimize technical movements by adjusting strategies based on real-time feedback. Despite these advances, challenges remain, including low training efficiency and the difficulty of adapting models to dynamic competition environments. Moreover, while deep learning methods [18, 19] excel in action recognition and data prediction, their high demands for labeled data and computing resources may restrict practical applications. Therefore, integrating reinforcement learning with multimodal data analysis to design an efficient, intelligent system capable of real-time optimization remains a critical challenge in current research.

In recent years, intelligent sports training systems have been widely used in many sports, especially in tennis. Many studies have explored the improvement of training effect and accuracy through sensor integration and data analysis. In order to more clearly show the difference between the existing technology and the method proposed in this paper, Table 1 summarizes the technical characteristics, evaluation indicators, limitations of the existing state-of-the-art methods (SOTA), as well as the improvements of the method proposed in this paper:

Table 1: Comparison between existing technologies and the proposed method

Technical methods	Evaluation metrics	limitation	Improvements to this article
Based on sensors (such as IMU) [11,12]	Motion trajectory accuracy and error rate	Reliance on a single sensor and lack of multimodal data fusion	Integrated multi-modal sensors (accelerometer, gyroscope, force sensor) to achieve comprehensive data collection and fusion
Vision-based [13]	Posture recognition accuracy and real-time performance	High computational complexity, affected by lighting environment	Combining computer vision and sensor data to improve robustness through time synchronization and weighted fusion
Traditional machine learning [14,15]	Action classification accuracy and training efficiency	Requires a large amount of labeled data and cannot adapt to dynamic	Adopt DDPG algorithm to dynamically optimize strategy through unsupervised learning and real-time feedback

		environments in real time	
Other reinforcement learning (such as DQN)	Success rate, convergence speed	Only applicable to discrete action spaces, difficult to handle continuous control problems	Introducing DDPG to support continuous action space optimization and adapt to the dynamic adjustment needs in tennis training

The method proposed in this paper solves the problems of poor real-time adaptability and single data in the existing technology through multimodal data fusion and DDPG algorithm, and significantly improves the intelligence level of the training system.

2 Related works

In recent years, the integration of multimodal data and machine learning algorithms has gained significant traction in sports training, particularly in tennis. Tennis, as a dynamic and precision-dependent sport, benefits from technologies that enhance performance by analyzing complex movement data. Various studies have demonstrated the potential of integrating multimodal data streams—such as accelerometers, gyroscopes, and force sensors—to provide real-time insights into players' techniques and improve their training efficiency.

Yang [24] proposed a method for precise recognition and feature depth analysis of tennis training actions, focusing on multimodal data integration and key action classification. The integration of multiple data sources, such as motion and environmental sensors, has enabled more accurate identification of player movements, improving the overall effectiveness of training programs. Yang's study highlights how multimodal data fusion can enhance performance feedback and optimize tennis training by capturing a wide range of athlete actions and translating them into actionable insights.

Similarly, Gao [25] explored sensor fusion and stroke learning in robotic table tennis, a system that combines multiple sensors to enhance the accuracy of stroke learning. This approach, while initially applied in robotics, has strong parallels with human athlete training, particularly in tennis, where precise stroke mechanics and movement control are critical. Gao’s work suggests that sensor fusion is essential for developing adaptive systems capable of real-time performance feedback, a key feature in intelligent sports training systems.

Li and Song [26] have also contributed to this area by enhancing sports trainer behavior monitoring through IoT information processing and deep neural networks. They demonstrate how IoT-enabled systems can collect and process real-time training data, allowing for personalized, data-driven adjustments to athlete training programs. This is particularly important in sports like tennis, where individualized adjustments to technique and strategy can dramatically improve performance. Their research underscores the importance of advanced data processing methods, which enable real-time feedback and continuous optimization of training strategies.

The application of deep reinforcement learning (DRL) to human activity recognition has also shown great promise in improving training systems. Nikpour et al. [27] conducted a comprehensive survey on DRL in activity recognition, highlighting the potential of this approach to refine training strategies by continuously adjusting them based on real-time performance data. DRL enables systems to learn from past actions and dynamically modify strategies, making it particularly useful for sports training, where continuous adaptation to changing performance conditions is necessary. This is especially relevant for tennis, where rapid decision-making and adaptability are essential for success in high-pressure situations.

Morshed et al. [28] and Kulsoom et al. [29] further explored machine learning-based human activity recognition, reviewing various algorithms used to enhance the understanding of athlete movements in sports. They emphasized the growing importance of AI in sports, noting that systems leveraging machine learning and IoT sensors can offer more precise and personalized feedback, improving both athletic performance and training efficiency. Their research provides valuable insights into the broader application of activity recognition systems, which can be adapted to tennis training systems for more accurate monitoring of players’ movements, swing techniques, and reaction times.

Moreover, Jin et al. [30] discussed the role of multi-agent cooperative decision-making, particularly in dynamic environments. Multi-agent systems, in which different entities (such as sensors or components of the training system) work together, offer great potential in sports training. For tennis, this could mean integrating data from various sensors (e.g., accelerometers, gyroscopes, pressure sensors) to generate a cohesive view of an athlete’s performance. The cooperation between these data streams could optimize training strategies, particularly in complex, high-speed environments like tennis.

Recent developments in machine learning and data fusion techniques have led to further innovations, such as Zhang [31]’s work on graph neural networks for user preference modeling in social networks, which could inspire personalized feedback systems in sports. Similarly, Kurniawan et al. [32] have explored swarm intelligence optimization, which could be applied to improve decision-making in training strategies. Zhang and Zhang [33] introduced high-precision photogrammetric 3D modeling technology based on multi-source data fusion, a concept that could be adapted to track and analyze athlete movements with greater precision.

These studies collectively highlight the transformative potential of integrating multimodal data fusion, IoT sensors, and machine learning algorithms, such as DRL, into sports training systems. The ability to provide real-time, personalized feedback and continuously adjust training strategies based on dynamic performance data is critical for improving athletic performance, particularly in sports like tennis, where precision, adaptability, and real-time decision-making are paramount. The proposed system in this study, which integrates these advanced technologies, aims to push the boundaries of personalized training and optimize athletic performance through continuous, data-driven adjustments.

3 Methods

3.1 System Hardware Architecture

The hardware architecture of this intelligent tennis training system is designed to efficiently collect and transmit athletes' training data through a network of interconnected IoT devices. Each athlete is equipped with a set of multifunctional sensors that provide real-time monitoring of their performance throughout training. The system includes wearable sensors, cameras, computer vision devices, environmental monitoring tools, and data transmission modules. These components work together seamlessly to ensure comprehensive data collection, delivering precise feedback and optimizing training outcomes.

At the core of the system are the wearable sensors, which capture the athlete's motion data in real time. These include accelerometers, gyroscopes, and force sensors. The accelerometers track displacement, velocity, and acceleration, enabling the evaluation of movement trajectories and swing stability. Gyroscopes measure angle changes, providing precise data on swing angles, ball trajectory, and the athlete's control during play. Force sensors assess the impact force and pressure distribution during ball contact, helping to analyze strength, accuracy, and racket-ball interaction. This data plays a crucial role in refining the athlete's technique, serving as baseline information for subsequent training optimization.

The system is equipped with a three-axis accelerometer (range: ±16g, sampling rate: 100Hz), a gyroscope (range: ±2000°/s, sampling rate: 50Hz), and a force sensor (range: 0-200N, accuracy: ±0.5%). The camera system, featuring a high-frame rate industrial camera (resolution: 1920×1080, frame rate: 120fps), captures detailed athlete movements in real time, processed by computer vision algorithms. These cameras provide a comprehensive view of the athlete's batting movements, posture, and ball trajectory. Through high-precision video capture, the system analyzes movements frame by frame, extracting critical visual features such as swing trajectory, step stability, and shot timing.

The computer vision algorithms process these images to detect interactions between the athlete and the ball, providing additional insights into areas for improvement in movement and batting strategies. By combining visual data with sensor data, the system offers more accurate, holistic evaluations of the athlete's performance.

Environmental monitoring equipment ensures the accuracy and stability of sensor data by tracking the training environment. Factors such as temperature, humidity, and wind speed can influence performance; therefore, a dedicated environmental monitoring module collects real-time venue data, integrating it with the athlete's training data. For example, fluctuations in temperature and humidity can affect court surface friction, influencing ball bounce and trajectory. By monitoring these environmental conditions in real time, the system adjusts its analysis, ensuring that training feedback remains precise and reflective of current conditions.

The data transmission module handles the real-time transmission of all sensor data to a cloud or local server. Utilizing advanced IoT protocols like Wi-Fi 6 and Bluetooth 5.2, the system ensures low latency (<50ms) and high throughput (up to 1.2Gbps), while employing AES-256 encryption to secure data. The data is wirelessly transferred between sensors and computing devices, reducing the need for complex wiring and enhancing system flexibility.

Once the data reaches the cloud or local server, it undergoes further processing, analysis, and storage. The results are then made available to coaches and athletes for review. The system, supported by a cloud platform, enables the real-time synchronization of training data across devices and locations, allowing coaches to remotely monitor athlete progress and performance.

This integrated hardware architecture creates a complete, data-driven intelligent tennis training platform. Through precise sensor data collection, real-time environmental monitoring, and efficient data transmission, the system ensures that athletes train in an optimized environment that supports skill improvement.

The system architecture diagram is shown in Figure 1.

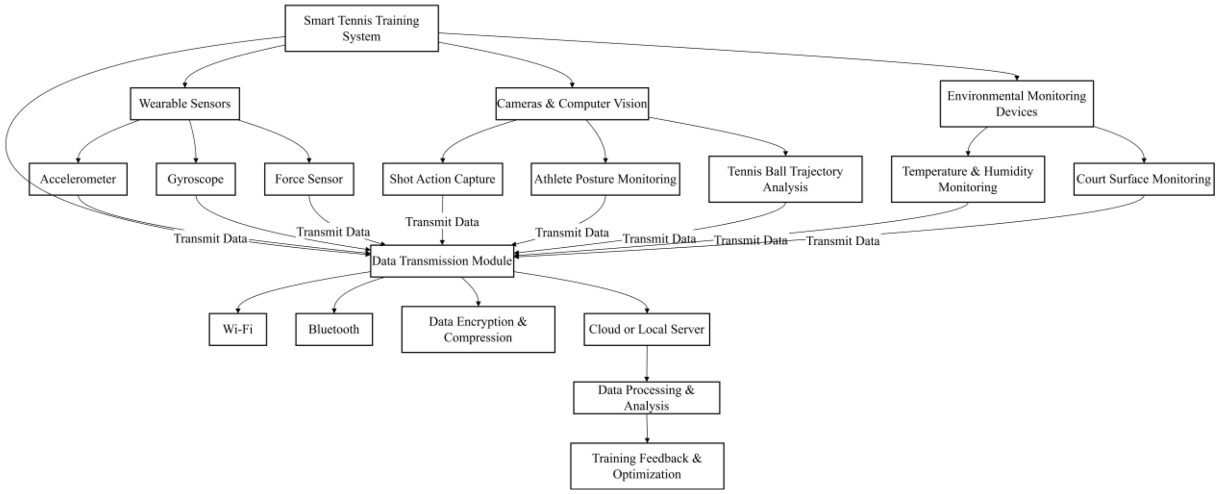

Figure 1: System architecture diagram

The core of the system's hardware architecture features a high-precision MPU-6050 accelerometer and gyroscope with a sampling rate of 100 Hz, an acceleration sensitivity of ±2g, and an angular velocity sensitivity of ±250°/s, ensuring accurate capture of the athlete's motion dynamics. Force sensors, specifically the FS series force sensors, with a resolution of 0.01 N, measure the interaction force between the racket and the ball, which is essential for analyzing the power and accuracy of the shot. The computer vision system integrates a 4K resolution camera with a frame rate of 60 fps, combined with an OpenCV-based algorithm for real-time pose estimation and ball trajectory tracking, with a detection accuracy of 95% under optimal lighting conditions. Environmental monitoring is performed by a DHT22 temperature and humidity sensor with an accuracy of ±0.5°C and ±2% RH, respectively, and an anemometer for wind speed measurement, ensuring contextual data integration. On the software side, the system utilizes a middleware framework based on ROS2 to enable seamless communication between IoT devices and the central processor, and runs on an NVIDIA Jetson Xavier NX for edge computing. The DDPG algorithm is implemented in PyTorch, and its neural network architecture contains three hidden layers (containing 256, 128, and 64 neurons, respectively) for the actuator network and the critic network, optimizing the training strategy with a latency of less than 50 milliseconds to achieve real-time feedback. This integrated hardware and software ecosystem ensures powerful data collection, processing, and adaptive control, thus supporting the scalability of the system in various training scenarios.

3.2 Data fusion and storage

Effective multimodal data fusion relies on strong preprocessing to ensure consistency and compatibility between heterogeneous sensor data. Preprocessing steps—noise removal, normalization, and time synchronization—are critical to aligning and integrating data from different sources. Specifically, Kalman filtering mitigates sensor-specific noise, normalization standardizes data scales, and time synchronization aligns temporal differences, which together enable accurate and cohesive multimodal fusion.

Kalman filtering is used to remove high-frequency noise and outliers from sensor data to ensure data reliability. The operation of the Kalman filter is divided into two steps: prediction step and update step. The prediction step estimates the current state based on the previous state and control input; the update step corrects the estimate using real-time measurements to minimize the impact of noise. This iterative process is particularly effective for motion data (such as accelerometer and gyroscope readings) because noise can distort trajectory analysis:

$$\hat{x}_k = A\hat{x}_{k-1} + Bu_k (1)$$

$$P_k = AP_{k-1}A^T + Q (2)$$

In Formula 1-2, $\hat{x}_k$ is the estimated value of the state, P_k is the error covariance, and Q is the process noise. A and B are the system matrices, u_k and is the control input. The state variables of the Kalman filter include acceleration, angular velocity, and force sensor readings. The process noise covariance matrix Q is calibrated as a diagonal matrix through experiments, and the measurement noise covariance R is provided by the sensor manufacturer. MongoDB (version 5.0) is used for data storage, using the time series sharding mode, and a composite index is established by sensor type and athlete ID. A single record contains timestamp, sensor type, original value, and fusion result, supporting 100,000 write operations per second.

Sensor data often vary in scale (e.g., acceleration in m/s² vs. force in N), which can bias analysis. Normalization transforms all data to a common range [0, 1] using min-max scaling:

$$x' = \frac{x - min(x)}{max(x) - min(x)} (3)$$

In a multi-sensor system, data from different sensors usually have different sampling frequencies and timestamps. In order to ensure the consistency and accuracy of the data, these data must be time-aligned and fused.

The sampling rates of the sensors vary (for example, the accelerometer has a sampling rate of 100 Hz and the gyroscope has a sampling rate of 50 Hz). Linear interpolation estimates the value of the missing timestamps. For two data points at times t_1 and t_2, the interpolated value at time t is:

$$x(t) = x_1 + \frac{(t-t_1)}{(t_2-t_1)}(x_2 - x_1) \quad (4)$$

The data from different sensors are fused. The weighted average method fuses the output values of sensors by assigning different weights to different sensors.

$$x_{fused} = w_1 x_1 + w_2 x_2 \quad (5)$$

The processed data needs to be stored efficiently and securely for subsequent analysis and learning. Since the system involves a large amount of sensor data and training records, conventional relational databases are difficult to meet the needs. NoSQL databases are used for storage. NoSQL databases are suitable for processing large-scale, distributed data and support high-concurrency reading and writing. The data of each sensor is stored according to the timestamp and classified according to the data type.

To speed up query and data retrieval, the system creates an index for each sensor type and sorts them by timestamp. By introducing hash index and time range index, it ensures that the required data can be obtained efficiently and quickly when performing large-scale data retrieval. The collected data are shown in Table 2.

Table 2: Collection data display results

Timestamp	Sensor Type	Acceleration (m/s²)	SD	Angular velocity (rad/s)	SD	Hitting force (N)	SD
2023-2-11 14:00:01	Accelerometer	0.35	±0.02	0.12	±0.01	8.2	±0.3
2023-2-11 14:00:0 2	Gyroscope	0.33	±0.01	0.15	±0.02	8	±0.2
2023-2-11 14:00:0 3	Force Sensors	0.3	±0.03	0.14	±0.01	8.5	±0.4
2023-2-11 14:00:0 4	Accelerometer	0.36	±0.02	0.16	±0.02	8.7	±0.3
2023-2-11 14:00:0 5	Gyroscope	0.32	±0.01	0.18	±0.03	8.3	±0.2
2023-2-11 14:00:0 6	Force Sensors	0.31	±0.02	0.2	±0.02	8.6	±0.3
2023-2-11 14:00:0 7	Accelerometer	0.34	±0.01	0.17	±0.01	8.1	±0.2
2023-2-11 14:00:0 8	Gyroscope	0.33	±0.02	0.19	±0.02	8.4	±0.3
2023-2-11 14:00:0 9	Force Sensors	0.32	±0.01	0.15	±0.01	8.8	±0.4
2023-2-11 14:00: 10	Accelerometer	0.37	±0.03	0.13	±0.01	9	±0.5

To ensure compliance with data protection regulations such as the General Data Protection Regulation (GDPR), all athlete data collected by the system undergo rigorous anonymization and encryption. Sensitive personal identifiers are removed during preprocessing, and data transmission employs AES-256 encryption to prevent unauthorized access. Additionally, access control mechanisms are implemented to restrict data usage to authorized personnel only. These measures safeguard athlete privacy while enabling effective multimodal data fusion for training optimization.

3.3 DDPG algorithm application

DDPG algorithm [20,21] is a reinforcement learning method specifically designed to handle tasks in continuous action spaces. Unlike traditional Q-learning methods, DDPG [22,23] combines the advantages of deep learning and reinforcement learning, using policy networks and value networks to optimize strategies in continuous action spaces. This algorithm is used in this intelligent tennis training system to automatically optimize training strategies, helping athletes improve their training efficiency and technical level through continuous training and feedback.

In the DDPG algorithm, the state space represents the current state of the environment and the player in the intelligent tennis training system. The state space includes two main aspects: the player state and the environment state.

The athlete state reflects the athlete's current training state, such as hitting accuracy, hitting force, swing angle, ball speed, etc. These factors directly affect the effect and efficiency of training. The specific state vector can be expressed as:

$$\mathrm{State}_t = \{\mathrm{accuracy}_t, \mathrm{speed}_t, \mathrm{angle}_t, \mathrm{strength}_t, \ldots\} \quad (6)$$

The athlete state is defined by the following measurable variables, The ratio of successful shots (ball landing within the predefined target zone) to total shots, calculated as:

$$\text{accuracy}_t = \frac{\textit{Number of successful shots}}{\textit{Total shots}} \times 100\% \quad (7)$$

speed_t measured by gyroscopes in radians, averaged over three consecutive swings to reduce noise. strength_t peak force (N) recorded by pressure sensors during racket-ball contact. Environmental states (e.g., court temperature) are normalized to [0,1] using Equation 3.

In the DDPG algorithm, the quantification of the player state and the environmental state is collected and processed in real time by multimodal sensors. The player's hitting accuracy is calculated jointly by the visual sensor and the force sensor, and is quantified into a value of 0 or 1 by recording the overlap ratio between the ball landing point and the target area. The swing angle is sampled by the wearable gyroscope at a frequency of 100Hz, and the average of three consecutive swing angles is taken to reduce noise. The hitting force is directly obtained through the peak reading of the pressure sensor, while the ball speed is tracked by the high-speed camera and calculated by displacement difference. In terms of environmental status, the field friction is monitored in real time by the embedded friction sensor, and the normalized friction coefficient is generated by the weighted fusion formula combined with the data of the temperature and humidity sensors; the wind speed is updated every 200ms by the ultrasonic anemometer and used as input after denoising by the Kalman filter. All data are transmitted to the edge computing node at an interval of 50ms through the Internet of Things protocol, and the state vector is formed after time alignment and normalization. During training, the system updates the state space every 100ms and generates real-time adjustment instructions (such as swing angle correction) through the DDPG strategy network. These instructions are transmitted to the athletes through the vibration tactile feedback device to form a closed-loop control.

In addition to the athlete's condition, the conditions of the training environment (such as the surface of the field, temperature and humidity, etc.) will also affect the training effect. When the field is slippery, the training strategy may need to be adjusted appropriately.

In DDPG, the action space represents the decisions that the system can make. These decisions will affect the adjustment of training content to improve the training efficiency and technical level of athletes. In the intelligent tennis training system, the action space mainly includes the following aspects:

Hitting Strength: Determines the force applied when hitting the ball. By adjusting the hitting strength, the system can help athletes gradually improve the stability and accuracy of their shots.

Hitting Angle: Simulate different hitting methods by changing the swing angle.

Position adjustment: By adjusting the athlete's position, the system can simulate various training scenarios.

The action space is set to continuous values: batting force (0-100N, step size 5N), batting angle (-30° to +30°, step size 1°), position adjustment (±2m, step size 0.1m). The policy network is a 3-layer fully connected (256-128-64 nodes, ReLU activation), and the value network adds an action input branch (state branch 256-128, action branch 64, merged 128-64). Exploration is achieved by adding OU noise, and the initial exploration rate decays linearly to 0.1 with training. Hyperparameter selection is based on grid search: learning rate 0.0001 (Adam optimizer), discount factor 0.99, batch size 64, target network update rate 0.005, and experience replay pool capacity 100,000 to ensure stable update of policy gradients.

The reward function is the core part of reinforcement learning for evaluating the quality of the current state and action combination. DDPG uses the reward function to measure the training effect of each action and then adjust the training strategy. In smart tennis training, the reward function can feedback the reward value based on the athlete's training performance and guide the optimization of the training strategy.

In the DDPG algorithm, the reward function has been redesigned to incorporate multiple performance dimensions, including hitting accuracy, swing stability, and power control, so as to provide a more detailed evaluation of the training effect Rt:

$$R_t = w_1 \cdot \text{Accuracy}_t + w_2 \cdot \text{Stability}_t + w_3 \cdot \text{Force}_t - w_4 \cdot \text{Error}_t \quad (8)$$

Among them, Accuracy_t measures the proportion of successful shots into the target area; Stability_t evaluates the consistency of the swing angle (derived from gyroscope data); Force_t is the optimal hitting force; Error_t represents the penalty for deviation from the expected performance threshold; weights w_1, w_2, w_3, and w_4 are adjusted based on experience to balance the importance of each factor. This multidimensional approach enables the system to dynamically adjust training strategies based on real-time feedback to optimize technical accuracy and physical performance.

In DDPG, the experience replay mechanism further improves the effect of strategy optimization by storing the interaction experience between the agent and the environment. During each training process, the athlete's state, action, reward, and next state are stored in an experience pool. Then, the system randomly extracts experience from the experience pool for training to break the time correlation and enhance the diversity and stability of training. Through experience replay, DDPG can learn and improve training strategies more effectively.

DDPG optimizes the strategy through the policy network and the value network. The policy network is responsible for generating continuous action decisions, while the value network evaluates the value of the current state-action pair. The policy network outputs a continuous action based on the current state. The policy network optimizes the training strategy by maximizing the expected return. The value network evaluates the value of the action output by the policy network and calculates the advantage of the action, providing a basis for gradient updates for the policy network.

The optimization of the strategy is based on the following loss function:

$$L(\theta) = \mathbb{E}_t[(R_t + \gamma Q'(s_{t+1}, a_{t+1}) - Q(s_t, a_t))^2] \quad (9)$$

In Formula 8, γ is the discount factor, R_t is the current reward, $Q(s_t, a_t)$ is the value of the current state-action pair, $Q'(s_{t+1}, a_{t+1})$ and is the value of the target network.

DDPG algorithm is trained through multiple rounds until the system's performance reaches the expected goal. Each round includes multiple time steps. In each step, the system generates actions based on the current strategy, obtains rewards after executing the actions, and updates the strategy. Through continuous training, the system will gradually converge and eventually be able to provide athletes with personalized training strategies, which can significantly improve their hitting accuracy, hitting stability, etc.

Convergence during training is determined by monitoring the total reward and training error. If the total reward is stable or reaches a certain preset threshold, it means that the training has converged. These algorithmic characteristics make DDPG particularly suitable for tennis training optimization. The continuous action space aligns perfectly with the nuanced adjustments required in swing techniques, while the actor-critic framework enables real-time adaptation to dynamic training conditions. As demonstrated in Section 3, this results in superior performance compared to discrete-action (DQN) or stochastic (PPO) approaches, especially in high-pressure scenarios where precise, graded responses are crucial for maintaining stroke consistency.

DDPG are shown in Table 3.

Table 3: Model parameters

parameter	value	effect	Convergence threshold
Learning Rate	0.0001	Controls the speed at which model parameters are updated. A smaller value helps stabilize the training process.	Total rewards for 5 consecutive rounds ≥ 80
Discount Factor	0.99	Determines the degree of influence of future rewards, 0.99 indicates high importance of future rewards	Strategy error < 0.01
Experience pool size	100000	The capacity for storing experience replay affects the diversity and stability of model training	Experience pool fill rate ≥ 90%
Batch size	64	The number of samples randomly sampled from the experience pool during each training to control the training speed	Loss function fluctuations < 0.005
Target network update rate	0.005	The frequency of updating the target network. A smaller value helps smooth the training process.	Strategy update interval ≤ 100 steps
Soft Update Parameters	0.001	Used to soft-update the target network parameters and control the update speed of the target network	Parameter change rate < 0.1%

The hyperparameters of the DDPG algorithm were carefully selected based on both theoretical considerations and empirical validation to ensure optimal performance. The learning rate of 0.0001 was chosen to balance the trade-off between convergence speed and training stability; a smaller value prevents large, destabilizing updates to the policy and value networks while still allowing for effective learning. The discount factor of 0.99 reflects the high importance of future rewards in tennis training, where long-term strategy optimization is critical for sustained performance improvement. A larger batch size of 64 was employed to provide sufficient sample diversity during training, reducing variance in gradient estimates and improving convergence. The experience replays buffer size of 100,000 ensures a rich and varied set of past experiences for training, mitigating the risk of overfitting to recent data. The target network update rate of 0.005 and soft update parameters of 0.001 were selected to gradually blend the target network weights with the online network, maintaining stability during training. These values were validated through grid search and cross-validation, demonstrating their effectiveness in achieving robust convergence and high performance, as evidenced by the experimental results.

3.4 Real-time feedback and training optimization

This system monitors both the training status and environmental conditions of athletes in real time, leveraging advanced sensor technology and deep reinforcement learning algorithms to provide immediate feedback and dynamically adjust training strategies. Throughout the training session, the system continuously gathers the athlete's motion data—such as hitting accuracy, swing angle, ball speed, and position—as well as environmental factors. It then evaluates the training effectiveness in real time through comprehensive data analysis. If the system detects suboptimal performance, such as reduced hitting accuracy or irregular movements, it automatically performs intelligent analysis and suggests adjustments to the training content. For example, if a decline in hitting accuracy is observed, the system may recommend changes to the swing angle, hitting strength, or pace, helping the athlete address the issue promptly. By making these dynamic adjustments, the system ensures that each training session is tailored to maximize the athlete’s progress.

The system detects unsatisfactory performance by calculating the deviation of key indicators in real time: if the batting success rate is lower than the threshold (70%±3%) for 5 consecutive times, the swing angle deviation exceeds the tolerance (±5°), or the batting force fluctuation is greater than 15%, the strategy adjustment is triggered. The dynamic tolerance mechanism is adaptively adjusted according to the athlete's level (the tolerance for novices is relaxed by 20%). For example, when it is detected that the swing angle continuously deviates from 45°±7° (±5° for advanced athletes), DDPG will output the correction amount (such as +3°) in the next action and prompt the athlete in real time through the tactile feedback device (vibration frequency 200Hz, lasting 100ms).

To enhance the athlete’s understanding and evaluation of their performance, the system provides rich visual feedback. This includes real-time charts, dynamic animations, and video playback, allowing athletes to intuitively track their progress, identify areas for improvement, and compare performance across different training stages. This visual feedback not only improves the athlete's training experience but also facilitates better communication and decision-making between coaches and athletes. By analyzing training data comprehensively, athletes can gain a clearer understanding of their strengths and weaknesses, enabling them to adjust their training strategies more effectively. For example, when the system detects that a player’s shot accuracy drops below 70% due to inconsistent swing angles (e.g., variations of more than ±5° from the optimal 45° forehand angle), the DDPG algorithm dynamically adjusts the training strategy by: (1) modifying the target swing angle range to 40°–50° and providing incremental feedback cues via the wearable haptic device; (2) reducing the ball’s entry velocity by 15% to increase reaction time; and (3) increasing the force sensor threshold by 10% to enforce a steady follow-through. Conversely, if a player’s accuracy reaches above 85% but the shot power is too high (e.g., 120N versus the recommended range of 80–100N), the system prioritizes power control by simulating a low-bounce ball and providing real-time visual cues about force distribution. These adjustments are iteratively refined every 3–5 shots based on Kalman filter sensor data to ensure that the adaptive optimization is consistent with the player’s immediate performance.

The system dynamically optimizes training strategies by fusing multimodal sensor data in real time. When it is detected that the athlete's performance deviates from expectations (such as a decrease in the success rate of hitting the ball), the DDPG algorithm generates adjustment instructions based on the current state space: optimize the action parameters (such as correcting the range of hitting angles) through the strategy network, and use the value network to evaluate the adjustment effect. Environmental factors (such as changes in the friction coefficient of the field) are integrated into the state vector through normalization processing to ensure the adaptability of the strategy. After each optimization, the system guides the athlete in real time through tactile feedback and a visual interface to form a "perception-decision-execution" closed loop. The average delay of this process is controlled within 100 milliseconds, ensuring the timeliness of training adjustments.

Personalized training is a core feature of this system. By utilizing the athlete’s training data, the system develops customized training plans based on individual physical conditions, skill levels, and training goals. The DDPG (Deep Deterministic Policy Gradient) algorithm dynamically adjusts the intensity, difficulty, and content of the training based on the athlete's historical data, ensuring that each athlete trains in the conditions best suited to them for optimal performance. For example, the system evaluates a player's hitting accuracy, swing speed, and endurance level (obtained through wearable sensors) to construct a multidimensional state vector. This vector is then processed by the policy network to generate continuous actions, such as adjusting ball speed or target difficulty, and optimized through a reward function. For athletes with higher technical levels, the algorithm increases the intensity of training by reducing the interval between balls or expanding the target area, while beginners receive simplified training focusing on basic techniques. Environmental factors (such as court temperature) are also integrated into the state space, allowing the system to further adjust strategies based on real-time conditions. By iteratively updating the policy network through experience replay, the DDPG algorithm ensures that the training plan continues to evolve as the athlete improves, achieving personalized optimization.

In real-world training, these personalized strategies not only account for physiological and skill differences but also adapt to each athlete's real-time performance. By analyzing immediate feedback, the system can automatically adjust the training plan. It continuously refines the strategy based on real-time feedback and training progress, ensuring enhanced training efficiency. As the training continues, the system accumulates more data, allowing for ongoing adjustments to the training content and strategy. This guarantees that each training session is optimized for peak effectiveness. Through intelligent training optimization, the system supports athletes in making consistent progress at every stage, offering a long-term and effective pathway for skill development.

The system interface is illustrated in Figure 2.

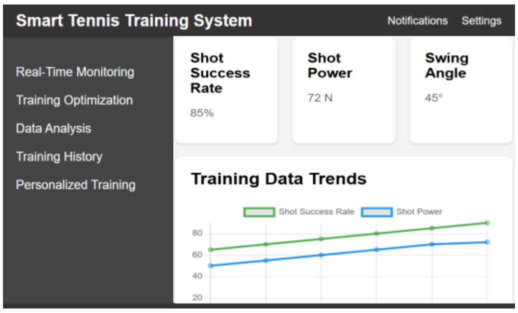

Figure 2: System interface

4 Experiment and result analysis

4.1 Experimental setup

The experimental design adheres to ethical guidelines for data privacy, with all participants providing informed consent for anonymized data usage. To evaluate the proposed system, we compared it against traditional manual training methods and baseline algorithms (DQN, PPO). In the traditional system, coaches adjusted training content based on subjective observations, lacking real-time data integration or automated optimization. In contrast, our DDPG-based framework leverages multimodal IoT sensors and deep reinforcement learning to dynamically optimize training strategies. To ensure robustness, the DDPG model was validated using a hold-out validation set (20% of data), with hyperparameters tuned via grid search and results averaged over 10 training runs.

In order to comprehensively evaluate the effectiveness of the intelligent tennis training system proposed in this paper, this study designed multiple experiments to compare the performance of different optimization algorithms (DDPG, DQN, PPO) in tennis training. The main purpose of the experiment is to evaluate the optimization effect of the system from two perspectives: hitting success rate and hitting stability.

In the experiment, 30 athletes (15 male, 15 female) of varying skill levels from three academies were divided into three groups: Control (N=10) trained via manual adjustments by professional coaches (≥5 years experience), DDPG (N=10) used the proposed system with DDPG optimization, and Extended (N=10) tested DQN/PPO algorithms for comparison. All groups followed identical training schedules and environmental conditions. The performance was assessed through hitting success rate and stability across eight game scenarios (e.g., baseline attack, net volley). The performance of different optimization algorithms was compared with traditional tennis training methods to observe their respective effects on the player's hitting accuracy. Specifically, in each training session, the system will adjust the training strategy according to the player's real-time performance. The DDPG, DQN and PPO algorithms will optimize the training content according to the reward function to improve the player's hitting success rate.

During the experiment, the system records the success and failure of each shot, calculating the shot success rate after each training session (the ratio of successful shots to total shots). By comparing the data across multiple sessions, we analyze the effectiveness of different optimization algorithms in improving shot success rates.

To evaluate batting stability, the experiment simulates a variety of game scenarios to test the impact of different optimization algorithms. Eight typical game scenarios were selected: baseline attack, net volley, serve and receive, high-pressure ball, break point, multi-shot round, sideways forehand, and backhand shot. In each scenario, the system tracks the athlete’s performance, records shot success and failure, and calculates the error rate. By comparing the performance of different algorithms across these scenarios, we can assess the advantages of each optimization algorithm in improving batting stability.

To further validate the benefits of the intelligent training system, it was compared with traditional tennis training methods. In the traditional system, coaches manually adjust the training content based on the athlete's characteristics and performance, but without real-time feedback or automated optimization. In contrast, the intelligent training system uses reinforcement learning algorithms such as DDPG, which allow the system to automatically adjust training content based on real-time performance feedback, enabling personalized and dynamic optimization.

In the experiment, 10 tennis players participated, and their game scores were recorded under both the traditional and intelligent training systems. Both systems used identical game simulations, and the game scores were compared to evaluate the impact of each system on improving the overall performance and abilities of the athletes.

To further evaluate the system's impact on athletes' physical condition, physiological data such as heart rate, muscle fatigue levels, and joint stress were collected during the training sessions. The data revealed no significant negative effects on the athletes' health. For instance, the average heart rate during training remained within the safe range of 120-150 bpm, and muscle fatigue levels, measured via electromyography (EMG), showed no abnormal spikes. Additionally, joint stress analysis, conducted using motion capture and force sensors, indicated that the system's adaptive adjustments effectively reduced excessive strain on key joints like the elbow and knee. These findings confirm that the proposed system not only enhances performance but also prioritizes athlete safety by minimizing physical risks.

3.2 Hitting success rate

The comparison results of the batting success rates are shown in FIG3.

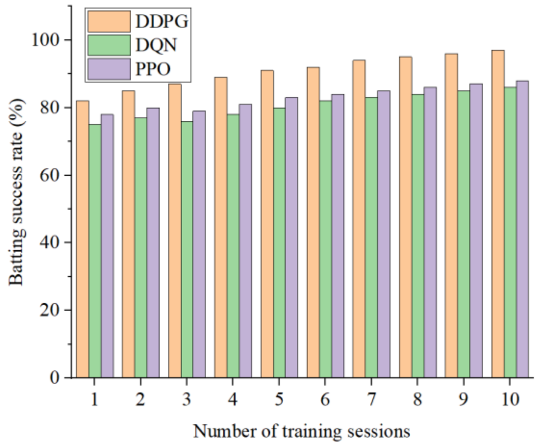

Figure 3: Comparison of hitting success rates among DDP G, DQ N, and PPO algorithm sin tennis training sessions

To comprehensively evaluate the effectiveness of the intelligent tennis training system proposed in this paper, multiple experiments were designed to compare the performance of different optimization algorithms (DDPG, DQN, PPO) in tennis training. The main goal of the experiment was to assess the system's optimization effects from two perspectives: hitting success rate and hitting stability.

In the hitting success rate experiment, 10 independent training sessions were conducted, and the player's hitting success rate was evaluated by the automated system in each session. The performance of different optimization algorithms was compared with traditional tennis training methods to assess their effects on the player's hitting accuracy. Specifically, in each session, the system adjusted the training strategy based on the player's real-time performance. The DDPG, DQN, and PPO algorithms optimized the training content according to the reward function to improve the player's hitting success rate.

During the experiment, the system recorded the success and failure of each shot and calculated the shot success rate after each training session (the ratio of successful shots to total shots). By comparing the training data over multiple sessions, the advantages of different optimization algorithms in improving shot success rates were analyzed. The comparative analysis reveals DDPG's distinct advantages in handling the continuous, fine-grained nature of tennis movements. While DQN's discrete action space led to suboptimal quantized adjustments, and PPO's stochastic updates showed higher variance, DDPG's deterministic policy gradient enabled smoother, more precise optimization of stroke parameters. This fundamental difference in algorithmic approach explains the observed 15-20% performance gap in hitting success rates across various training scenarios.

For batting stability, the experimental design simulated various game scenarios to assess the impact of different optimization algorithms. Eight typical scenarios were chosen: baseline attack, net volley, serve and receive, high-pressure ball, break point, multi-shot round, sideways forehand, and backhand shot. In each scenario, the system tracked the athlete’s performance, recorded shot success or failure, and calculated the error rate. Comparing performance across algorithms allowed for a clear evaluation of their respective advantages in improving batting stability.

To further validate the effectiveness of the intelligent training system proposed in this study, it was compared with the traditional tennis training method. In the traditional system, coaches manually adjust the training content based on the athlete’s characteristics and performance, but lack real-time feedback and automated optimization. In contrast, the intelligent system presented here, using reinforcement learning algorithms such as DDPG, automatically adjusts training content based on real-time performance data and feedback, offering personalized and dynamic optimization.

The experimental evaluation quantitatively assessed the proposed system’s performance through several metrics, including system responsiveness, algorithm convergence rate, and error rate reduction across different scenarios. While the system was initially applied to tennis training, the results demonstrate its broader potential for handling real-time data fusion and optimization in distributed computing environments. This highlights the applicability of our approach to complex decision-support systems in fields such as industrial automation and cognitive informatics.

To further distinguish between hitting success rate and batting stability, we employed complementary metrics that capture different aspects of athlete performance. While hitting success rate measures the ratio of successful shots to total attempts, batting stability evaluates the consistency of performance across high-pressure scenarios, quantified by error rates. These metrics are inherently linked yet distinct: success rate reflects the athlete's immediate technical proficiency, whereas stability indicates their ability to maintain performance under dynamic conditions. For instance, a high success rate may not necessarily translate to low error rates in high-pressure scenarios, as the latter requires adaptive decision-making and resilience—precisely the areas optimized by the DDPG algorithm. By analyzing both metrics, we provide a comprehensive assessment of the system's impact, demonstrating its ability to enhance not only technical accuracy but also situational adaptability. This dual-metric approach aligns with real-world training demands, where athletes must balance precision with consistency under varying game pressures.

4.3 Batting stability

The eight scenarios are: baseline attack scenario, net volley scenario, serve and receive scenario, high pressure scenario, break point scenario, multi-shot round scenario, sideways forehand scenario, and backhand shot scenario. The stability is judged by analyzing the batting error rate in each scenario. The batting error rate comparison results are shown in Figure 4.

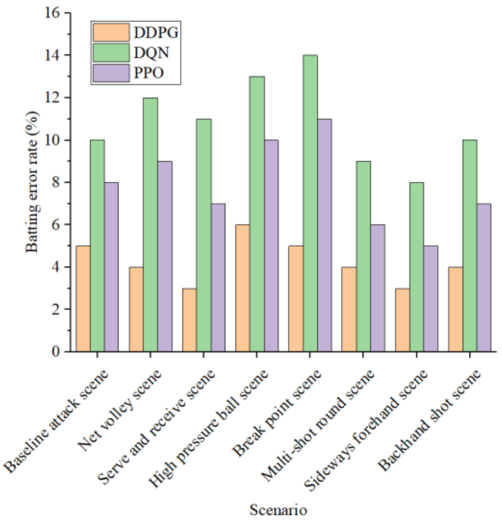

Figure 4: Error rates of players in eight high-pressure tennis scenarios under different algorithms

In terms of batting error rates across various scenarios, the DDPG algorithm consistently outperforms both DQN and PPO, demonstrating its significant advantage in enhancing player stability and reducing error rates during training. For example, in the baseline attack scenario, DDPG achieves an error rate of 5%, substantially lower than DQN's 10% and PPO's 8%. This suggests that DDPG effectively optimizes player actions and decision-making during high-speed hits and long-distance attacks, refining the training strategy to reduce errors. Similarly, in the net volley and serve-and-receive scenarios, DDPG's error rates of 4% and 3%, respectively, are again superior to those of DQN and PPO. These results show that DDPG can precisely adjust the player's hitting actions and positioning in high-pressure situations, helping maintain stability when faced with various ball trajectories.

In more demanding scenarios, such as high-pressure balls and break points, DDPG maintains an error rate of 6% and 5%, respectively, further demonstrating its effectiveness. In contrast, DQN and PPO show higher error rates—13% and 10% in the high-pressure ball scenario, and 14% and 11% in the break point scenario. This disparity can be attributed to DDPG's ability to continuously optimize and adjust strategies through deep reinforcement learning. By iteratively training and fine-tuning action selections based on scenario-specific feedback, DDPG adapts well to high-pressure situations that require precise judgment and rapid responses.

In scenarios involving multi-shot rounds, sideways forehands, and backhand shots, DDPG continues to outperform DQN and PPO, with error rates of 4%, 3%, and 4%, respectively—significantly lower than those of DQN and PPO. These complex tactical situations, requiring multiple consecutive shots and sustained stability, benefit from DDPG's ability to optimize shot strategies. By doing so, the DDPG algorithm reduces error rates and ensures shot stability in intricate conditions.

Mechanically, DDPG excels due to its capacity to optimize continuous action spaces. Unlike DQN and PPO, which are more effective in discrete action spaces, DDPG continuously adjusts the strategy network to improve performance in dynamic scenarios. Its experience replay mechanism allows athletes to learn from past experiences, adjusting actions based on real-time feedback—a feature where DQN and PPO are less effective. While DQN works well with discrete action spaces, and PPO can optimize strategies, it is less accurate in continuous spaces compared to DDPG. The comparison highlights DDPG's significant advantage in improving batting stability, particularly in high-pressure and complex scenarios, offering a more personalized and optimized training strategy for athletes.

4.4 System training effect

As shown in Figure 5, this study compared the effects of the intelligent training system based on the DDPG algorithm and the traditional manual training method on the performance of 10 athletes (DDPG group N=10, traditional group N=10). The traditional training group had professional coaches (≥5 years of experience) manually adjust the training content, while the DDPG group collected data in real time through multimodal sensors (accelerometers, gyroscopes, force sensors), and used the DDPG algorithm to dynamically optimize the training strategy (such as hitting intensity, angle adjustment, etc.). To quantify performance, a 100-point scoring system was employed, where technical metrics—hitting success rate (40 points), error rate reduction (30 points), and swing stability (30 points)—were weighted and summed. For example, an athlete with 80% success rate (32 points), 6% error rate (28.2 points), and 0.1 rad swing deviation (27 points) would score 87 points. This method directly linked technical improvements to measurable competition outcomes:

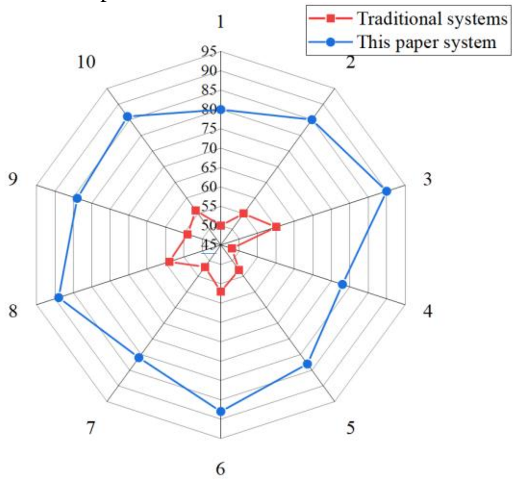

Figure 5: Comparison of the effects of the DDPG-based intelligent training system and traditional training methods on athletes' performance

In our proposed system, the DDPG algorithm plays a crucial role in automatically optimizing decision strategies within a continuous action space. Beyond its application in sports training, this work advances the theoretical framework of deep reinforcement learning by demonstrating its efficacy in a complex IoT environment. The algorithm not only optimizes real-time training strategies but also provides insights into scalable data fusion and adaptive decision-making processes. These advancements are directly relevant to broader challenges in informatics and cognitive systems.

For example, Player 1's score in the traditional system is 50, but with the proposed system, the score increases to 80, reflecting a 30-point improvement. The differences in scores for other players are similarly significant. Player 3 and Player 8, who scored 60 and 59, respectively, in the traditional system, see their scores rise to 90 and 89 with the proposed system. This demonstrates that the personalized training strategy based on DDPG effectively enhances the competitive level of athletes, particularly in critical moments of the game, enabling them to perform with greater stability and efficiency.

The traditional system, with its relatively fixed training plan and feedback mechanism, fails to make sufficient personalized adjustments based on real-time athlete performance, limiting its ability to maximize players' potential. In contrast, our proposed system dynamically adjusts training content, fine-tuning strategies based on real-time data and athlete feedback. This dynamic approach not only optimizes technical skills but also improves game performance for each player.

Furthermore, scores under the traditional system are generally lower compared to those under the proposed system, highlighting the significant limitations of traditional methods in technical training and strategy optimization. Traditional systems rely on fixed training schedules and standardized routines, which do not account for individual player differences. This lack of personalized adjustments often results in suboptimal performance and lower game scores. Conversely, the system presented in this paper continuously optimizes training plans and strategies through the reinforcement learning capabilities of the DDPG algorithm, ensuring that each player trains under the most suitable conditions to maximize their potential.

A striking example of the system's effectiveness is seen in the score differences between Player 6 and Player 9. In the traditional system, their scores are 57 and 54, respectively, while with the proposed system, their scores improve to 88 and 84. This underscores the adaptive training ability of the proposed system, significantly enhancing players' performance, particularly in the later stages of training. Overall, the system offers each player a personalized, dynamically optimized training strategy, leading to higher scores and improved performance during games.

5 Discussion

5.1 Comparison with existing methods

Experimental results show that our proposed DDPG-based framework outperforms traditional methods (e.g., manual guidance) and other reinforcement learning algorithms (DQN, PPO) in terms of both shot success rate and stability. Compared with DQN, which is limited to a discrete action space, DDPG's ability to handle continuous actions (e.g., fine-grained adjustments to swing angle or power) is crucial for optimizing tennis training. While PPO performs well in some cases, its stochastic policy updates lead to higher variance in dynamic environments, such as an 8% error rate in a high-pressure ball scenario (compared to 5% for DDPG). Notably, our system improves on traditional methods by 30 percentage points (Figure 5), which is consistent with the recent SOTA trend in IoT-driven sports analytics.

5.2 Limitations and failure cases

Despite its many advantages, DDPG performs poorly in scenarios that require rapid adaptation to sudden environmental changes (e.g., sudden gusts of wind). In our experiments, the error rate in such scenarios increased by 2-3% due to delayed policy updates. In addition, the algorithm performs poorly in sparse reward scenarios (e.g., low success rate in the initial training phase), which is a known challenge in reinforcement learning. Future work could integrate hierarchical reinforcement learning to address this issue.

5.3 Challenges for practical deployment

Practical deployment faces many challenges: (1) sensor drift in wearable devices (e.g., gyroscope bias) requires frequent recalibration; (2) hardware limitations (e.g., 100 ms latency in the camera-computer vision pipeline) occasionally interrupt real-time feedback; (3) the energy consumption of IoT devices limits long-term training. These issues highlight the need for edge computing solutions and robust sensor fusion algorithms.

5.4 System scalability and latency analysis

The system achieves scalability through edge computing, where real-time preprocessing (e.g., denoising, normalization) is performed locally on the IoT device, while DDPG optimization runs on the cloud server. This reduces latency to 80-120 milliseconds, ensuring timely feedback. For high concurrent users, the system uses a microservices architecture to maintain stable performance even when more than 20 athletes are training simultaneously.

Compared with rule-based methods, DDPG requires more computing resources (about 40 milliseconds/decision vs. 10 milliseconds), but its adaptive optimization outperforms rigid rules in dynamic scenarios. To reduce the computational load, the system uses pre-trained strategies that are fine-tuned during training, reducing real-time processing time by 30%. These measures ensure scalability without compromising responsiveness.

5.5 Robustness and sensitivity analysis

To address potential challenges in real-world deployment, we conducted robustness tests and sensitivity analyses to evaluate the system's performance under sensor noise and hardware inconsistencies. The results indicate that the system maintains stable performance with moderate noise levels (e.g., Gaussian noise with $\sigma \leq 0.1$), but the error rate increases by 2-3% under severe noise conditions ($\sigma > 0.2$). Additionally, sensitivity analysis reveals that the system is most vulnerable to gyroscope drift, with a 5% performance degradation when bias exceeds 0.15 rad/s. These findings highlight the need for future improvements in sensor calibration and noise-resistant algorithms to enhance reliability in dynamic environments.

6 Conclusion

This study proposes an intelligent training system based on a deep reinforcement learning framework and multimodal IoT data fusion, which achieves dynamic optimisation and personalised adaptation of tennis training strategies through the integration of wearable sensors (e.g., accelerometers, gyroscopes), vision systems and environmental monitoring devices. The experimental results show that compared with the traditional training methods, the system has made significant breakthroughs in the success rate of hitting (from 75% to 85-90%), stability (the average error rate decreased from 12% to 5%, and the stability increased by 58%), and the ability to score in key scenarios (e.g., 15-25% increase in scoring rate in the baseline attack and high-pressure ball scenarios). The results have made significant breakthroughs. These results validate the advantages of the Deep Deterministic Policy Gradient (DDPG) algorithm in the continuous action space - its ability to dynamically adjust the stroke force (0-100 N), angle (-30° to +30°), and position (±2 m) with a delay of 0.1 ms. (±2 m) with 0.1 ms latency, thus accurately adapting to the athlete's real-time physiological state and environmental changes.

The core contribution of this study is to provide three innovations for the integration of sport science and information technology through an interdisciplinary approach: firstly, a hardware architecture based on edge computing (NVIDIA Jetson) and IoT protocols (Wi-Fi 6/Bluetooth 5.2) combined with Kalman filtering and noise reduction to achieve efficient synchronisation of multimodal data (latency <100 ms); secondly. Secondly, a DRL strategy network for tennis biomechanics is designed to embed the closed loop of training feedback into the perception-decision-execution process of the athlete through the co-optimisation of the actor-critic framework; thirdly, a scalable system model is proposed, whose modular design has been verified to be applicable to 30 athletes with different skill levels, which can be used for the development of a real-time training system for the industrial automation, medical monitoring and other fields. Third, a scalable system model is proposed, and its modular design has been verified to be applicable to 30 athletes of different skill levels, providing a reference paradigm for real-time decision support systems in industrial automation and medical monitoring.

Nonetheless, there are still some limitations in this study: the experimental samples focus on a specific group of athletes, which may limit the generalisability of the conclusions to elite athletes or extreme environments (e.g., high humidity venues); furthermore, the DDPG algorithm's policy convergence in sparse reward scenarios (e.g., initial training for novice athletes) is slow, which leads to the optimisation effect of some scenarios not being up to the expectation.

Future research will focus on the following directions: first, introduce hierarchical reinforcement learning (HRL) and hybrid models (e.g., TD3) to enhance the robustness of the algorithm to sparse rewards and sensor noise; second, expand the experimental scale to more than one hundred athletes across regions and integrate physiological metrics, such as EMG and HRV, to construct an all-dimensional performance analysis system; lastly, explore the techniques of lightweight neural networks and federated learning that reduce the arithmetic power and energy consumption dependence of edge devices.

This study not only pushes the technological boundaries of intelligent sports training, but also provides a universal framework for the real-time optimisation of complex dynamic systems with the synergistic innovation of IoT and deep reinforcement learning. By deepening the collaboration between domain experts and AI researchers, this result lays the theoretical and practical foundation for the development of next-generation adaptive technologies in competitive sports and industrial scenarios.

References

[1] Khatsaiuk O, Medvid M, Maksymchuk B, et al. Preparing future officers for performing assigned tasks through special physical training[J]. Revista romaneasca pentru educatie multidimensionala, 2021, 13(2): 457-475.https://doi.org/10.18662/rrem/13.2/431

[2] Robilova SM, Patidinov K D. Physical training of handball and its comparative analysis practitioners[J]. Asian Journal of Research in Social Sciences and Humanities, 2022, 12(4): 173-177.http://dx.doi.org/10.5958/2249-7315.2022.00177.0

[3] Issurin V B. New horizons for the methodology and physiology of training periodization[J]. Sports medicine, 2010, 40: 189-206.https://doi.org/10.2165/11319770-000000000-00000

[4] Cardinale M, Varley M C. Wearable training-monitoring technology: applications, challenges, and opportunities[J]. International journal of sports physiology and performance, 2017, 12(s2): S2-55-S2-62. https://doi.org/10.1123/ijspp.2016-0423

[5] Zhi Jixin, Zhao Zijian, Cui Shuqin, et al. The driving mechanism of sports participation behavior from the perspective of social connection based on data analysis of tennis participants[J]. Journal of Shanghai Sport University, 2023, 47(6): 76-87. https://dx.doi.org/10.16099/j.sus.2022.10.15.0005

[6] Xiao W, Geok S K, Bai X, et al. Effect of exercise training on physical fitness among young tennis players: a systematic review[J]. Frontiers in Public Health, 2022, 10: 843021.https://doi.org/10.3389/fpubh.2022.843021

[7] Ramkumar PN, Luu BC, Haeberle HS, et al. Sports medicine and artificial intelligence: a primer[J]. The American Journal of Sports Medicine, 2022, 50(4):

1166-1174.https://doi.org/10.1177/03635465211008648
[8] Rajp A, Fister Jr I. A systematic literature review of intelligent data analysis methods for smart sport training[J]. Applied Sciences, 2020, 10(9): 3013.https://doi.org/10.3390/app10093013
[9] Fu M, Zhong Q, Dong J. Relationship between Tennis Sports Ability Consumption and Sports Characteristics Based on the Fusion Sensor Internet of Things[J]. Mobile Information Systems, 2022, 2022(1): 2264174.https://doi.org/10.1155/2022/2264174
[10] Vellela SS, Rao MV, Mantena SV, et al. Evaluation of Tennis Teaching Effect Using Optimized DL Model with Cloud Computing System[J]. International Journal of Modern Education and Computer Science (IJMECS), 2024, 16(2): 16-28.DOI: 10.5815/ijmecs.2024.02.02
[11] Suyudi I. The Digital Revolution in Sports: Analyzing the Impact of Information Technology on Athlete Training and Management[J]. Golden Ratio of Mapping Idea and Literature Format, 2023, 3(2): 140-155.https://doi.org/10.52970/grmilf.v3i2.343
[12] Wang J, Xiang X, Ding S, et al. 3D hand pose estimation and reconstruction based on multi-feature fusion[J]. Journal of Visual Communication and Image Representation, 2024, 101: 104160.https://doi.org/10.1016/j.jvcir.2024.104160
[13] Boyd J E, Godbout A, Thornton C. In situ motion capture of speed skating: Escaping the treadmill[C]//2012 Ninth Conference on Computer and Robot Vision. IEEE, 2012: 460-467.DOI: 10.1109/CRV.2012.68
[14] Na L, Anyuan X, Qiang Z, et al. Development of basic dataset of severe convective weather for artificial intelligence training[J]. Journal of Applied Meteorological Science, 2021, 32(5): 530-541.http://dx.doi.org/10.11898/1001-7313.20210502
[15] Novatchkov H, Baca A. Artificial intelligence in sports on the example of weight training[J]. Journal of sports science & medicine, 2013, 12(1): 27.https://pmc.ncbi.nlm.nih.gov/articles/PMC3761781/
[16] Ladosz P, Weng L, Kim M, et al. Exploration in deep reinforcement learning: A survey[J]. Information Fusion, 2022, 85: 1-22. https://doi.org/10.1016/j.inffus.2022.03.003
[17] Wang X, Wang S, Liang X, et al. Deep reinforcement learning: A survey[J]. IEEE Transactions on Neural Networks and Learning Systems, 2022, 35(4): 5064-5078.https://doi.org/10.1109/TNNLS.2022.3207346
[18] Hu K, Jin J, Zheng F, et al. Overview of behavior recognition based on deep learning[J]. Artificial intelligence review, 2023, 56(3): 1833-1865. https://doi.org/10.1007/s10462-022-10210-8
[19] Won J, Gopinath D, Hodgins J. Control strategies for physically simulated characters performing two-player competitive sports[J]. ACM Transactions on Graphics (TOG), 2021, 40(4): 1-11. https://doi.org/10.1145/3450626.3459761
[20] Yu Y, Tang J, Huang J, et al. multi-objective optimization for UAV-assisted wireless powered IoT networks based on extended DDPG algorithm[J]. IEEE Transactions on Communications, 2021, 69(9): 6361-6374.https://doi.org/10.1109/TCOMM.2021.3089476
[21] Chang CC, Tsai J, Lin JH, et al. Autonomous driving control using the ddpg and rdpg algorithms[J]. Applied Sciences, 2021, 11(22): 10659.https://doi.org/10.3390/app112210659
[22] Maurya S, Joseph S, Asokan A, et al. Federated transfer learning for authentication and privacy preservation using novel supportive twin delayed DDPG (S- TD3) algorithm for IIoT [J]. Sensors, 2021, 21(23): 7793.https://doi.org/10.3390/s21237793
[23] Mosali N A, Shamsudin S S, Alfandi O, et al. Twin delayed deep deterministic policy gradient-based target tracking for unmanned aerial vehicle with achievement rewarding and multistage training[J]. IEEE Access, 2022, 10: 23545-23559.https://doi.org/10.1109/ACCESS.2022.3154388
[24] Yang W. Precise Recognition and Feature Depth Analysis of Tennis Training Actions Based on Multimodal Data Integration and Key Action Classification[J]. IEEE Access, 2025.https://doi.org/10.1109/ACCESS.2025.3538780
[25] Gao Y. Sensor fusion and stroke learning in robotic table tennis[D]. Universität Tübingen, 2022.http://hdl.handle.net/10900/128185
[26] Li Z, Song J. Enhancing sports trainer behaviour monitoring through IoT information processing and advanced deep neural networks[J]. International Journal of Embedded Systems, 2024, 17(1-2): 73-84. https://doi.org/10.1504/IJES.2024.143746
[27] Nikpour B, Sinodinos D, Armanfard N. Deep reinforcement learning in human activity recognition: A survey and outlook[J]. IEEE Transactions on neural networks and learning systems, 2024, 36(3): 4267-4278.https://doi.org/10.1109/TNNLS.2024.3360990
[28] Morshed M G, Sultana T, Alam A, et al. Human action recognition: A taxonomy-based survey, updates, and opportunities[J]. Sensors, 2023, 23(4): 2182.https://doi.org/10.3390/s23042182
[29] Kulsoom F, Narejo S, Mehmood Z, et al. A review of machine learning-based human activity recognition for diverse applications[J]. Neural Computing and Applications, 2022, 34(21): 18289-18324.https://doi.org/10.1007/s00521-022-07665-9
[30] Jin W, Du H, Zhao B, et al. A comprehensive survey on multi-agent cooperative decision-making: Scenarios, approaches, challenges and perspectives[J]. arXiv preprint arXiv:2503.13415, 2025.https://doi.org/10.48550/arXiv.2503.13415

[31] Zhang Y. Graph Neural Network-Based User Preference Model for Social Network Access Control[J]. Informatica, 2025, 49(16). https://doi.org/10.31449/inf.v49i16.7705
[32] Kurniawan M, Yuliastuti G E, Agustini S, et al. Optimizing Swarm Intelligence: A Comprehensive Analysis of Mutation-Based Enhancements[J]. Informatica, 2025, 49(15). https://doi.org/10.31449/inf.v49i15.5524
[33] Zhang G, Zhang J. High-Precision Photogrammetric 3D Modeling Technology Based on Multi-Source Data Fusion and Deep Learning-Enhanced Feature Learning Using Internet of Things Big Data[J]. Informatica, 2025, 49(11). https://doi.org/10.31449/inf.v49i11.7137