SIN-CHN SCIENTIFIC PRESS PTE.LTD.

Molecular & Cellular Biomechanics 2025, 22(3), 668.
https://doi.org/10.62617/mcb668

Article

Biomechanical insights and optimization in the teaching design of badminton games based on motion capture and adaptive virtual reality video coding

Qi Tian*, Jiping Tang

College of Physical Education, Chengdu University, Chengdu 610106, China
*** Corresponding author:** Qi Tian, 19983127272@163.com

CITATION

Tian Q, Tang J. Biomechanical insights and optimization in the teaching design of badminton games based on motion capture and adaptive virtual reality video coding. Molecular & Cellular Biomechanics. 2025; 22(3): 668. https://doi.org/10.62617/mcb668

ARTICLE INFO

Received: 30 October 2024
Accepted: 19 November 2024
Available online: 13 February 2025

Abstract: The application of games in sports not only brings new development to sports, but also brings new requirements to sports. To maintain and enhance students' learning motivation and interest, more effective individualized teaching is needed. This study focuses on students' individualized knowledge structure, tracking the differences in sports skills among different types of students, and designing a sports game model based on sports games. A Bayesian network model was used to model learners' knowledge and establish an adaptive badminton competition mode. Then, a new feature intersection method is studied, and a method for deep knowledge tracking is established using feature embedding and attention mechanisms. Finally, on this basis, this method is combined with adaptive learning methods to establish a badminton game model based on adaptive learning and improved deep knowledge tracking. Additionally, this study explores the biological principles underlying sports skill learning. When students learn badminton skills. When students learn badminton skills, the body's proprioceptive system constantly provides feedback. This biological feedback is vital as it helps students adjust their movements unconsciously. Motion capture technology can capture the kinematic data of students' badminton movements, such as joint angles and limb velocities. By integrating this data with the biological feedback, we can optimize learning outcomes by enabling more precise identification of skill deficiencies and more effective remediation. The experiment shows that the acceleration Z is the maximum, approximately 100. The acceleration of X is the smallest, approximately between −200 and 200. The most unstable is the acceleration Y. After positive compensation, the value of the adaptive quantization parameter cascade algorithm increases. And the quantized adaptive quantization parameter cascade algorithm value does not have a significant impact on the evaluation of the reconstructed image. The average values of each scale and sub dimension are above 0.70, and the constituent reliability values are above 0.90, indicating that the internal quality of each scale is good. And the internal consistency between questions is also good, all of which have passed the validity test. The survey method used in this experiment has strong practicality and can effectively achieve game design, which has great practical value in teaching practice.

Keywords: motion capture; virtual reality; badminton sports; video encoding; adaptive technology; biomechanical principles

1. Introduction

Sports competition is an important teaching method aimed at sports competition, which combines the dual attributes of competition and entertainment [1]. In game design, existing motion principles should be used as theoretical support to achieve a balance between motion and the game. Through competition, the sports effect is achieved [2]. Sports have developed rapidly in recent years. Because it can create a game scenario with interesting storylines and rich dynamic interactions, allowing

students to interact in the same situation, thereby making it easier and more enjoyable to improve their physical fitness [3,4]. This provides favorable conditions for personalized learning implementation. At present, there are two problems in building an adaptive sports game learning environment. Firstly, in a gamified learning environment, learners may intentionally or unintentionally overlook some knowledge points, resulting in incomplete knowledge. Secondly, students may encounter some difficulties in completing specific tasks in the game, which can cause them to feel frustrated. This can provide personalized and adaptive sports support for students to improve their exercise efficiency. In the game environment, it is necessary to design a game stand that conforms to the individual characteristics of learners. Through relevant technologies and algorithms, real-time recording of learners' ability changes and path choices during the game process was carried out. Therefore, Adaptive learning algorithm is proposed to be introduced into the motion game. Bayesian network (BN) model was used to characterize the probability relationship between concepts and evaluate the individual knowledge state of learners. BN is a probabilistic graphical model used to represent probabilistic relationships between variables. In the field of education, BN models can be used to model students' knowledge structures and learning paths, as well as to predict students' learning outcomes. In badminton game teaching, the BN model can be used to assess students' mastery of badminton skills and adjust the game difficulty based on this information for personalised teaching. And the characteristics and attention mechanisms of motion dimension are embedded in knowledge tracking model, thus establishing a motion game model that can track students' knowledge structure. The research purpose of this topic is to meet the needs of modern sports, maximize the sports value of sports, so that the quality education and individualized teaching of high-performance sport can be realized.

The novelty of the study lies in focusing on students' individualised knowledge structures and designing a sports game-based model of sports games by tracking differences in sports skills between different types of students. A BN model is used to simulate learners' knowledge and an adaptive badminton competition model is developed to assess learners' individual knowledge states and incorporate features of the sport dimension and attention mechanisms to build a sport game model that can track students' knowledge structures. Meanwhile, the study proposes a new feature intersection method and establishes a deep knowledge tracking method using feature embedding and attention mechanism. A motion capture (MC) algorithm based on long- and short-term memory (LSTM) and a video coding optimisation algorithm based on adaptive quantization parameter (AQP) algorithm are proposed with the aim of bringing the virtual reality (VR) in teaching badminton games. This combination provides a new perspective and technical means to improve the interactivity and efficiency of physical education.

This study will be conducted in four parts. The first is an overview of badminton sports game instructional design based on MC and adaptive VR video encoding. Secondly, there is a research on the teaching design method of badminton sports games based on MC and adaptive VR video encoding. Next is the experimental verification of the second part. Finally, there is a summary of the research content and an indication of shortcomings.

2. Related works

In recent years, MC and VR have become increasingly mature, providing technical support for the development of many fields. However, there are still many problems with MC and VR, and many experts and scholars have made innovative research on this. Simonetto et al. [5] proposed a new methodological framework in their research on motion capture and VR, which locates the application site and usage safety. The new method can improve the clarity of the usage location to a certain extent, and it has better robustness than traditional methods [5]. VR has important application value in traditional education fields such as military training, and traditional VR requires high virtual system costs. In response, Maciejewski et al. [6] propose a VR tracking device and combined it with traditional VR systems. They have optimized key technologies in existing VR and constructed experimental prototypes. This algorithm can significantly improve the performance of VR systems and has a lower cost compared to conventional algorithms [6]. Zhuo and Sharma [7] believe that a new motion capture technology is introduced, that is, the 3D amplitude tracking method in inverse kinematics is used to locate human motion. This algorithm has a significant improvement in capturing conventional and static motion compared to conventional motion capture algorithms, and the accuracy of motion tracking is also greatly improved [7]. Kadirvelu et al. [8] addressed the problem of long and insufficiently accurate disease assessment in Friedreich's ataxia by proposing a method to capture whole-body kinematic data by wearable sensors and analyse digital behavioural features in conjunction with machine learning, which enabled longitudinal prediction of clinical scores in patients with FA with significantly higher prediction accuracy, providing new possibilities for disease course tracking and clinical trial optimisation offer new possibilities [8]. Obukhov et al. [9] addressed the problem of accurately capturing human motion under variable environmental conditions by proposing a method for combining and synchronising data from three motion capture systems (virtual reality trackers, motion capture suits, and cameras using computer vision techniques), which enables a comprehensive assessment of the human body's position and state, and optimises the digital shadowing under the influence of environmental factors such as electromagnetic interference and changes in illumination creation process [9].

Ning and Na [10] argued that 3D sports action recognition can better handle uncertainties in time scales. However, due to the increase in training samples, its recognition performance will decrease. Based on motion sequences, a motion recognition framework was established to classify movements. This method comprehensively analyzes motion data and effectively integrates motion data with motion data. This method has significantly improved recognition accuracy and efficiency [10]. Cai et al. [11] believe that conventional motion capture techniques cannot effectively capture moving targets or characters. For this reason, this paper intends to take sports movements in videos as an example to study a deep historical LSTM. The network will collect and classify each historical information based on LSTM. This method has high efficiency in the implementation process [11]. Hu et al. [12] studied a clustering-based action collection method based on traditional action collection. This algorithm has better generalization performance while ensuring the

accuracy of image segmentation [12]. Buchman-Pearle and Acker [13] proposed a method for evaluating markers during motion capture. This method can effectively reduce the configuration of flags, thereby reducing the number of flags. This algorithm can effectively reduce the false appearance of external clustering and improve clustering accuracy [13]. Uslu et al. [14] believe that although traditional VR can be achieved through robots, research progress in this area is slow due to its high cost. Therefore, applying the least squares technique to motion capture systems can reduce the overhead of the system. Through overall hardware testing, it has been proven that the proposed algorithm can effectively improve the accuracy of motion capture [14]. Vlahek et al. [15] believes that conventional motion capture techniques cannot provide an evaluation tool for imaging. This project aims to use motion capture technology instead of traditional imaging technology to achieve synchronous analysis of motion information of various parts of the human body. This method can effectively reduce the error of the algorithm and improve its accuracy when applied [15].

In summary, MC technology and VR have been applied in many fields. However, during the research process, there are still many problems with this technology. The accuracy of traditional motion capture technology and VR systems is not high enough, and the error is relatively large. The cost of popularizing this technology is too high, and there is still an imprecise phenomenon in motion capture. Therefore, this study is based on traditional virtual technology and MC technology to optimize and improve the teaching algorithm for badminton games, thereby achieving an improvement in the teaching quality of badminton game video encoding.

3. A badminton game teaching plan based on MC and VR video coding

To promote badminton game teaching development, a LSTM based MC algorithm and an AQP based video encoding optimization algorithm are proposed. This proposed plan aims to apply VR to badminton games teaching. Firstly, MC algorithm was used to record VR videos of badminton teaching, and this proposed video encoding algorithm was used to assist students in learning in VR environments.

3.1. Motion capture algorithm based on LSTM recurrent network

LSTM recurrent network is a model that can handle long-term dependencies in sequential data [16]. Based on this, researchers have proposed a new neural network model and introduced it into different gates. This can achieve effective processing of strongly correlated material information and enhance network's memory and generalization capabilities. The information transmission in a cyclic LSTM network is completed by three gates. The input gate controls the update of the content of the memory unit. The 'forgetting' gate controls when the contents of memory units are cleared. In a memory unit, when information is output, output gate is controlled. The cooperation of these three grid control elements enables network to have a good learning ability for long time sequence information, and can well overcome the problems such as gradient disappearance in the traditional recurrent neural network. Cyclic LSTM network is widely used in language modeling, Machine translation,

speech recognition, image description, etc. **Figure 1** shows the neuron and network diagram of LSTM recurrent network.

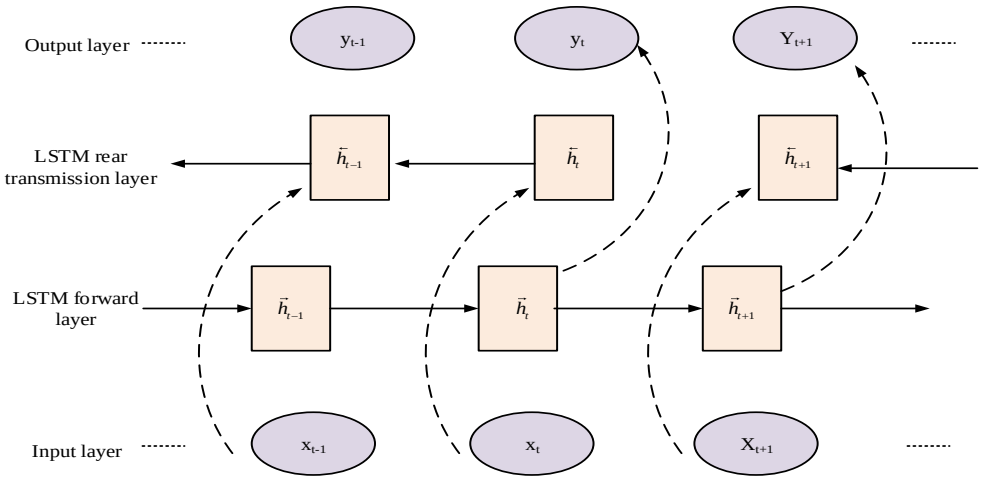

Figure 1. Schematic diagram of neurons and networks in cyclic LSTM.

In **Figure 1**, the LSTM implicit layer consists of multiple neural nodes, and a weight matrix exists between the nodes to link the nodes together. The LSTM model consists of three layers with 200 neurons in each layer. The structure can effectively handle long-term dependencies while keeping the complexity of the model within reasonable limits and performs well in similar sequence data processing tasks. However, compared to general neural networks, LSTM neurons have stronger forgetting ability [17]. Among them, the forgetting mechanism mainly determines whether the previous moment state is saved to next moment by controlling the storage unit in neuron. Due to its ability to effectively handle long-term correlations, LSTM can effectively describe long-term correlations [18]. Compared with other neural networks, LSTM has better memory performance, can effectively solve the long-term correlation of sequence data, and is not prone to common optimization problems such as gradient loss. Based on the structure of the LSTM network shown in the figure, the study uses the SIGMOD function and the tanh function as excitation functions. Where the tanh function is able to change the linear output of the LSTM into a nonlinear form, thus enabling the LSTM to simulate nonlinear functions. The sigmoid function is used in the gating mechanism to determine the flow of information while the tanh function is used in the output layer in order to limit the output value to between −1 and 1, which helps to maintain the stability of the values. Equation (1) is the expression of tanh function.

$$(\frac{e^x + e^{-x}}{e^x - e^{-x}})^{-1} = \tanh(x) \tag{1}$$

x in Equation (1) is function's independent variable. Except for tanh, SIGMOD function is another important excitation function used in Equation (2).

$$\sigma(x) = (e^{-x} + 1)^{-1} \tag{2}$$

where $\sigma(x)$ is the SIGMOD function. **Figure 2** shows the tanh function, and **Figure 2** shows the SIGMOD function. There are significant differences between these two

types of function mappings, as well as their characteristics and parts of action. Although the curve shapes of these two methods are similar, SIGMOD function is non negative and not symmetric. Tanh function has (−1, 1) origin symmetry. It can smooth the gradient when the input amount is too large or too small. And SIGMOD function is within the interval of (0, 1). In SIGMOD function, when the input is near the origin, this function has a larger slope and changes faster [19]. As the input distance from origin increases, the gradient becomes smoother. This gradient law is more in line with logic gate function's requirements. Because logic gate function controls data entry and exit in LSTM neurons. Therefore, SIGMOD function is used as the excitation function for gate function.

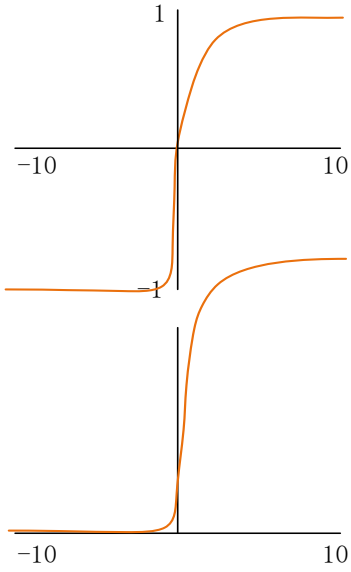

Figure 2. Two types of excitation functions.

To better guarantee the performance of LSTM, the least squares network's loss function is also designed. Among them, two most commonly used loss functions are mean squared error and cross entropy, which can better reflect the difference between model and reality [20]. However, in the context of the proposed LSTM network, the two common error functions reduce the weight update rate due to the small function gradient caused by the SIGMOD function as an excitation function, so the category cross entropy is used here as the loss function. It is suitable for multicategorical problems and can effectively measure the difference between model predictions and actual labels. This is shown in Equation (3).

$$L_t = -\sum_k T_{tk} \log(p_{tk}) \tag{3}$$

In Equation (3), t and k represent the data where the function is located and the classification where the prediction results are located, respectively. p_{tk} represents the predicted value, while $T_{t,k}$ is its actual category. Compared with the single cross entropy function and mean squared error function, categorical cross entropy function can effectively avoid the weight update rate reduction. Finally, to address the problem that the loss function is difficult to carry out accurate minimisation during the solution process, the study uses the gradient descent method for optimisation. Gradient descent

method is an optimal algorithm to polarise the objective function, which can find the minimum value of the loss function and thus improve the accuracy of the model. On this basis, a solution method based on the principle of least squares was proposed. The gradient descent method is a commonly used method, in which objective function is generally composed of a sample set and model parameters. The basic idea is to continuously adjust the gradient direction of objective function when solving, to find the optimal solution. Equation (4) is the mathematical expression of gradient descent optimization method.

$$\phi = \phi - b\frac{\partial J(\phi)}{\partial \phi} \tag{4}$$

In Equation (4), ϕ is the parameter of objective function, and $J(\phi)$ is a objective function. b represents the learning rate of LSTM. In gradient descent, objective function parameters are iterated continuously in negative gradient direction to obtain objective function' minimum value [21]. Overall, **Figure 3** shows the flowchart of a behavior capture algorithm based on an LSTM loop network. Compared to traditional machine learning methods, LSTM behavior capture method in this project has higher accuracy and precision in theory. During collection, three different types of mobile sensors are used to ensure collection results integrity. On this basis, data processing was carried out based on LSTM loop network, and finally presented to user.

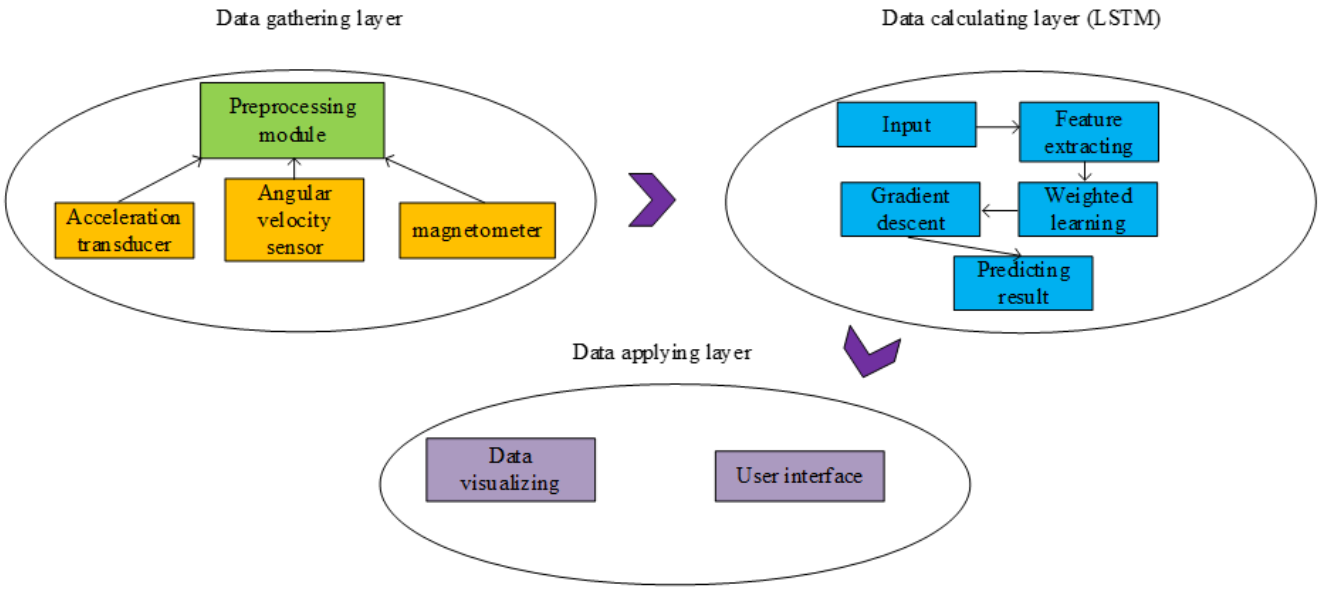

Figure 3. LSTM Motion capture.

3.2. Adaptive QP video coding and its application in badminton game teaching

Quantization parameter (QP) is a parameter that plays an important regulatory role in video encoding, which affects the compression performance of video files. Due to the multiple dispersion quantization of the amplitude of the input signal in this method, there is a loss of information, which puts higher requirements on the selection of quantization parameters. As this study focuses on VR scenarios, High Efficiency

Video Coding (HEVC) mode was adopted and corresponding QP was used for processing. The common quantization method in HEVC is Rate Distortion Optimization Quantization (RDOQ), whose main idea is to match the rate distortion optimization criteria during the quantization process in Equation (5).

$$l_i = \arg\min_{k=1,\ldots,m} \left\{ D\left(C_i, l_{i,k}\right) + \lambda \mathrm{g} R\left(l_{i,k}\right) \right\} \tag{5}$$

In Equation (5), C_i is a transformation coefficient, $l_{i,k}$ is a quantized value, $D\left(C_i, l_{i,k}\right)$ is the quantized distortion, $R\left(l_{i,k}\right)$ is the quantized coding bits, and λ is Lagrange factor. The formula was pre quantified based on the quantified values in Equation (6).

$$\left|l_i\right| = round\left(\frac{\left|C_i\right|}{QP_{step}}\right) \tag{6}$$

In Equation (6), QP_{step} is the quantization step size. $round\left(\mathrm{g}\right)$ is the quantized value after rounding. Based on RDO, the optimal quantization value of coefficients in current Transmission Unit (TU) was determined, traversed in *Z*-scan order, and finally determined. Therefore, whether it is to increase quality policy (QP) or reduce QP, it needs to be set according to the specific situation and cannot be arbitrarily increased or decreased. After analyzing VR video features, targeted processing can be carried out on QP. By projecting a two-dimensional image from a 360° perspective, the resulting two-dimensional image exhibits non-uniform sampling characteristics. Currently, when using existing video encoding standards, VR 360-degree videos need to be projected onto a flat surface before being encoded. When viewed, flat images need to be projected onto spherical videos. In fact, what users see is spherical videos. Considering the characteristics of spherical videos, there may be issues with the focus area. It is inaccurate to directly use the original Peak Signal to Noise Ratio (PSNR) evaluation index to evaluate flat images. Therefore, it is necessary to design new standards to measure video quality. The quantitative image quality assessment criterion weighted to spherically uniform PSNR (WS-PSNR) is a metric for assessing the quality of spherical video, which weights the traditional PSNR values by considering the visual importance of different regions. This weighting method can more accurately reflect the human eye's perception of video quality, especially in 360-degree video and VR applications. The acquisition process of WS-PSNR is shown in Equation (7).

$$\begin{cases} WS-PSNR = 10\log\left(\dfrac{MAX^2}{WMSE}\right) \\ WMSE = \sum\limits_{i=0}^{width-1}\sum\limits_{j=0}^{height-1}\left(y\left(i,j\right)-y'\left(i,j\right)\right)^2 \cdot W\left(i,j\right) \end{cases} \tag{7}$$

In Equation (7), the calculation of WS-PSNR involves a comparison between the original and distorted images. Specifically, it calculates the difference between the original and reconstructed pixel values and weights each pixel according to its visual importance. The weighting factor is determined based on the enterprise resource

planning (ERP) projection weighting map. The degree of image distortion is measured by comparing the difference between the original image and the compressed/distorted image, as seen in Equation (8).

$$W(i,j)=\frac{w(i,j)}{\sum_{i=0}^{width-1}\sum_{j=0}^{height-1}w(i,j)} \tag{8}$$

In Equation (8), MAX is image' maximum value. $y(i,j)$ and $y'(i,j)$ are the original pixel values and the reconstructed pixel values, respectively. $W(i,j)$ is the normalized sphere' weight ratio. $width$ and $height$ are video' width and height, respectively. $w(i,j)$ is a weight scaling factor in Equation (9).

$$w(i,j)=\cos\left(\left(j-\frac{height}{2}+\frac{1}{2}\right)\cdot\frac{\pi}{height}\right) \tag{9}$$

In Equation (9), the displayed image's weight distribution is related to color depth distribution. The ERP projection weighting diagram is shown in **Figure 4**, which shows the visual attention at different latitudes.

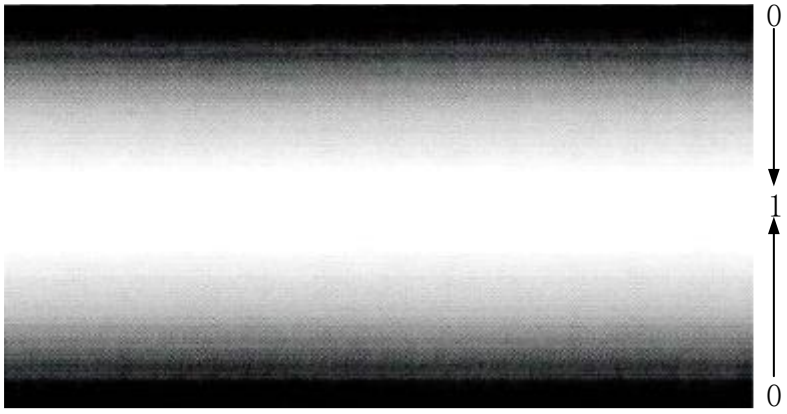

Figure 4. ERP projection weight distribution map.

In **Figure 4**, the color is darker, it is closer to 0. Conversely, it is closer to 1. In the equatorial region, due to its high level of attention, to reduce distortion degree, a negative compensation method was adopted to reduce the quantization step size. In polar region, QP forward compensation will be applied to this area and the quantization step will be increased. Due to the good correspondence between weight w of WS-PSNR and the distortion distribution of ERP projection videos, this experiment considered using the weight w of WS-PSNR as a reference to compensate for QP. A self-adaptive QP based on ERP features was proposed at F meeting of JVET Association, with proposal number F0038. The idea of the proposal is to adjust QP based on latitude. The polar region compensates positively for the set QP, while the equatorial region uses the set QP. The proposal adopts Equation (10) to adjust QP.

$$QP_{new}=QP-3\times\log_2(w) \tag{10}$$

In Equation (10), w is related to latitude, φ is the latitude arc length value, and y is the latitude angle value. F0038 uses $\cos(\pi y)$ as w. For ERP, WS-PSNR uses the weight calculation in Equation (11).

$$w(i,j)=\cos\left(\left(j-\frac{N}{2}+\frac{1}{2}\right)\times\frac{\pi}{N}\right) \tag{11}$$

In Equation (11), N is CTU height and j is pixel position's height. Equation (12) represents the modified weight value.

$$\begin{cases} w_{new}=\dfrac{w_{Index\_Of\_Ctu\_Mean\_Weight}}{Total\_Ctu\_Mean\_weight}\times Num\_Of\_Ctu\_Height \\ QP_{new}=QP-3\times\log_2(w_{new}) \end{cases} \tag{12}$$

In Equation (12), each row's w_{new} of CTU positions has been calculated and its value range is [0, 1.57]. In a VR 360-degree video with a resolution of 3328 × 1664, **Figure 5** shows the magnitude of each row's w_{new} of CTU based on the height variation.

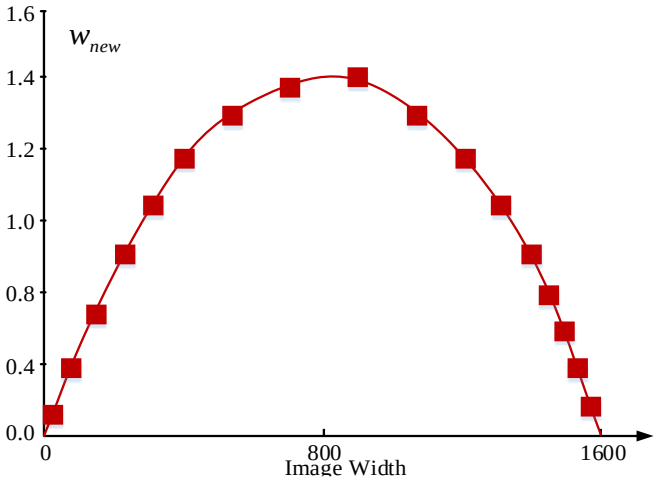

Figure 5. Compensation value based on WS-PSNR weight.

In **Figure 5**, the weighted average values of each elevation point in CTU were obtained through adaptive QP compensation. To ensure that the compensated QP is within the allowable range, the maximum QP value should not exceed 51.

3.3. Teaching design of badminton games integrating MC and adaptive VR video coding

VR is a computer system that provides immersion, interactivity, and presence. It utilizes computers to generate a virtual environment and allows users to enter it through specific devices. In this space, users can have an immersive, realistic, and three-dimensional feeling. In this virtual space, users can operate and control different actions with a mouse. The aim is to explore a badminton tournament teaching model based on motion capture and dynamic imaging as the main tool. Deep Knowledge Tracing (DKT) model has been proven to be an effective knowledge tracking model,

which is more suitable for evaluation scenarios with large user data scales. An improved DKT model was proposed, which enriches the input data of the model through feature embedding and introduces an attention mechanism in **Figure 6**.

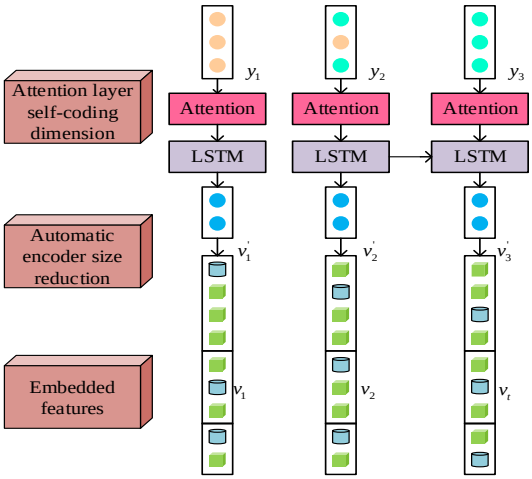

Figure 6. Deep knowledge tracking model based on feature embedding and attention mechanism.

In **Figure 6**, students' answer results are not only influenced by their own answering behavior, but also by the difficulty coefficient of question. Due to its ability to capture complex characteristics in data, DKT model can be feature embedded to improve its performance. Secondly, from students' historical answers, the answers obtained from the previous time point will also have a certain impact on the answers obtained at the current time point. The last hidden state of LSTM is student's hidden learning state at the current moment. Once the time is long, LSTM may cause some important information to be lost. Therefore, this study introduces Attention Mechanism (AM), which weights and aggregates all historical information to reduce important information's loss. Based on this, this study expands model's input data information. The characteristics of students' answering behavior and difficulty coefficients are embedded into the original input information, becoming a more meaningful historical interaction sequence $v_t^{'}$. Therefore, the input historical interaction sequence of this model is $V^{'}=\left(v_1^{'},v_2^{'},\ldots,v_t^{'}\right)$. The output is a probability vector that predicts the probability of students answering the corresponding questions correctly. There are usually three behavioral characteristics in students' answering process, namely the number of times they try to answer the question (Attempt count), whether they request help (First action) during the answering process, and the times they request prompts (Hint count). The additive model was studied to calculate attention scores in Equation (13).

$$s\left(v_i^{'},v_t^{'}\right)=\tilde{v^T}\tanh\left(W_1v_i^{'}+W_2v_t^{'}\right) \tag{13}$$

In Equation (13), $v_i^{'}$ represents the encoded historical sequence information after dimensionality reduction. $v_t^{'}$ represents the input information encoding after

dimensionality reduction at time t. $W, \tilde{v}$ represent the network parameters of science department. The attention state s_t is represented as the weighted sum of H in Equation (14).

$$s_t = \sum_{j=1}^{t} \alpha_j h_j \tag{14}$$

In Equation (14), by combining s_t and h_t, the probability of students mastering all knowledge points in Equation (15) can be calculated.

$$y_t = \sigma\left(W\left(s_t \oplus h_t\right) + b\right) \tag{15}$$

Cross entropy can be used to measure the difference between the real tag and predicted result's probability distribution in machine learning. So, it is used as DKT-FA model's loss function in Equation (16).

$$L = -\sum_{t}\left(a_{t+1}\log y_t^T \delta\left(q_{t+1}\right) + \left(1 - a_{t+1}\right)\log\left(1 - y_t^T \delta\left(q_{t+1}\right)\right)\right) \tag{16}$$

In Equation (16), a_{t+1} represents the true probability distribution, and $y_t^T \delta\left(q_{t+1}\right)$ represents the predicted probability distribution. By tracking students' learning progress, problems can be identified, guided, practiced, diagnosed, and corrected. This process is called one adaptation tracking, and the focus of adaptation is on matching teacher's pattern level. Based on the different behavioral characteristics of students, difficulty levels were set to match their knowledge structure to reflect their personalized learning pathways. Adaptive learning theory emphasises adjusting teaching content and methods according to learners' individual differences (e.g., learning styles, ability levels and interests). In badminton game teaching, adaptive learning theory can help teachers design game levels with different difficulty levels to suit the needs of different students, so as to improve students' learning efficiency and motivation. **Figure 7** shows the framework structure of an entrepreneurial sports game based on adaptive learning and DKT-FA model.

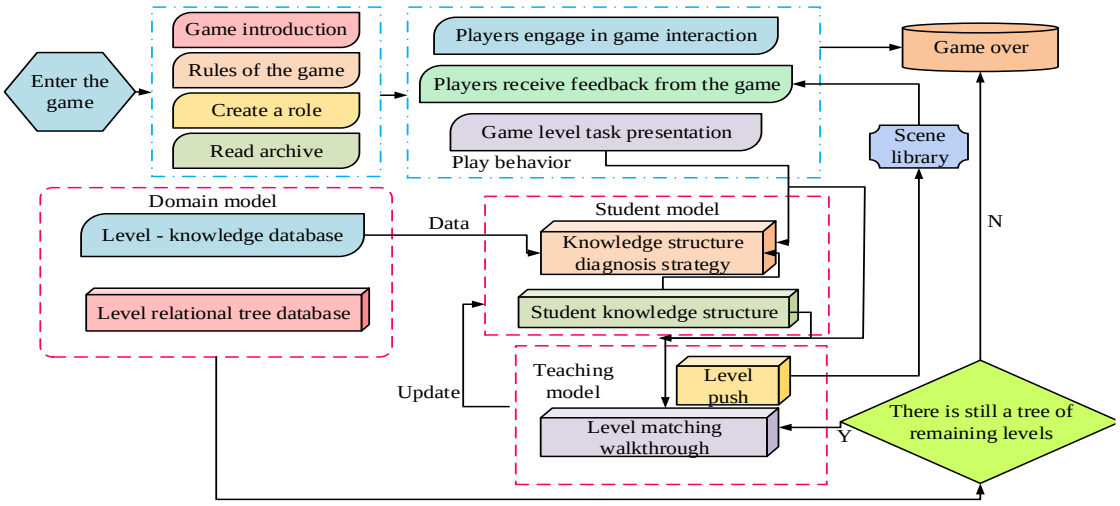

Figure 7. Entrepreneurship education game frame structure based on adaptive learning and DTK-FA model.

In **Figure 7**, firstly, students are familiarized with the basic knowledge of this game and have an understanding of basic game content. Then, they will create a game character and start the game process. The rendering layer is used to display scene effects such as visuals and sounds, and to interact with players and games. Among them, the guidance layer is game's backend computing system, which provides feedback on students' knowledge structure and level matching through the behavioral data of the presentation layer. In guidance layer, it consists of three parts: domain, student, and teacher modules, which respectively complete the establishment of game backend database, evaluation of student knowledge structure, and matching of student adaptive levels. This game starts with the first number of levels stored in the domain model and continues through the last level tree. The student model first presents the behavioral data output by middle school students in presentation layer for knowledge structure diagnosis, and outputs their possible knowledge structures. Based on this, by analyzing the output behavioral data and knowledge structure of middle school students in the presentation layer, corresponding level matching strategies are proposed. The adaptive level deduction and feedback information were output to rendering layer, enabling students to continuously participate in game level and complete student model update. With the BN model, teachers can identify students' weaknesses in badminton skills and design targeted game levels to reinforce these skills. Adaptive learning algorithms can adjust the level difficulty in real time based on students' performance in the game, ensuring that students are always learning at an appropriate level of challenge, and are neither bored by being too easy nor frustrated by being too difficult.

4. Analysis of badminton game teaching design based on MC and adaptive VR video coding

To test whether the performance and function of this system are normal, a badminton club was selected for test. Five hundred volunteers were selected as experimental participants for the study to ensure that the sample was representative of the general population. All participants had no motor impairments and none had received professional badminton training prior to the experiment. Data was collected using a NOKOV branded motion capture system that was capable of capturing participant movement data at a frequency of 120 Hz. In addition, a VR device, PICO 4 Ultra MR, was used to record participants' interactions in the virtual environment. During data collection, each participant completed three sets of specific badminton movements under the guidance of an instructor, including a preparation movement, a foot-side glide, and a hand-leading movement. Data preprocessing included steps such as removing noise, filling in missing values, and data normalisation. The normalisation process involves scaling all the data into the [0, 1] interval to eliminate the effect of different magnitudes. In the model training phase, the dataset is divided into training set (70%) and test set (30%). The training set is used to train the LSTM network model and the test set is used to evaluate the performance of the model. In this case, the results of system identification, as shown in **Figure 8**.

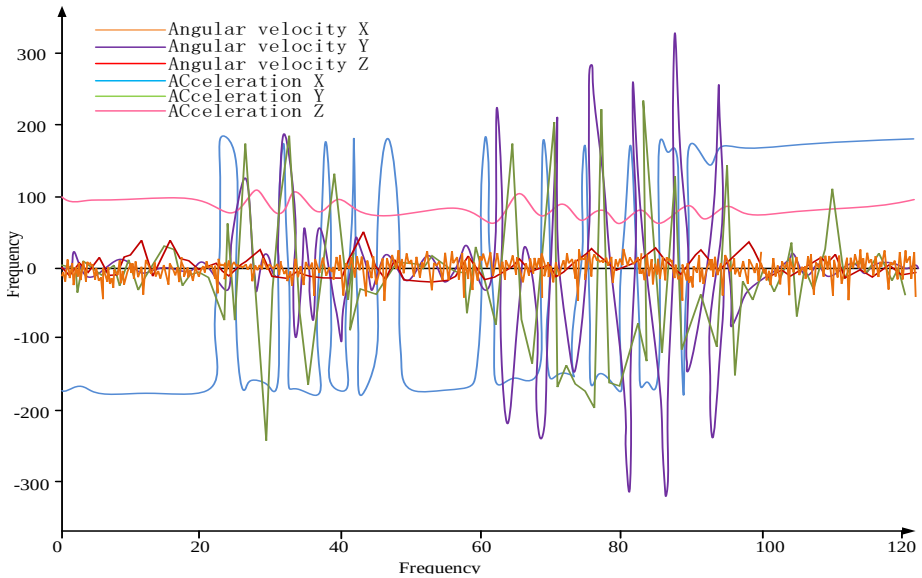

Figure 8. Recognition system results for foot side sliding.

In **Figure 8**, the collected data is labeled and input into LSTM recurrent neural network for Supervised learning to generate recognition model. The model was loaded into this system, and the original unlabeled data was uploaded to this system for recognition. The visual motion curve and recognition results were output. Acceleration Z is the maximum, at which point its value is around 100. Acceleration X is the slowest, fluctuating around −200 to 200. Acceleration Y is the most unstable and the fluctuation amplitude is the largest. Taking into account video reconstruction quality in polar and equatorial regions, it has ultimately improved. The encoding performance improvement of AerialCity video test sequence, Broadway video test sequence, Balboa video test sequence, and Gaslamp video test sequence is not significant. Due to the high texture complexity of video content in polar regions within these above four sequences, the increased QP quantization after compensation has a significant impact on the reconstructed video quality. The use of smaller QP quantization after compensation in the equatorial region improves the reconstructed video quality. However, after comprehensive calculation, it can be concluded that the improved reconstructed video quality has limitations. **Table 1** shows the data of the improved adaptive QP algorithm.

Table 1. Improved adaptive QP algorithm data.

Resolution ratio	Sequence	BD-rate	ΔWS-PSNRy
4K	AerialCity	−1.20%	0.08
	PoleVault	2.40%	0.07
6K	BranCastle2	−1.90%	0.07
	Landing2	2.90%	0.02
	Broadway	−1.10%	0.11
	Balboa	0.70%	0.18

Table 1. (*Continued*).

Resolution ratio	Sequence	BD-rate	ΔWS-PSNRy
8K	Gaslamp	−1.20%	0.06
	Trolley	2.00%	0.05
	Harbor	2.20%	0.03
	KiteFlite	2.00%	0.04
	ChairliftRide	−1.70%	0.05
	SkateboardInLot	−1.50%	0.06
Average value		−1.77%	0.015

In **Table 1**, compared with HM16.16, the algorithm showed an average decrease of 1.77% in BD rate and an average increase of 0.015% in WS-PSNR. Among them, the encoding performance of PoleVault video test sequence, Landing2 video test sequence, Harbor video test sequence, and Trolley video test sequence has been significantly improved. This is because among four sequences mentioned above, the texture complexity of video content in bipolar region is lower and the bipolar region's weight is lower. QP becomes larger after positive compensation, and quantization has less impact on the evaluation of reconstructed video quality. The weight of equatorial region is relatively high, and QP becomes smaller after compensation, resulting in more detailed quantization and a gain in reconstructed video quality. The BD rate of four test sequences, AerialCity, Landing2, ChairliftRide, and SkateboardInLot, were compared with the original coding framework HM16.16 in **Figure 9**.

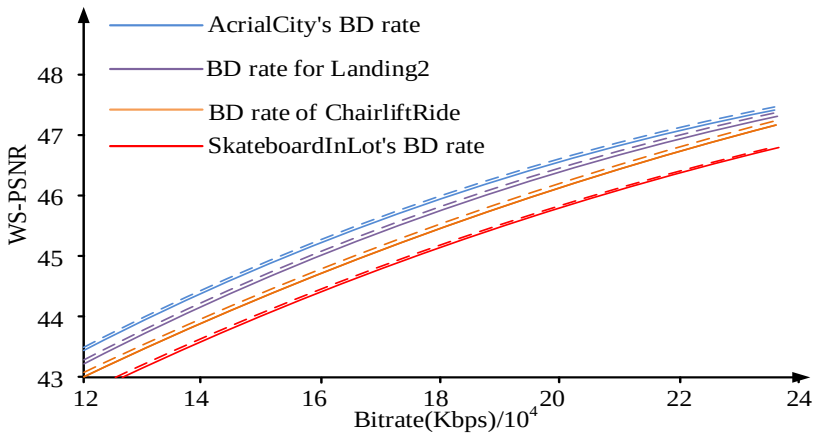

Figure 9. Comparison diagram between different test sequences and the original coding framework.

In **Figure 9**, BD rate curves are higher than those of HM16.16, indicating that this proposed method can improve the quality of reconstructed images. This is because this algorithm fully considers the characteristics of 360-degree VR image ERP projection format. That is, in 360-degree VR images, the polar region is too extended and attention is not high. On this basis, WS-PSNR weighting is used (WS-PSNR weighting w and ERP image distortion distribution have good consistency, that is, the polar region is relatively low, and the equatorial region is relatively high). The weighted values for each CTU were calculated based on their units. And adaptive

compensation was applied to each CTU weighting value, making the equatorial region of each CTU weighting value (which has a significant impact on the reconstructed image) more finely processed. For equatorial regions that are less sensitive (with a small impact on reconstructed images), they are processed more roughly to improve image reconstruction quality and coding performance. **Table 2** shows the dimensional characteristics and effectiveness results of badminton sports.

Table 2. Dimensional characteristics and effectiveness results of badminton games.

Serial number	Dimension	Factor loading degree	CR	AVE
1	Tutorial/practice	0.868	0.915	0.730
2	Interaction	0.815		
3	Feedback	0.871		
4	Identity	0.863		
5	Immersion	0.946	0.957	0.848
6	Happy frustration	0.916		
7	Sense of control	0.901		
8	Increase in difficulty	0.920		
9	Rule	0.894	0.921	0.795
10	Learning content	0.892		
11	Learning objective	0.889		
12	Teaching effect	0.806	0.948	0.701
13	Reading efficiency	0.769		
14	Communication	0.808		

According to **Table 2**, in badminton, the factor load of each sport dimension is greater than 0.75. Therefore, all factor loads can reflect the movement information. The AVE and CR of each scale and its sub dimensions were greater than 0.70 and 0.90, respectively. This indicates that each scale has good internal quality and good internal consistency between issues, and has passed the validity test. Using SEGR scale and Likert five-point method, each sports project's execution effectiveness was evaluated, and its average score was obtained. **Figure 10** shows the performance impact of different hidden layer nodes on the model, as well as the impact of training set size on model's performance.

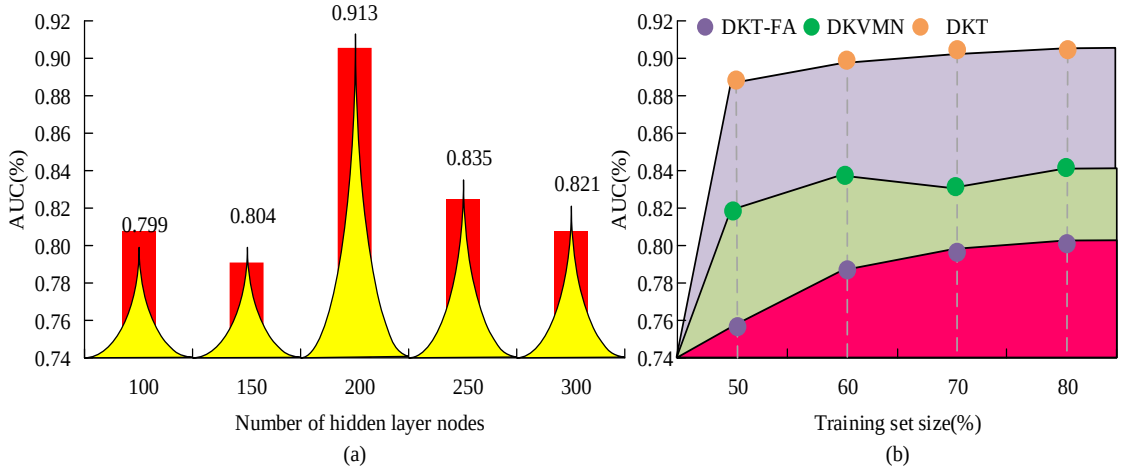

Figure 10. Influence of node number of different hiding layers and training set size on the model. **(a)** AUC values of node number models of different hidden layers; **(b)** Effect of training set size on model effect.

In **Figure 10a**, as hidden layer nodes increase, research model's AUC value also continuously increases. When node reaches 200, its AUC value is the highest, at 91.3%. After more than 200 nodes, this model may produce overfitting due to too many nodes. According to **Figure 10b**, DKT-FA model outperforms other two models in various sample sets. And when training sample size changes, the performance difference between two algorithms is not significant. This proves model's stability and practicality. **Figure 11** shows the satisfaction comparison results between the research model and classroom teaching, as well as the time spent on badminton performance of each group.

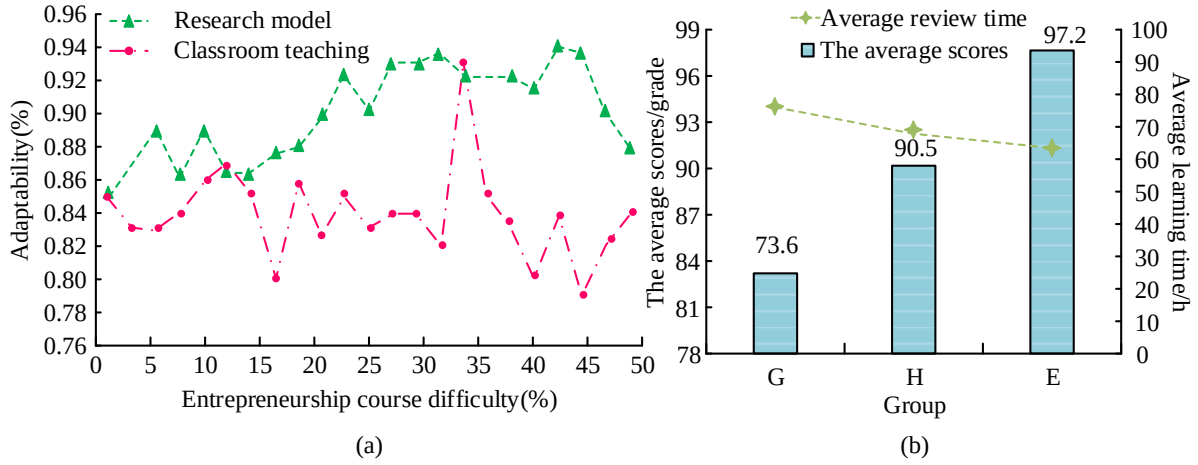

Figure 11. Comparison between the accuracy of recommendation algorithms and learners' badminton performance. **(a)** Comparison of student adaptability of two teaching methods; **(b)** Learner's entrepreneurial education achievement and learning time.

In **Figure 11**, Group A was treated as a blank group without participating in the game. The students in Group B who only conducted classroom teaching served as the control group. The students in Group C who conducted classroom teaching and participated in games served as the experimental group. In **Figure 11a**, badminton course difficulty gradually increases, and students' adaptation level also undergoes

significant changes. The controllability of badminton games among students is showing an increasing trend, and it is higher than classroom teaching. The difference in adaptability between it and the class students is also gradually increasing, ranging from 2% to 11%. From **Figure 11b**, it can effectively improve students' badminton performance and save learning time. It has been confirmed that the designed badminton game mode is effective and can meet the individualized teaching needs of students in badminton. In order to determine whether the game is suitable for both male and female students, gender was used as the independent variable in the experiment and an independent sample t-test was used. The learner experience differences of different genders in various game dimensions were tested. **Table 3** shows the statistical results of t-test for independent samples.

Table 3. Statistical results of independent sample *t*-test.

Dimension	Gender	*N*	Mean value	Standard deviation	*T*	*P*
Tutorial/practice	Male	26	4.00	0.75	−0.666	0.514
	Female	23	4.15	0.49		
Interaction	Male	26	4.34	0.49	−2.260	0.030
	Female	23	4.69	0.45		
Feedback	Male	26	4.76	0.45	−0.075	0.947
	Female	23	4.78	0.58		
Identity	Male	26	3.49	0.50	−0.944	0.357
	Female	23	3.65	0.58		
Immersion	Male	26	3.48	0.59	−0.110	0.928
	Female	23	3.47	0.73		
Happy frustration	Male	26	4.13	0.75	−0.505	0.618
	Female	23	4.28	0.69		
Sense of control	Male	26	4.23	0.57	−1.557	0.137
	Female	23	4.48	0.77		
Increase in difficulty	Male	26	4.05	0.64	−2.415	0.025
	Female	23	4.56	0.44		
Rule	Male	26	2.98	0.84	−1.906	0.065
	Female	23	3.35	0.54		
Learning content	Male	26	4.47	0.52	−0.885	0.372
	Female	23	4.62	0.55		
Learning objective	Male	26	3.95	0.66	−1.514	0.137
	Female	23	4.25	0.68		
Teaching effect	Male	26	3.65	0.73	−2.307	0.036
	Female	23	4.07	0.74		
Motion effects	Male	26	3.75	0.69	−0.284	0.787
	Female	23	3.77	0.63		
Communication	Male	26	1.45	0.55	−0.385	0.706
	Female	23	1.47	0.56		

In **Table 3**, on 17 sport game dimensions, the average scores of girls are slightly higher than those of boys. In terms of improving difficulty, the average score difference between male and female students is the largest, at 0.45. Girls have a significantly higher level of identification with this dimension than boys. In the feedback dimension, both girls and boys scored an average of over 4.3 points. The results indicate that from this perspective, the game feedback information has strong applicability, and the game design goals can be achieved. This study collected the mastery of various knowledge points through the entrepreneurial sports game from learners, and obtained empirical research results on the performance of each research model in **Figure 12**.

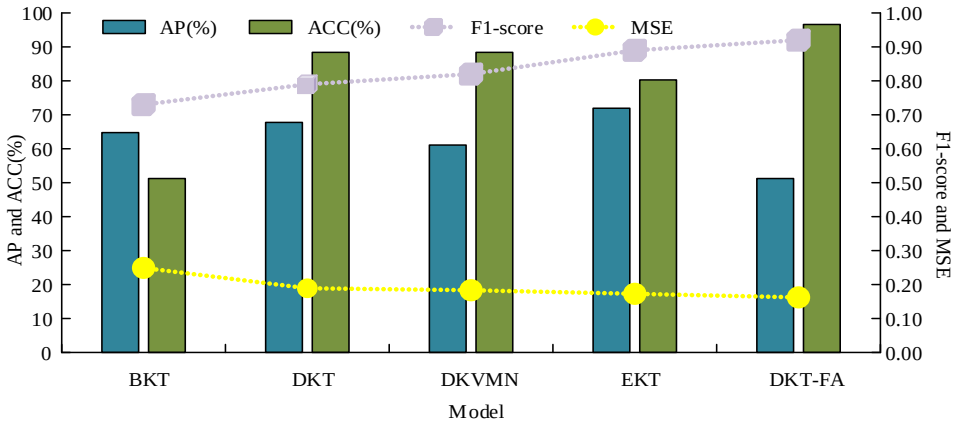

Figure 12. Empirical research results.

In **Figure 12**, the prediction accuracy of motion game model constructed by combining adaptive learning and knowledge tracking reached 94.8%, which is 5.0% – 26.1% higher than other models. But among all tested models, its mean squared error is the smallest, only 0.1623. In addition, among all tested modes, this mode scored the highest, reaching 0.92.

5. Discussion

The badminton sports game model based on LSTM network constructed in the study achieved 94.8% in terms of prediction accuracy, a result that was improved by 5.0%–26.1% over other models. The features used in the model, including the characteristics of students' question-answering behaviours and question difficulty coefficients, reflect students' knowledge acquisition and learning paths, and are crucial to the model's predictive ability. In particular, the study enhanced the model's learning and utilisation of these features through feature embedding and attention mechanisms, thereby improving the model's prediction accuracy.

In contrast to traditional MC and VR research, the study combines these two technologies with adaptive learning and DKT to provide a new solution for teaching sports games. This technology integration showed significant advantages in improving teaching effectiveness. The model exceeded most models reported in the existing literature in terms of predictive accuracy. This enhancement may be attributed to the study's optimisation of feature selection and model structure, as well as its focus on data pre-processing. In addition, the study considered the impact of gender differences

on learning outcomes, an area that is less well addressed in the existing literature. Findings from the study indicated that girls scored slightly higher than boys on average on some dimensions, which may be related to gender differences in motor skill learning. This is in line with the findings of Maurer [22]. In summary, the study shows innovation and practicality in terms of technology integration, model performance enhancement, and empirical study of teaching effectiveness compared to the existing literature. Future research can further explore how these factors affect teaching effectiveness and how these findings can be applied to a wider range of educational scenarios.

6. Conclusion

From the perspective of sports, an adaptive learning model based on sports theory was established with ordinary university students as the research object. On this basis, BN is used to model learners' knowledge, and adaptive learning and improved deep tracking methods are combined to establish corresponding motion game models. By combining adaptive learning and knowledge tracking, the accuracy of the constructed motion game model reached 94.8%, which increased by 5.0%–26.1% compared to other models. But among all the tested models, its Mean squared error is the smallest, only 0.1623. In addition, among all tested modes, this mode scored the highest, reaching 0.92. As badminton courses difficulty gradually increases, students' adaptability also undergoes significant changes. However, overall, students' controllability of badminton games shows an increasing trend and is higher than that of classroom teaching. And the adaptability difference between it and class students is gradually increasing, ranging from 2% to 11%. Compared with HM16.16, the average BD rate of the algorithm decreased by 1.77%, and the average WS-PSNR increased by 0.015a. Among them, the encoding performance of PoleVault video test sequence, Landing2 video test sequence, Harbor video test sequence, and Trolley video test sequence has been greatly improved. Although there have been good research results, due to the narrow selection of domain knowledge corresponding to sports games, there are still some shortcomings in communication, immersion, and other dimensions. This is also a need for further improvement in future research.

Author contributions: Conceptualization, QT and JT; methodology, QT; software, JT; validation, QT and JT; formal analysis, QT; investigation, JT; resources, QT; data curation, JT; writing—original draft preparation, QT; writing—review and editing, JT; visualization, QT; supervision, JT; project administration, QT; funding acquisition, JT. All authors have read and agreed to the published version of the manuscript.

Ethical approval: Not applicable.

Conflict of interest: The authors declare no conflict of interest.

Abbreviations

Bayesian Network	BN
Motion Capture	MC
Virtual Reality	VR

Long Short-Term Memory	LSTM
Adaptive Quantization Parameter Algorithm	AQP
Quantization Parameter	QP
High Efficiency Video Coding	HEVC
Rate Distortion Optimization Quantization	RDOQ
Transmission Unit	TU
Quality Policy	QP
Peak Signal to Noise Ratio	PSNR
Weighted To Spherically Uniform PSNR	WS-PSNR
Enterprise Resource Planning	ERP
Deep Knowledge Tracing	DKT
Attention Mechanism	AM

References

1. Saleem N, Gao J, Khattak MI, Rauf HT, Kadry S, ShafiM, DeepResGRU: Residual gated recurrent neural network-augmented Kalman filtering for speech enhancement and recognition. KBS. 2021; 238:107914.1-107914.13.
2. Wang X, Zhang P, Gao W, Li Y, Wang Y, Pang H. Misfire Detection Using Crank Speed and Long Short-Term Memory Recurrent Neural Network. Energies. 2022;15(1):300-310.
3. Lee C, Ng K, Chen C, Lau H, Chung S, Tsoi T. American Sign Language Recognition and Training Method with Recurrent Neural Network. Expert Syst. Appl. 2021;167(1):114403.1-114403.14.
4. Mekruksavanich S, Jitpattanakul A. Deep Convolutional Neural Network with RNNs for Complex Activity Recognition Using Wrist-Worn Wearable Sensor Data. Electronics. 2021;10(14):1685-1685.
5. Simonetto M, Arena S, Peron M.A methodological framework to integrate motion capture system and virtual reality for assembly system 4.0 workplace design Safety Sci. 2022;146(2):105561.1-105561.14.
6. Maciejewski M, Piszczek M, Pomianek M, Norbert P. Design and Evaluation of A Steamvr Tracker For Training Applications - Simulations and Measurements. MetrolMeas Syst. 2020;27(4):601-614.
7. Zhuo N, Sharma A. 3D Visual Motion Amplitude Tracking Simulation Method for Sports. Recent AdvElectr El. 2021;14(7):718-726.
8. Kadirvelu B, Gavriel C, Nageshwaran S, Chan J P, Nethisinghe S, Athanasopoulos S, Ricotti V, Voit T, Giunti P, Festenstein R, Faisal AA. A wearable motion capture suit and machine learning predict disease progression in Friedreich's ataxia. Nature medicine. 2023, 29(1): 86-94.
9. Obukhov A D, Volkov A A, Vekhteva N A, Patutin K I, Nazarova A O, Dedov D L. The method of forming a digital shadow of the human movement process based on the combination of motion capture systems. Informatics and automation. 2023, 22(1): 168-189.
10. Ning B, Na L. Deep Spatial/temporal-level feature engineering for Tennis-based action recognition, FCGS. 2021;152(6):188-193.
11. Cai J, Hu J, Tang X, Hung T,Tan Y. Deep Historical Long Short-Term Memorys for Action Recognition. Neurocomputing. 2020;407(24):428-438.
12. Hu X, Bao X, Xie S, Wei G. Unsupervised motion capture data segmentation based on topic model. ComputAnimatVirt W. 2020;32(1):2-12.
13. Buchman-Pearle J, Acker S. Estimating Soft Tissue Artifact of the Thigh in High Knee Flexion Tasks Using Optical Motion Capture: Implications for Marker Cluster Placement. J Biomech. 2021;127(8):5-17.
14. Uslu T, Gezgin E, Zbek S, Guzin D, Etin L. Utilization of Low Cost Motion Capture Cameras for Virtual Navigation Procedures: Performance Evaluation for Surgical Navigation. Measurement. 2021;181(5):109624.
15. Vlahek D, Stoi T, Golob T, Milos K, Teja L, Matjaz V, Domen M. Method for estimating tensiomyography parameters from motion capture data. SAI. 2021;45(2):213-222.
16. Min H, Chen Z, Fang B, Xia Z, Song Y, Wang Z, Zhou Q, Sun F, Liu C. Cross-Individual Gesture Recognition Based on Long Short-Term Memory Networks. Sci Program. 2021; 5:6680417.1-6680417.11.

17. Xia Z, Xing J, Wang C, Li X. Gesture Recognition Algorithm of Human Motion Target Based on Deep Neural Network. MIS. 2021; 2:2621691.1-2621691.12.
18. Venkatnarayan R, Mahmood S, Shahzad M. WiFi based Multi-User Gesture Recognition. IEEE Trans Mob Comput. 2021;20(3):1242-1256.
19. Wonderen EV, Unsworth S. Testing the validity of the Cross-Linguistic Lexical Task as a measure of language proficiency in bilingual children. J Child Lang. 2021;48(6):1101-1125.
20. Lee W, Chun J, Lee Y, Park K, Park S. Contrastive learning for knowledge tracing//Proceedings of the ACM Web Conference. 2022;2330-2338.
21. Mohammed S A, Al-Haddad L A, Alawee W H, Dhahad H A, Jaber A A, Al-Haddad S A. Forecasting the productivity of a solar distiller enhanced with an inclined absorber plate using stochastic gradient descent in artificial neural networks. Multiscale and Multidisciplinary Modeling, Experiments and Design. 2024, 7(3): 1819-1829.
22. Maurer M N. Correlates of early handwriting: Differential patterns for girls and boys. Early Education and Development. 2024, 35(4): 843-858.