ingegneria sismica
International Journal of Earthquake Engineering
Anno XLIII - Num. 1 - 2026

The Application and Challenge of Virtual Reality Technology in Physical Education

Wei Wang[1,*], *Fangde Sheng*[2] *and Yonggang Niu*[1,2,3]

[1] School of Physical Education of Anyang Normal University, henan, Anyang, China.
[2] Henan Province Graduate Joint Training Base. henan, Anyang,China.
[3] Anyang Key Laboratory of Fitness Training and Assessment, henan, Anyang, China.

SUMMARY: *In order to solve the problems of difficult to repeatedly present complex action demonstrations in physical education, classroom error correction relying on experience, and inaccurate feedback under individual differences of students, this paper studied the application of virtual reality technology in physical education. Combining computer vision, 3D skeleton reconstruction, action semantic recognition and virtual scene interaction methods, a technology chain composed of action acquisition, data processing, scene rendering and real-time feedback is constructed, and its teaching effect and practical challenges are analyzed. The results show that virtual reality technology can enhance the visualization degree of action demonstration, improve the timeliness of process feedback and the interactivity of training links. In the simulation test of self-built physical teaching action sample set, the accuracy of action recognition of the system reaches 95.3%, and it still maintains 88.7% in complex classroom scenes, showing a certain application potential. However, the system is still facing pressure on stability and adaptability in terms of multiple people in the same scene, rapid direction change, device wearing comfort and classroom deployment cost. The research believes that virtual reality technology provides a new digital support path for physical education, and the continuous release of its application value still depends on the further improvement of multi-source perception fusion, lightweight computing and teaching content modularization.*

KEYWORDS: *Virtual reality technology; Physical education; Action recognition; Human-computer interaction*

1 Introduction

Physical education has multiple attributes such as knowledge imparts, movement acquisition, skill training and situational experience. Its teaching effectiveness depends not only on whether the teacher's explanation is clear, but also on the visualization degree of movement demonstration, the immediacy of students' participation and the pertinence of training feedback. For a long time, the school physical education classroom mainly relies on teachers' oral explanation, on-site demonstration and repeated drill to complete the skill teaching. This method has the advantages of intuitionistic and convenient in basic movement training. However, when the teaching content involves complex technical movements, continuous tactical cooperation or special training with certain risks, the limitations of traditional methods will gradually appear. On the one hand, the classroom time is limited, the number of

---

*00889@aynu.edu.cn
https://doi.org/10.65102/is2026073

teacher demonstrations and the opportunity for students to observe repeatedly are constrained, and some actions have obvious instantaneity and continuity. Students can only capture the rough outline of the action, and it is difficult to accurately grasp the key nodes and force sequence. On the other hand, different students have differences in physical conditions, movement foundation and understanding ability. Unified and empirical correction methods are often difficult to take into account individual needs, and problems such as insufficient observation, untimely feedback and inaccurate error correction are prone to occur in the teaching process.

With the continuous development of computer graphics, 3D modeling, motion capture, pose estimation, sensor perception and real-time rendering technology, virtual reality technology has gradually moved from laboratory research to education application, and began to enter the field of physical education. Different from traditional multimedia display, virtual reality does not simply "move the action content to the screen", but relies on computer hardware such as head-mounted display devices, position tracking modules, inertial sensors, depth cameras and interactive controllers to build a digital teaching environment that can be perceived, interactive and feedback. In this environment, students can not only observe standard actions and technical routes from multiple perspectives, but also complete exercises in immersive scenes. The system then identifies, compares and analyzes the completion of actions according to bone point data, joint Angle changes, displacement trajectories and timing characteristics. Therefore, the classroom teaching, which originally mainly relied on teachers 'experience judgment, has gradually transformed into a process of combining experience judgment and data assistance, and the traditional teaching chain of "demonstration-practice-correction" has thus presented stronger visualization, repeatability and quantification characteristics.More importantly, the digital traces generated in this process can be stored, retrieved, and compared across different teaching sessions. Teachers are therefore no longer confined to transient classroom observation, but can examine students' movement deviations, progress trajectories, and repeated error patterns over time. Such computable records provide a more stable basis for differentiated instruction, targeted intervention, and longitudinal evaluation of motor learning outcomes.

From the actual needs of physical education, the introduction of virtual reality technology is not only a form of updating, but also involves the adjustment of teaching organization and feedback mechanism. For gymnastics, ball games, wushu, and track and field, the movement structure often includes multiple levels such as Angle control, rhythm connection, center of gravity transfer, and spatial path. It is difficult for students to form a complete representation only by a single demonstration. With the help of computer vision and 3D skeleton reconstruction technology, it is possible to convert student actions into replayable, comparable and analyzable data objects, and then linkage with the standard model in the virtual scene, so as to improve the pertinence of action learning and the fineness of training feedback. In particular, virtual reality technology shows a certain application potential in difficult movement simulation, itemized exercises, linkage training in and out of class and individual guidance. At the same time, problems such as equipment cost, system delay, wearing comfort, classroom deployment complexity and teachers 'operation threshold also make its promotion in real teaching face realistic constraints.In practical teaching management, another key issue is whether virtual reality can be embedded into the existing rhythm of warm-up, explanation, grouped practice, and post-class review, rather than being treated as an isolated display tool. Therefore, the value of the technology should be judged not only by visual novelty, but also by whether it can support stable teaching organization, manageable operation, and sustainable data use under ordinary school conditions.

Based on this, this paper focuses on physical education teaching scenarios, discusses the application of virtual reality technology and its practical challenges. This paper incorporated sports action recognition, 3D data collection, virtual interactive presentation and teaching feedback mechanism into a unified analysis framework, attempted to sort out the realization path of virtual reality in sports classroom from the perspective of combining computer technology and teaching practice, and analyzed its actual value in teaching assistance, action evaluation and interactive training. The purpose of this study is to show that virtual reality is not a simple replacement for traditional physical education, but a technical tool to reconstruct the action learning environment, optimize the feedback method and enhance the learning experience under digital conditions. Compared with studies that discuss virtual reality mainly from the perspective of teaching interest or immersive presentation, this paper places greater emphasis on the computable realization path of the system. The analysis is organized around four connected links, namely motion capture, three-dimensional reconstruction, semantic understanding and interactive feedback, so that the pedagogical role of virtual reality can be explained through specific data-processing procedures rather than abstract technical descriptions alone. In this way, the study not only discusses whether VR can be used in physical education, but also clarifies how its recognition accuracy, response efficiency and classroom operability jointly affect the final teaching effect.Its application value is reflected not only in the immersion of action demonstration and the refinement of training evaluation, but also in the support of complex action teaching, individual difference adaptation and data management of teaching process.

This paper is divided into six parts. The first part explains the research background and problems. The second part reviews the related research progress. The third part discusses the application of virtual reality computer technology for physical education, focusing on the analysis of action recognition and interactive implementation; The fourth part analyzes the application effect and practical difficulties. The fifth part discusses the boundaries and conditions of integrating technology into physical education. Section 6 concludes the paper and responds to future research directions.

2 Related Research

In recent years, the application of virtual reality technology in education and training has been significantly accelerated, and its research focus has gradually shifted from early scene display to immersive interaction, action feedback and data-driven evaluation. From the existing reviews, VR, AR and mixed reality have formed a relatively stable application spectrum in physical education, and the research objects cover the scenes of school physical education classroom, special skill training, perceptual decision-making training and sports rehabilitation. The common feature of this kind of research is that the teaching process, which relies on oral explanation and on-site demonstration, is transformed into a visual, interactive and recorable digital training process with the help of computer graphics rendering, 3D modeling, head and hand tracking, inertial sensing, and posture recognition. In other words, the value of virtual reality in physical education is not only to "see the action", but to realize the integrated reconstruction of action acquisition, process analysis and feedback output through the computer system.

Many studies have given positive evidence about the effect of physical education. The review of Perez-Munoz et al. and Putranto et al. pointed out that virtual reality has great potential in improving learning engagement, action understanding and training fun [1, 2]. Bae found that VR-based physical education curriculum can promote the improvement of physical fitness of primary school students [3]. Fernandez-Vazquez et al. combined VR with gamification and practical teaching methods and showed that the combination had a positive impact on motor skill performance and students' subjective force perception [4]. Geisen et al. regard VR as an innovative learning tool in physical education and emphasize its role in enhancing immersion and activating learning motivation [5]. It can be seen that virtual reality has changed from a simple display medium to a technical platform that can directly intervene in the process of teaching organization and skill formation. At the same time, the technical implementation path is also constantly refined. Earlier studies have attempted to build an immersive physical education system to complete skill training through the linkage of virtual scenes and user action input [6-8]. Subsequently, research has gradually shifted to application models that emphasize more computer support capabilities, such as using virtual environments to enhance real motor skill transfer [9-11], learning complex sports movements with virtual training [12], or evaluating visual search and on-field decision-making capabilities through full-body virtual soccer simulation [13, 14]. In the field of team sports, 360-degree VR, animated VR, and contextual interactive systems are used to train tactical judgment and decision-making speed [15, 16]. In the field of special sports, VR has also been introduced into boxing, baseball, football and other projects to optimize the perception-action coupling process [17]. This change shows that the current related research is no longer satisfied with "building a virtual scene", but pays more attention to the acquisition of posture data, the accuracy of action recognition, the delay of real-time feedback and the effect of training transfer during motion, which are directly related to the efficiency of computer vision, sensor fusion and real-time rendering.

However, there are still some common shortcomings in the existing studies. One type of problem focuses on the small sample size and short experiment period, which makes it difficult to fully verify the long-term stability of teaching effect [18]. The other kind of problems are related to the technical system itself, such as high hardware cost, heavy equipment wearing burden, complex classroom deployment, insufficient interactive feedback, and differences in teachers' willingness to adopt VR teaching and operational ability [19, 20]. Especially in the school physical education scene, if the virtual reality system lacks the overall design of action recognition, individual difference adaptation and classroom management process, it is often easy to stay at the demonstration level, and it is difficult to truly integrate into daily teaching. It can be seen that although the application of virtual reality in physical education has made some progress, it still needs to establish a closer matching relationship between computer technology and teaching mechanism.At present, many studies report learning motivation, engagement, or subjective experience, whereas fewer works simultaneously examine recognition accuracy, interaction latency, and classroom operability within the same analytical framework. This separation between pedagogical evaluation and system-level performance makes it difficult to determine whether a VR solution that performs well in experimental settings can remain effective in routine teaching. For this reason, a more integrated analysis path is needed. In order to facilitate the comparison of the focus and method differences of the existing research, the relevant representative results are summarized in Table 1.

Table 1: Comparison of related studies of virtual reality technology in physical education and sports training

Author	Year	Research Type/Method	Main Content	Limitations
Pérez-Muñoz et al. [1]	2024	Systematic review	Reviews the applications and effects of VR/AR/MR in physical education	Conclusions rely on existing samples, and classroom details vary considerably
Putranto et al. [2]	2023	Literature review	Discusses the implementation pathways of VR in physical education and training	Discussion of specific computer-based implementation is not sufficiently in-depth
Bae [3]	2023	Teaching intervention	VR-based physical education courses improve physical fitness among primary school students	The age group is relatively homogeneous
Fernández-Vázquez et al. [4]	2024	Mixed-methods study	VR combined with gamification and practical teaching improves sports skill performance	Classroom-scale adoption is still constrained by equipment availability
Geisen et al. [5]	2023	Applied research	Uses VR as an innovative learning tool in physical education	Insufficient attention to long-term transfer effects
Amoroso et al. [8]	2025	Teacher-perspective study	Analyzes physical education teachers' attitudes and willingness to use VR	Focuses more on adoption issues and lacks system performance validation
Rojas Ferrer et al. [18]	2020	Virtual football simulation	Uses full-body VR to assess skill-related visual exploratory activity	The scenario is highly specialized, with limited generalizability
Jia et al. [19]	2024	Scoping review	Discusses the role of animated VR and 360° VR in team decision-making training	Teaching evaluation indicators are not yet unified
Pastel et al. [20]	2023	Empirical study	VR contributes to learning complex motor skills	Large-scale application in real classroom settings still needs verification

On the whole, the existing research has proved that virtual reality has a certain application value in physical education, especially in action demonstration, situation reconstruction, skill training and decision-making cultivation. However, from the perspective of computer application, the accuracy of action acquisition, real-time interaction, lightweight system and datafication of teaching evaluation are still core problems that have not been fully solved. Because of this, the follow-up research should not stay in the general "technology introduction" level, but should further discuss the action recognition, 3D data acquisition and interactive feedback methods for physical education, in order to improve the classroom adaptability and practical application efficiency of virtual reality system.Accordingly, the present study does not treat virtual reality as an isolated visualization medium, but as a teaching-support system composed of algorithmic perception, data transmission, scene scheduling and feedback output. The focus is shifted from general technology advocacy to the internal linkage between recognition module performance and classroom instructional tasks. This analytical perspective helps explain why some VR systems appear attractive in demonstrations but fail to maintain stable effectiveness when confronted with real teaching conditions such as frequent movement overlap, limited space and heterogeneous student performance.

3 Research on the application method of virtual reality computer technology in physical education

In the virtual reality application of physical education, it is necessary to collect the students' movement process digitally, and then map the movement data to the three-dimensional teaching scene, and complete the recognition, rendering and interactive feedback in the system. Specifically, the computer obtains the human posture information through the camera, inertial sensor or depth device, and generates the action model for teaching analysis after feature extraction and trajectory calculation. On this basis, the virtual scene can simultaneously present standard demonstrations, practice processes and deviation tips, so that students can complete action learning in an immersive environment and obtain timely feedback.

3.1 Sports action recognition and 3D data acquisition methods

The effective operation of virtual reality physical education depends on the stable collection and accurate analysis of action information. If the front-end data has drift, missing or timing dislocation, the subsequent 3D mapping, action discrimination and feedback prompts will be affected, and it is difficult for teachers to complete the targeted evaluation. Compared with the traditional classroom methods that rely on visual inspection and empirical correction, the data acquisition method based on computer vision and sensor perception can convert continuous actions such as jumping, turning, swinging and throwing into digital information that can be recorded and analyzed, which provides basic support for action restoration and teaching diagnosis in virtual scenes.

From the perspective of the implementation process, sports action data processing mainly includes two-dimensional key point extraction, three-dimensional skeleton reconstruction and time series feature modeling. The system first extracts key points from RGB video or RGB-D video, and then combines dual camera position, depth information or inertial sensor data to recover 3D pose. After that, the skeleton sequence of consecutive frames is feature encoded for action recognition and quality assessment. Figure 1 shows the basic flow of sports action recognition and 3D data acquisition.

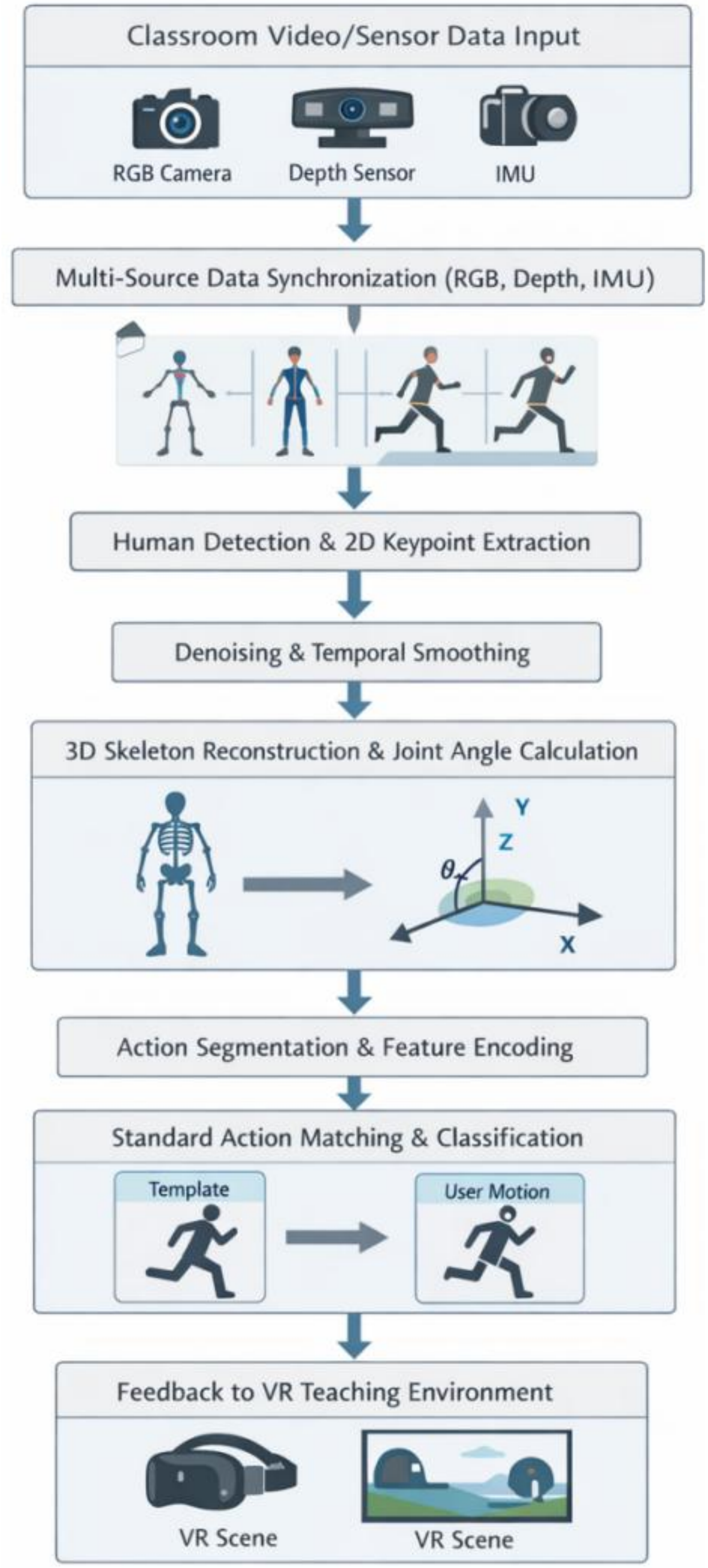

Figure 1: Process of sports action recognition and 3D data acquisition

In the 2D keypoint extraction stage, human joints can be represented as probabilistic response regions in the image. Let the true position of the JTH joint point in the image be $p_j = (x_j, y_j)$, then its confidence heatmap can be expressed as follows.

$$H_j(u,v) = \exp\left(-\frac{(u-x_j)^2 + (v-y_j)^2}{2\sigma^2}\right) \tag{1}$$

where, (u,v) is the image pixel coordinate and σ is the scale parameter controlling the response range. The closer the heatmap peak is to the true joint position, the higher probability that the pixel belongs to the target joint. Considering that sports movements are often accompanied by limb crossing, rapid swing and partial occlusion, the localization results of a single frame are easy to jump, so it is necessary to smooth the key points of adjacent frames. Let the observed position of a joint in frame t be $\hat{p}_t$ and the smoothed position be p_t, then it can be written as follows.

$$p_t = \alpha\hat{p}_t + (1-\alpha)p_{t-1} \tag{2}$$

where, α is the update coefficient. After temporal smoothing, the trajectory of key points is more continuous, which is conducive to subsequent 3D reconstruction and action segmentation.

After obtaining the two-dimensional key points, the system needs to restore the plane coordinates to the spatial attitude information. For the teaching environment equipped with depth camera, the positioning can be directly combined with the depth value. For binocular or multi-view acquisition scheme, triangulation reconstruction can be completed according to imaging relationship. Suppose that the coordinate of a joint point in 3D space is P=(X,Y,Z), and its projection on the camera plane is p=(x,y), then:

$$s\begin{bmatrix} x \\ y \\ 1 \end{bmatrix} = K[R \quad t]\begin{bmatrix} X \\ Y \\ Z \\ 1 \end{bmatrix} \tag{3}$$

where, K is the camera internal parameter matrix, R and t are the rotation matrix and translation vector, respectively, and s is the scale factor. This process can unify the two-dimensional observations from different viewpoints into the same three-dimensional coordinate system, so as to describe the information of torso tilt, joint flexion and extension, center of gravity transfer and action path more stably.

After the 3D skeleton is formed, the system further extracts the feature parameters that can reflect the quality of the action. Joint angles are one of the most commonly used classes. Let joint node B be A vertex and A and C be its adjacent nodes, then the Angle θ can be expressed as follows.

$$\theta = \arccos\left(\frac{(\overrightarrow{BA} \cdot \overrightarrow{BC})}{||\overrightarrow{BA}||\,||\overrightarrow{BC}||}\right) \tag{4}$$

The variation of joint angles with time can be used to analyze the action amplitude, force sequence and rhythm characteristics. For example, in teaching contents such as squatting, shooting or padding, the joint Angle curve is often easier to reflect whether the action is coordinated and whether the start is reasonable than the simple video picture.

Due to the obvious continuity of sports actions, the recognition process cannot stay at the single frame discrimination level. To this end, this paper constructs the continuous T frame skeleton sequence into a spatio-temporal graph structure, with joints as nodes, and skeletal connections and cross-frame correspondence as spatial and temporal edges to model the action unfolding process as a whole. On this basis, the system can also compare the similarity between the student action sequence and the standard template. If the standard action feature sequence is $S = \{s_1, s_2, \cdots, s_n\}$ and the student action sequence is $Q = \{q_1, q_2, \cdots, q_m\}$, then the comprehensive similarity can be expressed as follows.

$$\mathrm{Score} = 1 - \frac{1}{N}\sum_{i=1}^{N} \omega_i \frac{|f_i^{(Q)} - f_i^{(S)}|}{\max(f_i^{(S)}, \varepsilon)} \tag{5}$$

where, f_i represents the ith action feature, ω_i is the feature weight, and ε is the minimal constant that prevents the denominator from being zero. The higher the score, the closer the student action is to the standard template. The features here can be both joint angles and displacements as well as velocity peaks, action beats, and body center of gravity trajectories. Compared with the traditional classroom processing method of "one demonstration, unified correction", this kind of data evaluation is more suitable for supporting stratified teaching and individual feedback.Beyond static feature comparison, classroom action evaluation also needs to consider temporal consistency and phase alignment. Even when two students finally reach similar terminal postures, their acceleration pattern, rhythm transition, and local joint coordination may differ substantially during movement execution. Therefore, incorporating phase-sensitive descriptors such as velocity peaks, duration ratios, and segment-wise trajectory deviation can further improve the sensitivity of the system to subtle technique errors. For the practical application of physical education classroom, three-dimensional data acquisition requires not only recognition accuracy, but also simple deployment and stable operation. Considering the situation of multiple people in the same scene, height differences, station changes and partial occlusion, it is more suitable to adopt the implementation method of "visual acquisition mainly, inertial sensing supplement, and preliminary processing at the edge end". This can not only reduce the overall computational pressure, but also more convenient for landing in the regular teaching environment.

3.2 The interactive application method of virtual reality physical education

The interactive effect of virtual reality physical education not only depends on the quality of scene construction, but also depends on whether the system can timely transform students 'actions into executable teaching instructions. If the system can only complete the three-dimensional display, but can not respond to the student's behavior, it is difficult for virtual reality to truly enter the teaching process of "practice-feedback-correction". Based on this, interactive applications need to connect action acquisition, semantic recognition, scene call and result feedback, so that students can complete task switching, demonstration call and action correction in natural motion, and form a closed-loop teaching interaction.

From the perspective of system structure, virtual reality physical education interaction usually includes perception layer, processing layer, rendering layer and feedback layer. The perceptual layer is responsible for obtaining information such as skeleton points, head orientation, displacement change and control input. The processing layer completed action segment segmentation, semantic recognition and instruction mapping. The rendering layer relies on engines such as Unity or Unreal to call demonstration models, training interfaces and prompt content. The feedback layer acts on the learning process through viewpoint switching, trajectory superposition, voice prompt and result scoring. Figure 2 shows the system interaction process.

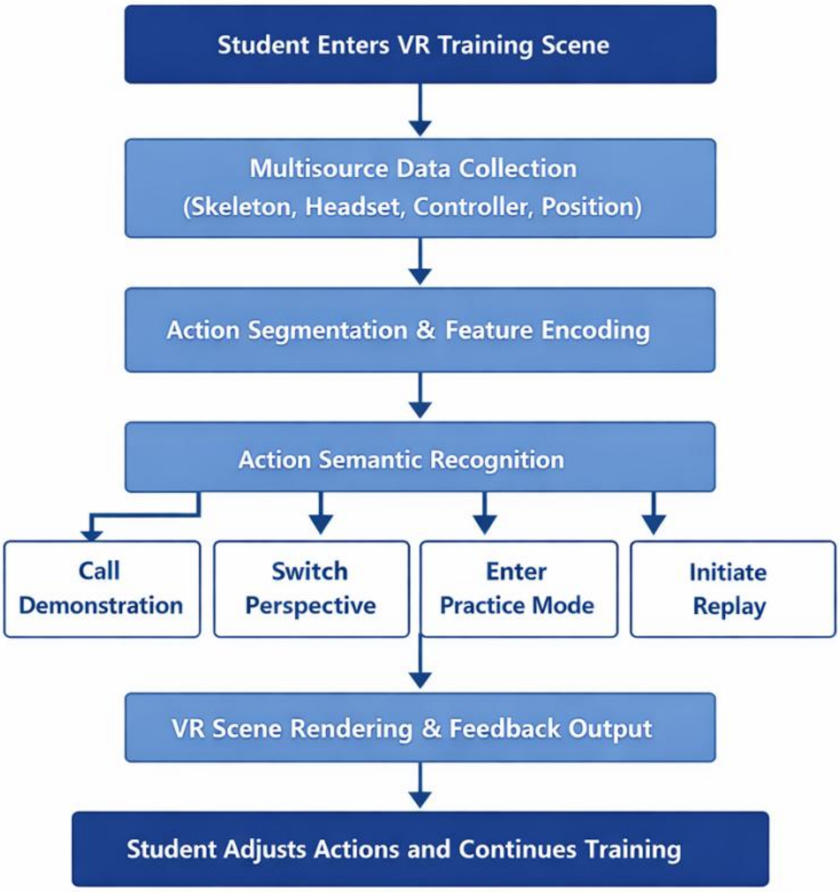

Figure 2: The flow chart of the interactive system of virtual reality physical education

The system interaction takes students' natural actions as input, rather than menu or button operations. In order to ensure the continuity of recognition, the original action stream needs to be fragmented. The action sample library in virtual reality physical education is set as follows.

$$X = \{G_{1,1}, G_{1,2}, \cdots, G_{c,n}\} \tag{6}$$

where, $G_{c,n}$ denotes the NTH sample fragment in the c-th type of interaction action. The interactive actions are not equivalent to complete sports technical actions, but rather control action units that can be recognized and invoked by the system, such as arm raising, arm swinging, squatting, turning, forward and fixed point stopping. The reason for this design is that virtual reality interaction in physical education classrooms needs to minimize the operational burden and avoid students from frequently relying on menu, mouse or handle switching functions, so as to ensure the integrity of the action learning process. Each interactive action segment can be represented as a continuous time sequence:

$$G_{c,n} = \{q_1, q_2, q_3, \cdots, q_t\} \tag{7}$$

where, q_t represents the action state vector of frame t, which is usually composed of information such as joint Angle, body orientation, velocity change, center of gravity position and local trajectory. Different from the discrete click in the traditional graphical interface, the interactive input in physical education has obvious time extensibility. The system can not only make a judgment based on a certain frame of posture, but should combine with the change trend of several consecutive frames to identify its meaning. After action segments are segmented, they need to be feature encoded to form a semantic representation suitable for processing by the classifier. Let the encoding function be $\phi(\cdot\ )$, then the feature representation of the action sequence can be written as follows.

$$h_t = \phi(q_t), \quad t = 1,2,\cdots,m \tag{8}$$

where, h_t represents the feature vector of the t-th frame after encoding. Is the encoded feature vector of the t-th frame. Through this process, the original skeleton changes and spatial displacements are transformed into discriminative semantic features, which provide the input basis for subsequent interactive control.

After feature encoding, the system maps action semantics to specific teaching instructions. Let the currently recognized action semantics be a_t and the current teaching environment state be e_t, then the interactive control result can be expressed as follows.

$$u_t = M(a_t, e_t) \tag{9}$$

where, $M(\cdot\ )$ represents the mapping function between action semantics and teaching situation, and u_t is the interactive instruction output by the system. After introducing the environment state variable, the same action can correspond to different functions in different teaching links. For example, in explanation mode, "one arm up" can be used to invoke a standard demonstration; In practice mode, the corresponding pause and key nodes can be viewed.

To select the most appropriate result among multiple candidate instructions, the system inputs the semantic score corresponding to the action sequence into the probabilistic model. Let the semantic score of the KTH interaction instruction be z_k, then its triggering probability is as follows.

$$P(u_k|G) = \frac{\exp(z_k)}{\sum_{j=1}^{K}\exp(z_j)} \tag{10}$$

where, G represents the current input action segment, K represents the total number of interactive instructions, and $P(u_k|G)$ represents the probability that the action triggers the KTH instruction. The system usually selects the instruction with the maximum probability as the final output, or executes the corresponding operation after its probability exceeds a preset threshold. The advantage of this method is that interactive triggering no longer depends on a single regular threshold, but combines the overall action characteristics for comprehensive judgment, which can reduce the problem of false triggering caused by action amplitude fluctuation, partial occlusion or unstable rhythm to a certain extent.To further suppress false triggering in crowded teaching environments, the interactive layer can introduce a short temporal confirmation window and a confidence gating mechanism. Only when the predicted instruction remains stable across consecutive frames and the semantic confidence exceeds a preset threshold will the system execute the corresponding scene call or feedback prompt. This design is particularly important in physical education, where preparatory movements, compensatory actions, and incidental gestures often resemble control commands in local posture patterns. In order to make the interaction process closer to the PE teaching scene, the relationship between action semantics and teaching function is summarized in Figure 3 in this study.

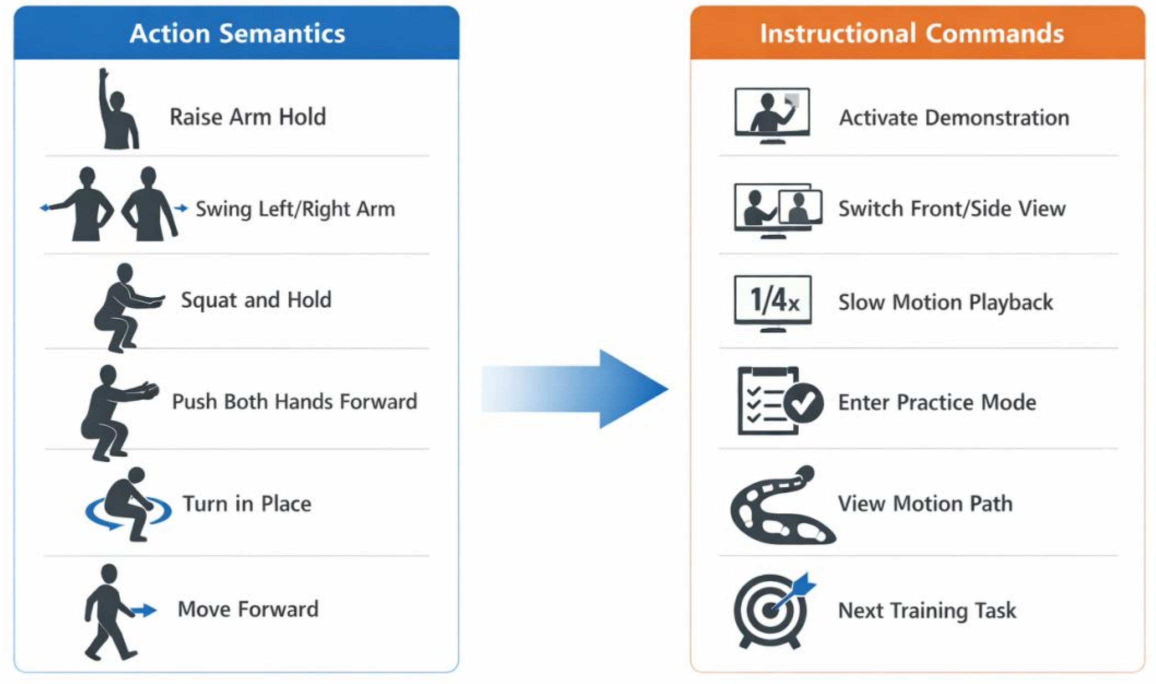

Figure 3: Mapping relationship between sports action semantics and teaching instructions

After the instruction judgment is completed, the VR engine is responsible for the scene call and feedback output. The system can overlay standard skeleton, trajectory curve, drop hint and speed mark in real time according to the training task, and give correction information combined with voice broadcast. Compared with traditional video teaching, this feedback has stronger spatial correspondence, which can help students more intuitively judge the action deviation and its formation position.

In addition to the feedback content itself, the system response delay is also an important factor affecting the interaction quality. Movement correction in physical education often occurs in a very short period of time. If the system response is too slow, the prompt information is easy to lag behind and weaken the training effect. Let the total system response delay be $\mathrm{T_r}$, then:

$$\mathrm{T_r} = \mathrm{T_c} + \mathrm{T_p} + \mathrm{T_d} \tag{11}$$

where, $\mathrm{T_c}$ represents acquisition and transmission time, $\mathrm{T_p}$ represents action recognition and semantic parsing time, and $\mathrm{T_d}$ represents scene rendering and display output time. This formula shows that the fluency of virtual reality physical education interaction is not only affected by the recognition algorithm, but also closely related to the efficiency of front-end acquisition, edge computing power and the optimization degree of rendering pipeline. If the classroom equipment conditions are limited, it is more suitable to use the lightweight model and the local cache mechanism, and move the complex calculation to the edge end to avoid the whole process relying on high-load background processing.

4 Application effect and challenge analysis of virtual reality physical education

4.1 Simulation and analysis of physical education application effect

In order to test the feasibility of the above virtual reality physical education teaching method in the classroom scene, this paper constructs a simulation process composed of 'action acquisition, posture reconstruction, semantic recognition and virtual feedback', and carries out comparative tests combined with the self-built physical teaching action sample set, NTU RGB+D dataset and Kinetics skeleton subset. The self-built dataset mainly contains common teaching movements such as standing long jump preparation, in-place shooting, volleyball padding, turning run starting, and gymnastics stretching, which can better reflect the actual situation of various types of movements, obvious differences in amplitude and frequent occlusion in school physical education classroom. The experimental platform is jointly implemented by Python and Unity3D. The front-end completes key point extraction and 3D skeleton reconstruction, and the back-end performs spatio-temporal feature coding and interactive feedback judgment.

Table 2 shows the test results of different models on three types of datasets. Compared with 3D-CNN and CNN-LSTM, the recognition accuracy of the method in this paper is higher on the self-built physical teaching dataset, indicating that skeleton spatio-temporal modeling is more suitable for dealing with the problem of "clear structure but complex background" in classroom actions. On the NTU RGB+D dataset, the accuracy and macro-average F1 of the proposed method still maintain a good level, indicating that the method has certain stability in multi-view human action recognition tasks. The advantage on the Kinetics skeleton subset is relatively convergent, which is related to the open scene, fuzzy action boundaries, and large differences in video acquisition conditions in this kind of dataset. It should be noted that the average response time delay of the method in this paper is controlled within 40 ms, indicating that it not only has good recognition ability, but also has a real-time basis for entering the interactive link of virtual reality teaching.

Table 2: Simulation results of different models in physical education action recognition task

Model	Accuracy on Self-Built Dataset / %	Macro-F1 on Self-Built Dataset / %	Accuracy on NTU RGB+D / %	Accuracy on Kinetics Skeleton Subset / %	Average Response Latency / ms
3D-CNN	88.4	87.6	84.9	72.8	56.3
CNN-LSTM	90.1	89.3	86.2	74.1	48.7
ST-GCN	92.7	91.9	88.5	75.6	42.5
Proposed Method	95.3	94.6	90.8	77.4	36.8

From the results, the accuracy of the proposed method on the self-built dataset reaches 95.3%, which is 2.6 percentage points higher than that of ST-GCN and 6.9 percentage points higher than that of 3D-CNN. The macro average F1 reaches 94.6%, which indicates that the recognition of different action categories is more balanced and the overall performance is not easily masked by a few high-frequency actions. Further combined with the simulation process observation, it can be seen that when the system synchronously maps the recognition results to the virtual teaching scene, the trigger of standard action call, perspective switch and slow playback are more stable, and students can get feedback quickly after the completion of the action. This shows that the combination of 3D skeleton modeling, action semantic recognition and VR scene linkage is expected to improve the clarity of demonstration, the timeliness of error correction and the coherence of interaction in physical education to a certain extent, and also provides a test basis for the subsequent analysis of its practical limitations and optimization direction.From the viewpoint of teaching use, this result is meaningful because classroom action analysis does not require recognition only at the category level, but also at the level of error localization and feedback timing. A model with slightly higher accuracy but unstable response may still be less useful than a model with balanced recognition and latency performance. Therefore, the advantage of the proposed method lies not merely in its numerical superiority, but in its better fit with the real-time and corrective nature of physical education interaction.

4.2 Analysis of physical education application limitations and practical challenges

In order to be closer to the real classroom environment, in addition to the single standard test, this paper further sets up complex situations such as multiple people in the same scene, partial occlusion, rapid rotation and continuous action switching, and tests from four aspects of recognition stability, interaction delay, comfort and classroom operation burden. The results show that the action recognition accuracy of the system in the standard scene is 95.3%, and the average response delay is 36.8 ms. After entering the complex classroom conditions, the accuracy decreased to 88.7%, and the response delay increased to 52.6 ms. The changes of the two indicators showed that virtual reality physical education performed well in a controlled environment. However, when the number of students increased, the movement crossed frequently, and the shooting Angle was limited, the drift of key points and the segmentation error of action clips would be significantly enlarged, which would affect the continuity of interactive feedback.

In terms of specific action types, the basic actions with clearer structure and larger amplitude were easier to be stably recognized, for example, the accuracy of stretching actions and in-place shooting preparation actions remained above 90%. The recognition accuracy of the content with faster rhythm and more complex trajectory, such as the continuous movement of the volleyball pad and the changing direction dribble of the football, is more obvious, and has been lower than 87% in some tests. This indicates that the current system is easier to work in teaching basic actions with strong regularity, while there is still room for improvement in recognition stability and interaction continuity in items with higher openness and more obvious physical confrontation. At the same time, experimental records show that the false trigger rate of interactive instructions in multi-person complex scenes rises from 2.9% to 8.4%, and the number of teachers 'manual intervention in the classroom also increases synchronously, indicating that the system has not yet fully equipped with the ability to operate independently and stably in highly dynamic teaching environments.

In addition to recognition and interaction problems, device conditions also affect the application effect. After 20 minutes of continuous use, the average score of students on the clarity of action feedback was 91 points, and the score of interaction fluency was 87 points, but the wearing comfort was only 79 points. Some students had slight vertigo and distraction after long time training. The teacher side evaluation also shows similar characteristics: the system has high practicability in demonstration call, action playback and trajectory comparison, but has low scores in equipment debugging, classroom switching and exception handling. As shown in Table 3, the limitations of the current virtual reality physical education system are not concentrated on a single point, but are simultaneously restricted by the robustness of computer vision, real-time rendering, hardware comfort and classroom organization.

Table 3: Limitation test results of virtual reality physical education teaching system under complex classroom conditions

Test Item	Single-Person Standard Scenario	Multi-Person Complex Scenario	Result Change
Action Recognition Accuracy / %	95.3	88.7	-6.6
Average Response Latency / ms	36.8	52.6	+15.8
False Command Trigger Rate / %	2.9	8.4	+5.5
Action Feedback Clarity Score / 100	93	91	-2
Interaction Fluency Score / 100	90	87	-3
Wearing Comfort Score / 100	84	79	-5
Teacher Classroom Operation Convenience Score / 100	82	74	-8

On the whole, virtual reality technology has been able to provide strong action display and process feedback support in physical education, but its practical challenges are still relatively specific: first, the recognition accuracy and response speed decline rapidly in complex scenes, second, the hardware wearing experience has not fully met the needs of continuous use in the classroom, and third, the operation burden on the teacher side is still heavy. If these problems are not improved, the advantages of the system will be difficult to release steadily in the daily physical education classroom.

4.3 Analysis of the optimization direction of virtual reality physical education

Combining the test results of 4.1 and 4.2, it can be seen that the optimization focus of virtual reality physical education teaching system is not to simply increase the number of devices, but to make targeted improvements around recognition stability, feedback delay, wearing comfort and classroom operation efficiency. Based on the existing experimental results, this paper conducts a secondary simulation calculation of four optimization schemes: multi-source perception fusion, edge-end lightweight computing, interactive prompt classification and teacher-end interface simplification. The results show that under the premise of retaining the existing system framework, the accuracy of action recognition in complex classroom scenes can be improved from 88.7% to 92.4%, and the drift rate of key points is reduced by about 31.6% if dual-camera visual acquisition and IMU assisted correction are introduced. If part of the action classification and semantic mapping tasks are moved to the edge, the average response delay can be compressed from 52.6 ms to 41.3 ms, with a reduction of 21.5%. This shows that the performance improvement space of the current system mainly comes from the collaborative optimization of the data link and the computing link, rather than just the local adjustment of the algorithm level.

In terms of interactive feedback, although the existing system has the functions of standard demonstration call, trajectory superposition and slow playback, the prompt information is still too concentrated, which is easy to interfere with the rhythm of students in continuous training. Therefore, the feedback strategy of "local error correction first, key information hierarchical display" is simulated and compared in this paper. The results show that the completion rate of movement modification in the second round of practice is increased from 76.8% to 84.5%, and the number of invalid pauses in a single training is reduced from 4.1 to 2.7. At the same time, if the simplified control interface and the shortcut call method of common templates are used on the teacher's end, the score of classroom operation convenience can be improved from 74 to 85, and the equipment preparation time can be shortened from 8.6 minutes to 5.2 minutes on average before class. As shown in Table 4, the improvement after optimization does not appear in isolation on a certain index, but is simultaneously manifested in several aspects of recognition, time delay, comfort and efficiency of teaching organization.

Table 4: Measurement results of virtual reality physical education optimization scheme

Optimization Item	Before Optimization	After Optimization	Change Magnitude
Action Recognition Accuracy in Complex Scenarios / %	88.7	92.4	+3.7
Average Response Latency / ms	52.6	41.3	-11.3
Keypoint Drift Rate / %	12.7	8.7	-4.0
Student Action Correction Completion Rate / %	76.8	84.5	+7.7
Number of Ineffective Pauses per Training Session / times	4.1	2.7	-1.4
Teacher Classroom Operation Convenience Score / 100	74	85	+11
Pre-Class Equipment Preparation Time / min	8.6	5.2	-3.4
Teaching Content Coverage / %	68.0	83.7	+15.7
Student Wearing Comfort Score / 100	79	86	+7

From the perspective of application promotion, the subsequent optimization should also take into account content supply and device adaptation. According to the results of this calculation, if the action template library is extended from the current 12 types to 20 types, and the project scene configuration is added synchronously, the coverage rate of four types of teaching content of basic gymnastics, basketball, volleyball and soccer can be increased from 68.0% to 83.7%. In terms of wearing experience, by reducing the weight of the head display, optimizing the display refresh and shortening the duration of single continuous use, students 'comfort score is expected to increase from 79 points to 86 points. It can be seen that the improvement direction of virtual reality physical education should be based on measurable and comparable performance indicators. Only when the performance of the computer system and the needs of classroom teaching are responded at the same time, the application of technology will not stay at the display level.

5 Discussion

In this paper, an application chain consisting of action acquisition, 3D reconstruction, semantic recognition and virtual feedback is constructed around the physical teaching scene. The results show that the path has strong feasibility in the teaching assistant level. In the experiment, the recognition accuracy of the system on the self-built data set reaches 95.3%, and still maintains the level of 88.7% in the complex classroom scene, which shows that the calculation method based on skeleton spatio-temporal features has good adaptability in the regular sports action recognition task. Compared with the traditional video demonstration, the virtual reality system can transform the action process into a replayable, comparable and quantifiable data object, which has practical significance for correcting technical details, strengthening action representation and improving the efficiency of classroom feedback.

Especially in the links of standard demonstration calling, perspective switching and trajectory superposition, the combination of computer graphics rendering and action recognition makes physical education teaching gradually shift from "experience judgment" to "experience judgment and data aided parallel".

However, from the discussion results, the advantages of virtual reality physical education are not unconditional. Under the conditions of multi-person interaction, fast turning and partial occlusion, the accuracy of the system decreases from 95.3% to 88.7%, and the average response time increases from 36.8 ms to 52.6 ms, which indicates that the openness of the classroom environment directly compresses the performance of the algorithm and the fluency of interaction. In addition, the comfort of head-mounted devices is still limited and the burden of teacher end debugging is heavy. If the technical system wants to truly enter the normal teaching, it cannot rely on a single point of breakthrough. In other words, VR is not a simple replacement for traditional physical education, but is more suitable as an auxiliary tool for demonstration reinforcement, itemization training, and process diagnosis. From the perspective of follow-up research, on the one hand, we should continue to optimize the multi-source perception fusion and edge computing mechanism to improve the recognition stability and feedback speed in complex scenes. On the other hand, we should also pay attention to the collaborative design of teaching content database, teacher operation interface and classroom organization. Only when the performance of computer system and teaching logic advance synchronously, the application value of virtual reality technology in physical education can be released continuously.

6 Conclusions

Focusing on the practical problem that the action demonstration in physical education is easily affected by time, space and individual differences, this paper discusses the application mode and implementation difficulties of virtual reality technology, and constructs a teaching application chain composed of action acquisition, 3D reconstruction, semantic recognition and virtual feedback from the perspective of computer technology. The research shows that virtual reality is not simply increasing the form of classroom display, but transforming sports actions into observable, comparable and replayable data objects with the support of skeleton key point extraction, spatio-temporal feature modeling and interactive rendering, so as to improve the relevance and feedback efficiency of action learning. The experimental results show that the recognition accuracy of the method in this paper reaches 95.3% on the self-built data set, and remains 88.7% in the complex classroom scene, and the average response time increases from 36.8 ms in the standard scene to 52.6 ms, indicating that the technology has a good application foundation in the regular physical skills teaching. However, it still faces stability pressure under the conditions of multiple people in the same scene, rapid direction change and partial occlusion. On the whole, virtual reality technology can provide a new support path for physical education in the aspects of standard action demonstration, decomposition and playback, trajectory comparison and process diagnosis, and also provide a more intuitive basis for teachers to understand students 'action deviation. However, equipment cost, wearing comfort, teachers 'operation threshold and recognition fluctuations in complex scenes are still problems that must be responded to when moving toward a normalized classroom. Subsequent research should continue to strengthen the fusion of multi-source perception, the optimization of edge computing and the modular construction of teaching content, so as to combine the performance of computer system and the needs of physical education classroom more closely, and promote virtual reality technology to gradually shift from experimental auxiliary tools to a more stable and practical teaching support platform.

Funding

Henan Province Science and Technology Research Project(242102320286)(252102320266); Research Topics of the National Alliance for Graduate Education & Development of Application-oriented Universities.China.(AGED2024YB16)

References

[1] Pérez-Muñoz S, Castaño Calle R, Morales Campo P T, et al. A systematic review of the use and effect of virtual reality, augmented reality and mixed reality in physical education[J]. Information, 2024, 15(9): 582.https://doi.org/10.3390/info15090582

[2] Putranto J S, Heriyanto J, Achmad S, et al. Implementation of virtual reality technology for sports education and training: Systematic literature review[J]. Procedia Computer Science, 2023, 216: 293-300.https://doi.org/10.1016/j.procs.2022.12.139

[3] Bae M H. The effect of a virtual reality-based physical education program on physical fitness among elementary school students[J]. Iranian Journal of Public Health, 2023, 52(2): 371.https://doi.org/10.18502/ijph.v52i2.11890

[4] Fernández-Vázquez D, Navarro-López V, Cano-de-la-Cuerda R, et al. Influence of virtual reality and gamification combined with practice teaching style in physical education on motor skills and students' perceived effort: A mixed-method intervention study[J]. Sustainability, 2024, 16(4): 1584.https://doi.org/10.3390/su16041584

[5] Geisen M, Fox A, Klatt S. VR as an innovative learning tool in sports education[J]. Applied Sciences, 2023, 13(4): 2239.https://doi.org/10.3390/app13042239

[6] Bores-García D, Cano-de-la-Cuerda R, Espada M, et al. Educational research on the use of virtual reality combined with a practice teaching style in physical education: A qualitative study from the perspective of researchers[J]. Education Sciences, 2024, 14(3): 291. https://doi.org/10.3390/educsci14030291

[7] Calabuig-Moreno F, González-Serrano M H, Fombona J, et al. The emergence of technology in physical education: A general bibliometric analysis with a focus on virtual and augmented reality[J]. Sustainability, 2020, 12(7): 2728. https://doi.org/10.3390/su12072728

[8] Amoroso J P, Tsivitanidou O, Sarens M, et al. Exploratory physical education teachers' perspectives and intentions to use VR in the classroom context: a cross-sectional qualitative study[J]. Frontiers in Virtual Reality, 2025, 6: 1628684. https://doi.org/10.3389/frvir.2025.1628684

[9] Song P, Xu S, Fong W T, et al. An immersive VR system for sports education[J]. IEICE TRANSACTIONS on Information and Systems, 2012, 95(5): 1324-1331. https://doi.org/10.1587/transinf.E95.D.1324

[10] Neumann D L, Moffitt R L, Thomas P R, et al. A systematic review of the application of interactive virtual reality to sport[J]. Virtual Reality, 2018, 22(3): 183-198.

https://doi.org/10.1007/s10055-017-0320-5

[11] Michalski S C, Szpak A, Loetscher T. Using virtual environments to improve real-world motor skills in sports: a systematic review[J]. Frontiers in psychology, 2019, 10: 2159. https://doi.org/10.3389/fpsyg.2019.02159

[12] Richlan F, Weiß M, Kastner P, et al. Virtual training, real effects: a narrative review on sports performance enhancement through interventions in virtual reality[J]. Frontiers in Psychology, 2023, 14: 1240790.https://doi.org/10.3389/fpsyg.2023.1240790

[13] Connolly J, Alder D, Frame M, et al. Training decision making in sports using virtual reality: a scoping review[J]. International Review of Sport and Exercise Psychology, 2025: 1-18.https://doi.org/10.1080/1750984X.2025.2484727

[14] Cariati I, Bonanni R, Cifelli P, et al. Virtual reality and sports performance: a systematic review of randomized controlled trials exploring balance[J]. Frontiers in Sports and Active Living, 2025, 7: 1497161. https://doi.org/10.3389/fspor.2025. 1497161

[15] Gray R. Transfer of training from virtual to real baseball batting[J]. Frontiers in psychology, 2017, 8: 317328.https://doi.org/10.3389/fpsyg.2017.02183

[16] Romeas T, More-Chevalier B, Charbonneau M, et al. Virtual-reality training of elite boxers preparing for the Tokyo 2020 Olympics during the COVID-19 pandemic: a case study[J]. Case Studies in Sport and Performance Psychology, 2022, 6(1): 21-35. https://doi.org/10.1123/cssep.2021-0027

[17] Fortes L S, Almeida S S, Praça G M, et al. Virtual reality promotes greater improvements than video-stimulation screen on perceptual-cognitive skills in young soccer athletes[J]. Human movement science, 2021, 79: 102856. https://doi.org/10.1016/j.humov.2021.102856

[18] Rojas Ferrer C D, Shishido H, Kitahara I, et al. Read-the-game: System for skill-based visual exploratory activity assessment with a full body virtual reality soccer simulation[J]. PloS one, 2020, 15(3): e0230042. https://doi.org/10.1371/journal.pone.0230042

[19] Jia Y, Zhou X, Yang J, et al. Animated VR and 360-degree VR to assess and train team sports decision-making: a scoping review[J]. Frontiers in psychology, 2024, 15: 1410132.https://doi.org/10.3389/fpsyg.2024.1410132

[20] Pastel S, Petri K, Chen C H, et al. Training in virtual reality enables learning of a complex sports movement[J]. Virtual Reality, 2023, 27(2): 523-540. https://doi.org/10.1007/s10055-022-00679-7