

Design of a Visual Feedback System with Hitting Timing Recognition for Table Tennis Service Return Training

Rui Cao*, Zhentao Xia, Xiaomiao Hu

Tianjin University of Science and Technology, Tianjin, China

**Corresponding Author*

Keywords: Computer Vision; Table Tennis Training; Hitting Timing Recognition; Real-time Visual Feedback; Intelligent Sports

Abstract: To address the difficulties in hitting timing judgment, ambiguous conventional verbal feedback, and low training efficiency in table tennis service reception training, this paper develops a hitting timing recognition and real-time visual feedback system. The system adopts a high-speed camera to capture ball motion trajectories, and constructs target detection and hitting recognition algorithms based on OpenCV. It detects the hitting moment through inter-frame coordinate changes of the ball center and divides motion stages by combining the ball's vertical motion trend and height proportion. A real-time visual feedback mechanism is established to provide dynamic visual prompts for hitting timing and automatically record training data. Three comparative experimental groups of no feedback, verbal feedback and visual feedback are set for validation. The experimental results show that the visual feedback mode increases service reception accuracy by 43.1%, which is significantly superior to traditional training modes. The proposed system can effectively enhance beginners' timing judgment ability, reduce cognitive load during training, and provide a reference for the design of intelligent table tennis training systems and the application of computer vision in intelligent sports.

1. Introduction

Table tennis demands rapid responses under extreme time pressure, with ball speeds reaching 20–30 m/s and athletes facing diverse spin, speed, and placement variations [1-2]. Service reception serves as a key link connecting service and rally, and directly determines the initiative of competition rounds. Athletes need to quickly judge the spin, speed, landing point and trajectory of the incoming ball within a very limited time and select a reasonable hitting timing to complete the return [3]. Hitting timing directly affects return quality, stability and subsequent technical connection, which is the core focus of table tennis service reception training [4].

Current service reception training mainly relies on coaches' verbal guidance and repetitive empirical training, which suffers from abstract feedback, insufficient quantification and obvious dependence on coaches' experience [5]. Most beginners lack mature perception of hitting timing and struggle to understand abstract professional concepts such as hitting at the highest point and pressing forward at the rising stage, resulting in a long training cycle and low efficiency of movement correction. Therefore, developing an intelligent training system that can identify hitting timing in real

time and provide visual feedback is of great practical significance for improving service reception training efficiency [6-8].

With the continuous development of computer vision, inertial sensing and intelligent interaction technology, sports auxiliary training has become a research hotspot in the field of intelligent sports. Existing studies mainly focus on motion posture recognition, racket trajectory analysis and hitting action evaluation. Erwin Wu et al. [9] adopted virtual reality for table tennis spin perception training and optimized beginners' spin judgment ability through temporal and velocity distortion. Yiming Guo et al.[10] analyzed upper limb force characteristics based on the neuromusculoskeletal model. Zhang H et al. [11] realized racket trajectory recognition using inertial data from smart watches. Guo Tian et al.[12] completed hitting position detection via embedded piezoelectric sensors. Li et al.[13] achieved real-time accurate localization of table tennis ball and stroke effectiveness evaluation based on computer vision and deep learning. Wang et al.[14] proposed a table tennis trajectory tracking method by fusing YOLOv8 detection and binocular vision system. Chen et al.[15] designed a vision-based table tennis action recognition model using TL-CStrans Net and CS-Transformer mechanism. Although existing achievements have been made in motion behavior recognition, most studies focus on hitting movements, racket postures and force rules, while real-time recognition and visual feedback of hitting timing are insufficient. Meanwhile, most systems have cumbersome equipment deployment, unintuitive feedback forms and lack of real-time visual guidance, which limits their application in actual training scenarios.

To solve the above problems, this paper proposes a hitting timing recognition and visual feedback system for table tennis service reception training. Based on monocular vision, the system collects ball motion trajectories through a high-speed camera, detects hitting events according to ball coordinate changes, divides motion stages combined with vertical motion trend and height proportion, and constructs a real-time visual feedback mechanism to dynamically prompt hitting timing, helping trainees build timing perception and improve training efficiency.

This paper makes four main contributions. First, a hitting event detection method is developed based on sudden lateral displacement mutations between frames. Second, a timing division model is built by fusing vertical motion trend with normalized height ratio. Third, these algorithms are integrated into a real-time visual feedback system for training scenarios. Finally, the system's effectiveness is validated through controlled experiments with novice players. [16-21]

2. Overall System Architecture

2.1 System Composition

To realize real-time recognition and feedback of hitting timing in table tennis service reception training, this system constructs an intelligent training hardware platform based on monocular vision. The hardware consists of four modular units: image acquisition, data processing, display feedback, and power supply & structural support.

The image acquisition module adopts a high-speed industrial camera as the core sensor to capture continuous image sequences of fast-moving table tennis balls. Given that the ball speed can reach 20–30 m/s, a camera with a frame rate no lower than 120 FPS is selected to ensure sampling continuity and temporal accuracy. The camera is mounted on the side of the table with a fixed bracket to reduce errors caused by viewing angle deviation. The data processing module is composed of an embedded computing unit or an external host computer, undertaking core tasks such as video stream decoding, image preprocessing, ball detection, trajectory analysis and hitting timing evaluation. The display feedback module adopts an independent screen or external display device to output real-time visualized results, including trajectory curves, hitting moment markers and training status prompts.

2.2 Image Acquisition Scheme

The camera is placed 1.0–1.5 m away from the side of the table, with an shooting angle of 30° – 45° relative to the table plane, so as to acquire both horizontal and vertical motion information of the ball. To suppress interference from ambient light and complex backgrounds, fixed exposure parameters and white balance locking are adopted. A high shutter speed is applied to reduce motion blur and improve edge definition of high-speed targets.

The collected original video stream is transmitted to the processing module in the form of frame sequences. Each frame contains the spatial position of the ball at the corresponding moment. All image frames are numbered in chronological order to establish a discrete trajectory model, providing basic data support for subsequent hitting event detection.

2.3 Real-Time Feedback Architecture

To output recognition results instantly and provide training assistance, a closed-loop real-time feedback architecture of “visual perception – algorithm analysis – result feedback” is established. It supports low-latency processing and dynamic information updating during ball movement, enabling timely motion correction in training.

The overall feedback procedure is shown in the corresponding figure. Firstly, continuous frame data is obtained by the image acquisition module and then sent for image preprocessing and target detection. In the algorithm analysis layer, hitting events are detected via coordinate variation of the ball center, and the whole motion process is divided into different timing stages according to trajectory characteristics.

In the feedback layer, recognition results are converted into visual output, including three main forms:

- (1) Hitting moment marker: pinpoints the contact frame on the time axis or trajectory curve.
- (2) Timing stage indicator: labels the hit as rising phase, peak phase, or falling phase.
- (3) Deviation gauge: shows the offset between actual hitting moment and the optimal hitting window.

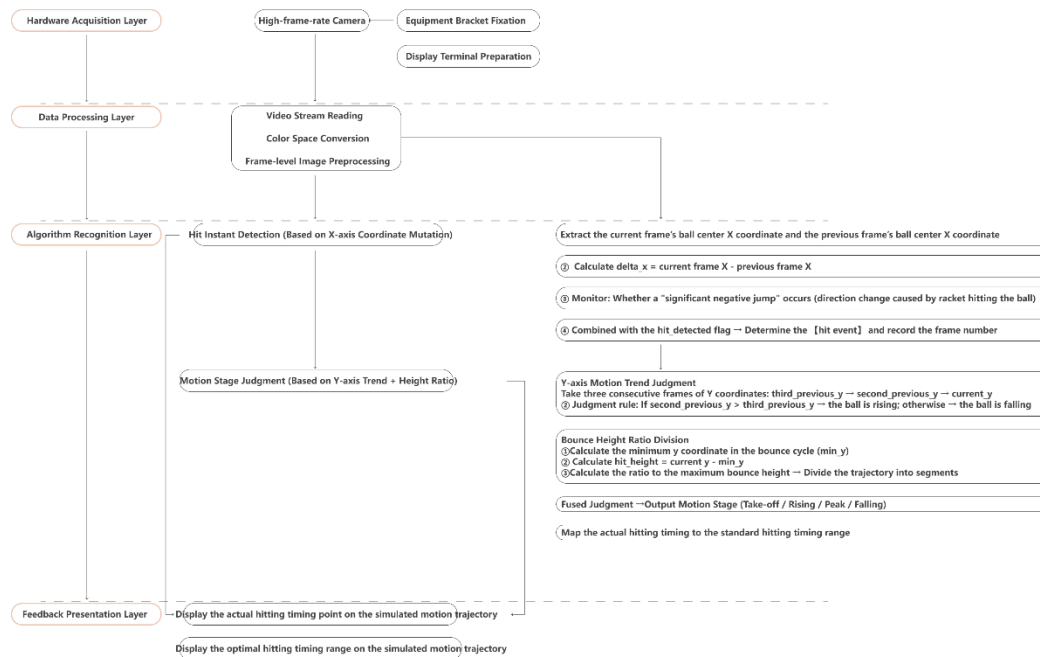


Figure 1. Overall algorithm flowchart of the proposed system

The overall workflow of the proposed system is illustrated in Figure 1, covering the complete closed loop from visual acquisition to real-time feedback.

3. Hitting Timing Recognition Method

3.1 Table Tennis Target Detection

Table tennis target detection is the fundamental prerequisite of the hitting timing recognition system, which aims to steadily acquire the spatial position coordinates of the ball in continuous video frames and provide basic input data for subsequent trajectory analysis and hitting event identification. Characterized by small size, fast motion speed and inevitable motion blurring during high-speed movement, table tennis targets are difficult to detect stably. Traditional detection methods based on complex deep learning models are limited in real-time performance and practical deployment cost. Accordingly, this paper adopts a lightweight visual detection approach based on color feature analysis and morphological constraints, which can achieve reliable target extraction while ensuring high real-time performance.

The input image is firstly converted into the HSV color space, and the initial candidate regions are segmented by setting a reasonable color threshold range corresponding to table tennis. Morphological opening and closing operations are subsequently adopted to eliminate background noise and interference. Connected regions in the image are further screened according to preset area and circularity constraints to retain valid ball targets. The two-dimensional coordinates of the table tennis ball in each frame are calculated by extracting the centroid of the qualified target region. The obtained coordinate sequence can be used as the fundamental data for subsequent trajectory modeling and analysis.

3.2 Hitting Event Recognition

Hitting event recognition focuses on accurately locating the moment when the racket makes contact with the table tennis ball within frame sequences.

Under normal flight conditions, the lateral displacement of the table tennis ball changes continuously and smoothly. When hitting occurs, the external force exerted by the racket will cause a sharp mutation in the horizontal motion state of the ball. The change in the ball's X-axis coordinate between consecutive frames is calculated as shown in Equation (1):

$$\Delta x_t = x_t - x_{t-1} \quad (1)$$

where x_t denotes the ball center's X-coordinate at the current frame, and x_{t-1} denotes the X-coordinate at the previous frame. Hitting behavior can be preliminarily judged when the variation of lateral displacement exceeds the empirical threshold. To suppress false detection caused by background occlusion or detection jitter, temporal consistency constraint is introduced to verify the continuity of candidate hitting states in adjacent frames. Only the frames that satisfy both the displacement mutation rule and temporal consistency constraint are defined as the effective hitting moment, and the corresponding frame index is recorded for subsequent timing analysis.

3.3 Motion Stage Division

After completing hitting event detection, it is necessary to divide the complete bouncing process of the table tennis ball into refined motion stages, so as to realize the semantic description of hitting timing. The stage division model is established based on the vertical motion trend and height variation characteristics of the ball.

The motion trend judgment rule based on three consecutive frames is defined as shown in Equation (2):

$$Trend = \begin{cases} Rising, & y_{t-2} < y_{t-1} < y_t \\ Falling, & y_{t-2} > y_{t-1} > y_t \end{cases} \quad (2)$$

where y_t , y_{t-1} , and y_{t-2} denote the vertical coordinates of the ball center at the current frame, previous frame, and frame before the previous one, respectively. By analyzing the positive and negative changes of vertical velocity, the rising and falling trends of the ball can be distinguished. Combined with the maximum bouncing height in the whole motion process, the normalized height ratio is defined to describe the relative position of the ball in the bouncing trajectory. The relative height and height ratio of the ball center at the current frame are calculated as shown in Equations (3) and (4):

$$h_t = y_t - y_{min} \quad (3)$$

$$Ratio_t = \frac{h_t}{y_{max} - y_{min}} \quad (4)$$

where y_{min} and y_{max} denote the minimum and maximum vertical coordinates of the ball center in the bounce cycle, respectively. Integrating the vertical motion trend and normalized height information, the entire bouncing process is divided into five complete motion phases, including the take-off stage, early rising stage, peak stage, early falling stage and late falling stage. The bouncing process of the table tennis ball is divided into five consecutive phases, with its trajectory and characteristics visualized in Figure 2. In particular, the peak stage is defined as the stable interval where the vertical velocity of the ball approaches zero. The refined five-stage division provides a reliable theoretical basis for subsequent quantitative evaluation and semantic recognition of hitting timing.

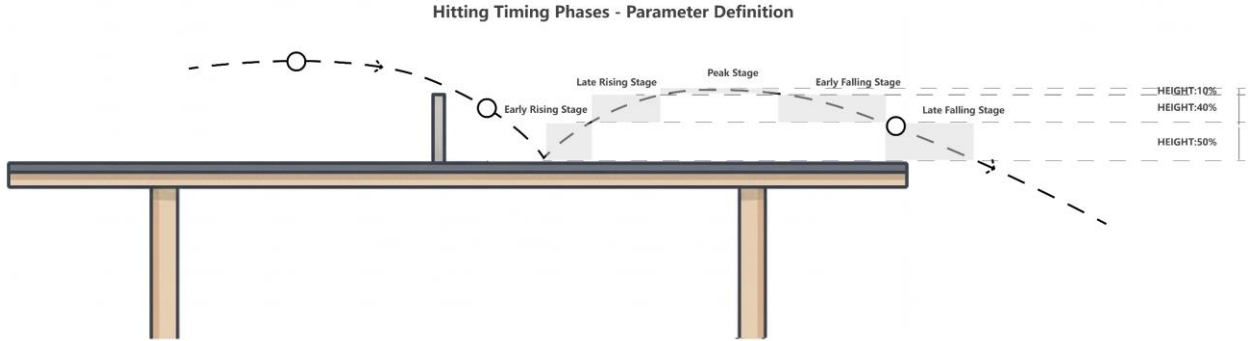


Figure 2. Schematic diagram of the five motion stages of a table tennis ball

3.4 Timing Classification Logic

Combined with the results of hitting event detection and motion stage division, this paper further constructs a hitting timing classification model to map actual hitting behavior into standardized technical semantic categories. The motion phase of the ball corresponding to the hitting moment is taken as the core basis for timing classification.

To improve the robustness of classification and avoid misjudgment caused by single-frame detection error, a time window constraint mechanism is adopted in this study. The final hitting timing category is determined by evaluating the motion stage stability within the adjacent time window. The standardized classification results are output in the form of unified labels, which can be directly

invoked by the subsequent real-time visual feedback module.

4. Training Feedback System Design

4.1 Real-Time Visual Feedback Mechanism

To realize the immediate application of hitting timing recognition results in the training process, a real-time feedback mechanism based on visual coding is constructed in this paper. The core goal of this mechanism is to transmit the system recognition results to trainees in a low cognitive load way without interfering with their normal hitting actions, thereby improving the efficiency of timing correction.

Recognition results are mapped into three visual information types: hitting moment marking locates key event frames on the trajectory diagram; timing stage prompt displays the current motion stage as rising, peak, or falling; deviation feedback quantifies the gap between actual and ideal hitting moments.

Let the ideal hitting timing interval be $T^* = [t_a, t_b]$ and the actual hitting moment be t_h , then the timing deviation is defined as $D = \min(|t_h - t_a|, |t_h - t_b|)$. The system divides the feedback into three levels according to the deviation value: the optimal hitting (green prompt) when $D \leq \epsilon_1$, the acceptable range (yellow prompt) when $\epsilon_1 < D \leq \epsilon_2$, and the deviated hitting (red prompt) when $D > \epsilon_2$. This grading mechanism realizes real-time quantitative feedback on hitting quality.

In addition, to avoid visual information overload, the system adopts a local dynamic highlighting strategy, which only displays feedback information within the key time window and keeps the interface simplified at other times, so as to ensure the concentration of visual attention during training.

4.2 Human-Computer Interaction Design

A low-interference human-computer interaction process is designed around four stages: preparation, execution, feedback acquisition, and correction. During preparation, users configure training mode, difficulty, and feedback intensity, after which camera calibration completes. During execution, ball trajectory is captured and hitting events trigger feedback automatically. During feedback acquisition, recognition results display timing stage, deviation level, and trajectory playback. During correction, trainees adjust subsequent strokes based on feedback, forming a closed "perception-feedback-correction" loop.

4.3 System Deployment Structure

The system supports deployment in three typical scenarios: standard gymnasium environment, teaching and training environment and temporary training environment. In the standard gymnasium environment, the system is fixedly installed on the side of the table, and the relative position between the camera and the table is ensured to be stable through the bracket structure; in the teaching and training environment, the system can be shared with multiple tables, and equipment rotation is realized through the quick disassembly and assembly structure; in the temporary training environment, the system supports portable deployment, and rapid construction is realized through mobile brackets and external power supply.

In addition, the system adopts a modular design at the hardware interface level, enabling fast connection and replacement between the visual acquisition unit, computing unit and display unit, thus improving the system expandability and maintenance efficiency.

5. Experimental Results and Analysis

5.1 Experimental Platform Establishment

To verify the effectiveness of the proposed hitting timing recognition and visual feedback system, an experimental platform for table tennis service reception training was constructed. The platform mainly consists of a high-speed visual acquisition module, a computational processing unit, an automatic serving device and a training display terminal.

All experiments were carried out in a standard indoor table tennis venue. The camera was fixed at approximately 1.2 meters beside the table, with the shooting angle maintained at about 35° relative to the table plane, so that the complete motion trajectory of the table tennis ball could be fully recorded. To eliminate environmental interference and ensure the reliability of experimental data, all tests were conducted under unified illumination conditions and fixed equipment parameters. The experimental platform configuration and typical training scenarios are illustrated in Figure 3.



(a) Typical training scenarios of different participants; (b) Overview of the experimental platform, including the high-frame-rate acquisition device and display terminal.

Figure 3. Experimental setup and training scenarios.

5.2 Comparative Experimental Design

A single-factor controlled experiment was designed to explore the influence of different feedback modes on the training performance of table tennis beginners. Taking feedback mode as the independent variable, three training conditions were set, including conventional training without feedback, manual verbal feedback, and real-time visual feedback provided by the proposed system. Hitting timing accuracy and training stability were adopted as core evaluation indicators to quantitatively compare the training improvement under different conditions.

A total of 15 beginners without professional table tennis training experience were recruited as participants and randomly divided into three groups with five members in each group. All participants received unified adaptive basic training before the formal experiment to ensure a consistent initial skill level. The overall experimental procedure included baseline pre-training test, grouped differential training, post-training performance evaluation, and experimental data statistical analysis. During the training process, the system automatically recorded the classification result of each hitting timing and quantitatively calculated the deviation between the actual hitting moment and the ideal time window, providing fundamental data support for subsequent comparative analysis.

5.3 System Recognition Performance Analysis

The system recognition performance was quantitatively evaluated from two aspects: hitting event detection and motion stage classification. The overall detection accuracy is calculated as shown in Equation (5):

$$Accuracy = \frac{TP + TN}{TP + TN + FP + FN} \times 100\% \quad (5)$$

where TP, TN, FP, and FN represent the number of true positives, true negatives, false positives, and false negatives, respectively. With manually labeled video frames as the ground truth, the system exhibited favorable detection stability and recognition accuracy. The proposed stage division method based on motion trajectory and height ratio achieved satisfactory engineering practicability and operational stability, with the overall average classification accuracy maintained above 92%.

The computational efficiency of the system also meets the requirement of real-time training. The average processing time per frame was controlled within 8–12 ms, corresponding to a stable operating frame rate of 80–120 FPS. The system can complete target detection, timing recognition and result output with extremely low latency, which is suitable for the application scenario of real-time auxiliary training in daily table tennis practice. The key performance indicators of the system in hitting event detection and motion stage classification are summarized in Table 1.

Table 1. System recognition performance indicators

Task	Value (%)
<i>Hitting event detection accuracy</i>	92.8
<i>False negative rate</i>	4.2
<i>False positive rate</i>	3.0
<i>Early rising stage classification accuracy</i>	91.5
<i>Late rising stage classification accuracy</i>	90.8
<i>Peak stage classification accuracy</i>	93.2
<i>Early falling stage classification accuracy</i>	91.0
<i>Late falling stage classification accuracy</i>	90.5

5.4 User Training Effect Analysis

The statistical results of hitting timing accuracy before and after training for the three groups under different feedback modes are presented in Table 2. Statistical comparison of hitting timing accuracy before and after training was conducted among the three groups. The visual feedback group achieved a performance improvement of 43.1% after training, which was remarkably higher than that of the verbal feedback group and the non-feedback group. The comparison of hitting timing accuracy under three different feedback modes is presented in Figure 4. The non-feedback group relied mainly on self-perception and trial-and-error adjustment, resulting in limited training improvement. In contrast, the presented visual feedback mode provides real-time intuitive quantitative prompts, achieving faster training convergence and higher movement stability.

In terms of the changing trend of average accuracy throughout the training process, the visual feedback group showed a smoother convergence curve, which effectively restrained the accumulation of incorrect movements and accelerated the learning process of technical skills. Statistical significance analysis indicated that the training effect of the visual feedback group was significantly different from the other two groups at the level of $(p < 0.05)$. The results further confirm that the designed real-time visual feedback system can effectively improve the training quality and learning efficiency of hitting timing for table tennis beginners.

Table 2. Comparison of training effects under different feedback modes

Group	Pre-training Accuracy (%)	Post-training Accuracy (%)	Improvement Rate (%)
<i>No-feedback Group</i>	42.5	48.3	13.6
<i>Verbal-feedback Group</i>	43	56.1	30.5
<i>Visual-feedback Group</i>	42.8	61.2	43.1

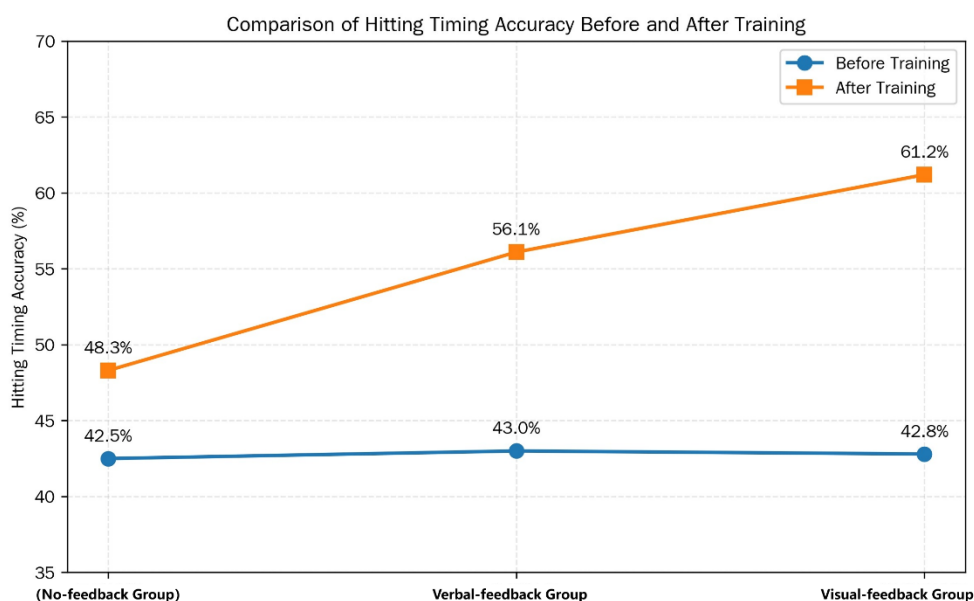


Figure 4. Comparison of hitting timing accuracy under different feedback modes

6. Conclusion

This paper proposed and implemented a hitting timing recognition and visual feedback system that addresses the quantitative evaluation gap in service reception training. Hitting events are detected from spatial coordinate variations of the ball, and timing stages are automatically classified by combining vertical motion trend with normalized height ratio.

A hitting event detection method was developed using sudden displacement changes between adjacent frames, improving recognition robustness. A timing division model fused motion direction with normalized height features to classify the rising, peak, and falling stages.

Comparative experiments across no-feedback, verbal feedback, and visual feedback modes demonstrated that visual feedback achieved a 43.1% improvement in hitting timing accuracy. The system maintained high recognition accuracy throughout, confirming its effectiveness and engineering feasibility.

In summary, the proposed hitting timing recognition and visual feedback system can effectively improve the ability of timing judgment in table tennis service reception training, and provides a feasible engineering implementation scheme for computer vision assisted sports training systems.

References

- [1] Dokic, K.; Mesic, T.; Martinovic, M. Table Tennis Forehand and Backhand Stroke Recognition Based on Neural Network. In *Proceedings of Advances in Computing and Data Sciences*; Springer: Singapore, 2020; pp. 24–35.
- [2] Lin, H.-I.; Yu, Z.; Huang, Y.-C. Ball Tracking and Trajectory Prediction for Table-Tennis Robots. *Sensors* 2020, 20, 333.

- [3] Rong, Z. Optimization of table tennis target detection algorithm guided by multi-scale feature fusion of deep learning. *Sci. Rep.* 2024, 14, 1401.
- [4] Wong, D.W.; Lee, W.C.; Lam, W.-K. Biomechanics of Table Tennis: A Systematic Scoping Review of Playing Levels and Maneuvers. *Appl. Sci.* 2020, 10, 5203.
- [5] Liebermann, D.G.; Katz, L.; Hughes, M.D.; Bartlett, R.M.; McClements, J.; Franks, I.M. Advances in the application of information technology to sport performance. *J. Sports Sci.* 2002, 20, 755–769.
- [6] Oagaz, H.; Schoun, B.; Choi, M.H. Performance Improvement and Skill Transfer in Table Tennis Through Training in Virtual Reality. *IEEE Trans. Vis. Comput. Graph.* 2022, 28, 4332–4343.
- [7] Messelodi, S.; Modena, C.M.; Ropele, V.; Marcon, S.; Sgrò, M. A Low-Cost Computer Vision System for Real-Time Tennis Analysis. In *Proceedings of Image Analysis and Processing-ICIAP 2019*; Springer: Cham, 2019; pp. 106–116.
- [8] Mateus, N.; Abade, E.; Coutinho, D.; Góñez, M.-Á.; Peñas, C.L.; Sampaio, J. Empowering the Sports Scientist with Artificial Intelligence in Training, Performance, and Health Management. *Sensors* 2025, 25, 139.
- [9] WU E, Piekenbrock M, Nakumura T, et al. Spinpong-virtual reality table tennis skill acquisition using visual, haptic and temporal cues[J]. *IEEE Transactions on Visualization and Computer Graphics*, 2021, 27(5): 2566-2576.
- [10] Guo Yiming, Sun Y, Ren Y, et al. Upper limb muscle force estimation during table tennis strokes[C]//2019 IEEE 16th International Conference on Wearable and Implantable Body Sensor Networks (BSN). IEEE, 2019: 1-4.
- [11] Tian GUO, Deng W, Gao Y, et al. Rich lamellar crystal baklava-structured PZT/PVDF piezoelectric sensor toward individual table tennis training[J]. *Nano Energy*, 2019, 59: 574-581.
- [12] Zhang Heng, Fu Z, Shu K I. Recognizing ping-pong motions using inertial data based on machine learning classification algorithms[J]. *IEEE Access*, 2019, 7: 167055-167064.
- [13] He, Z.; Yang, Z.; Xu, J.; Chen, H.; Li, X.; Wang, A.; Yang, J.; Chow, G.C.-C.; Chen, X. Real-Time Accurate Determination of Table Tennis Ball and Evaluation of Player Stroke Effectiveness with Computer Vision-Based Deep Learning. *Applied Sciences* 2025, 15, 5370.
- [14] Zhang, L.; Wang, H. A Method for Table Tennis Trajectory Tracking Combining YOLOv8 Object Detection Algorithm with Binocular Vision Acquisition System. *Sensors* 2024, 24, 3829.
- [15] Liu, Y.; Chen, X.; Li, J. TL-CStrans Net: A Vision Robot for Table Tennis Player Action Recognition Driven via CS-Transformer. *IEEE Access* 2024, 12, 89211–89224.
- [16] Capasa, L.; Zulauf, K.; Wagner, R. Virtual Reality Experience of Mega Sports Events: A Technology Acceptance Study. *Journal of Theoretical and Applied Electronic Commerce Research*, 2022, 17, 686–703.
- [17] Zadeh, A.; Taylor, D.; Bertson, M.; Tillman, T.; Nosoudi, N.; Bruce, S. Predicting Sports Injuries with Wearable Technology and Data Analysis. *Information Systems Frontiers*, 2021, 23, 1023–1037.
- [18] Giménez-Egido, J.M.; Ortega, E.; Verdu-Conesa, I.; Cejudo, A.; Torres-Luque, G. Using Smart Sensors to Monitor Physical Activity and Technical–Tactical Actions in Junior Tennis Players. *International Journal of Environmental Research and Public Health*, 2020, 17, 1068.
- [19] Michalski, S.C.; Szpak, A.; Saredakis, D.; Ross, T.J.; Billingham, M.; Loetscher, T. Getting your game on: Using virtual reality to improve real table tennis skills. *PLoS ONE*, 2019, 14, e0222351.
- [20] Brocherie, F.; Dinu, D. Biomechanical estimation of tennis serve using inertial sensors: A case study. *Frontiers in Sports and Active Living*, 2022, 4, 962941.
- [21] Wei, S.; Huang, P.; Li, R.; Liu, Z.; Zou, Y. Exploring the Application of Artificial Intelligence in Sports Training: A Case Study Approach. *Complexity*, 2021, 2021, 4658937.